

Comparative Analysis between Statistical and Artificial Intelligence Models in Business Failure Prediction

José Pereira¹, Mário Basto² & Amélia Ferreira-da-Silva³

¹ Polytechnic Institute of Cávado and Ave, School of Management, Portugal

² Polytechnic Institute of Cávado and Ave, School of Technology, Portugal

³ Polytechnic Institute of Porto, Institute of Accounting and Administration of Porto, Portugal

Correspondence: Amélia Ferreira-da-Silva, Institute of Accounting and Administration of Porto, S. Mamede de Infesta, 4465-004 Matosinhos, Portugal. Tel: 351-229-050-000. E-mail: ameliafs@gmail.com

Received: December 31, 2013

Accepted: January 20, 2014

Online Published: February 25, 2014

doi:10.5539/jms.v4n1p114

URL: <http://dx.doi.org/10.5539/jms.v4n1p114>

Abstract

A growing number of predicting corporate failure models has emerged since 60s. Economic and social consequences of business failure can be dramatic, thus it is not surprise that the issue has been of growing interest in academic research as well as in business context.

The main purpose of this study is to compare the predictive ability of five developed models based on three statistical techniques (Discriminant Analysis, *Logit* and *Probit*) and two models based on Artificial Intelligence (Neural Networks and Rough Sets). The five models were employed to a dataset of 420 non-bankrupt firms and 125 bankrupt firms belonging to the textile and clothing industry, over the period 2003–09.

Results show that all the models performed well, with an overall correct classification level higher than 90%, and a type II error always less than 2%. The type I error increases as we move away from the year prior to failure.

Our models contribute to the discussion of corporate financial distress causes. Moreover it can be used to assist decisions of creditors, investors and auditors. Additionally, this research can be of great contribution to devisers of national economic policies that aim to reduce industrial unemployment.

Keywords: bankruptcy prediction, financial distress, statistical models, artificial intelligence

1. Introduction

Bankruptcy, financial distress, insolvency, and so on, all these concepts are closely connect, almost synonymous. In general, they all mean business failure, i.e., company's life time has finished or it is under threat to.

How can we anticipate that? How can you prevent that? What are the main causes of business failure? These are some of the questions that have been asked by researchers, mostly since Beaver (1966) and Altman's (1968) pioneering works. However, the causes of business failure are too many, too complex and very interconnect.

Various methods have been used to construct prediction models. Research on the issue can be classified according to the country, industrial sector and period of data, as well as the financial ratios and models or methods employed (Dimitras, Zanakis, & Zopounidis, 1996).

Last four decades were very rich in publications that brought theoretical and methodological improvements to the debate of this issue. However, the new predicting models proposed in literature didn't always mean better results. That is probably the reason why researchers have looked for alternative techniques and tools aiming at incorporating a higher level of accuracy and usefulness.

In this paper, we will compare the predictive ability of five predicting corporate failure models: three models based on statistical techniques (Discriminant Analysis, *Logit* and *Probit*) and two models based on Artificial Intelligence (Neural Networks and Rough Sets). And because industry-specific characteristics can contribute to a firm's financial distress, we restrict our analyses to textile and clothing industry.

Human beings are the unique who wear clothes. However, clothing and textiles industrial is much more than just clothes or decorative textiles. In some sense, clothing and textiles industrial history reflects civilizations history.

Industrial revolution brought significant improvements in methods of textile production and working

organization. Since then, textile engineer on materials, equipment, methods, design, and so on, have been continually evolving. Nevertheless, the production of clothes continued being hand-made until sewing machines appeared (19th) and it started being widespread in the factories' industry.

Nowadays, the textile and clothing industry covers a wide variety of activities. According European Commission (2013), the sector is an important part of the European manufacturing industry. It plays a crucial role on the economy and social well-being in numerous regions of the EU-27.

In Portugal, mostly in the last four decades, clothing and textiles have been an important industry for national economy. Following the latest structural data available (Associação Têxtil e Vestuário de Portugal [ATVP], 2013), in 2012 there were approximately 6.000 companies employing 150.000 workers and generated a turnover of €4.910 million. The textile and clothing sector accounts for 9% of total Portuguese exportations and 20% of employment in transformation industry. But, ten years before, the numbers were quite different. Indeed, in 2002 the industry generated a turnover of €8.198 million and employed 243.264 workers. Besides, according to a study from the Bank of Portugal (2012), the rate of turbulence in the textile and clothing industry has been at around 15%, with the mortality rate recurrently exceeding the birth rate. Thus, the number of companies active in this sector has declined, falling by 28% since 2004. Moreover, by comparison with the group of non-financial companies over the last ten years, the sector of Textiles and Clothing has presented more negative natural balances, justifying the reduction of its weight in the whole Portuguese economy.

The world economy is undergoing a profound structural change that is testing companies' resistance to competitiveness. Definitely, "globalization has altered the competitive dynamics of nations, firms, and industries" Gereffi, G. (1999). The introduction of textile and clothing sector to the general rules of the General Agreement on Tariffs and Trade since 1 January, 2005, created great pressure on Portuguese companies. Without a doubt, liberalization has been controversial because both textiles and clothing contribute to employment in developed countries, particularly in regions where alternative jobs may be difficult to find, (Nordås, 2004) and Portugal is certainly one of those cases.

In such circumstances, it is worth to ask: what are the main differences between Portuguese textile and clothing successful and unsuccessful companies? Can we predict their business failure? Can we prevent their business failure?

This article will address these questions by doing a comparative analyze of five predictive business failure models using a dataset of 420 non-bankrupt firms and 125 bankrupt, over the period 2003-09.

2. Brief Literature Review

There are several undesirable consequences of business failures. Its economic and social cost can be significant. So, it is quite natural that this issue has occupied a significant part of researcher's agenda. In spite of recent growing interest on non-financial attributes in explaining business failures, traditionally investigation on this issue has been focused on financial attributes. In most of the works statistical or artificial intelligence techniques were applied to the accountancy data of the companies, aiming at obtaining prediction models that would indicate whether the company would or would not reach a bankruptcy situation in the future (Beaver, 1966; Altman, 1968; Martin, 1977; Tan & Kiang, 1992).

In a study on corporate bankruptcy prediction, one of the aspects we immediately need to clarify is the concept of bankruptcy we shall use. In specialized literature the term has been used in different ways by different authors: legal bankruptcy, insolvency, inability to do payments or continued losses. As we lack a general theory on corporate bankruptcy, there is also no unique definition for this concept. This is an important limitation, since the sample's selection, both in terms of firms that have and have not "bankrupt", depends on the definition of corporate bankruptcy used.

Throughout the last four decades several techniques were used to design models regarding this issue. The Altman model is the classical model on prediction of corporate bankruptcy most referred to in literature. It was developed in the end of the 70's using the discriminant analysis. The purpose of this technique was to obtain an indicator or «Z» score (variable dependent on a function) that was the result of the linear combination of several independent variables (ratios or financial indicators).

After the publication of Altman's *Z-Score* in 1968, most studies published in the decade that followed also used the discriminant analysis (Blum, 1974; Elam, 1975; Altman et al., 1977; Moyer, 1977; Norton & Smith, 1979).

The emergence of critiques emphasizing the limitations of this theory may have influenced researchers to try new techniques, namely *logit* and *probit*.

With the application of *logit* models to corporate bankruptcy prediction, it is possible to estimate the probability that a certain event will happen, as well as the probability of failure or corporate bankruptcy, considering the values of certain indicators of the company (Ohlson, 1980; Keasey & Watson, 1987).

The *probit* model is associated to the cumulative function of normal probability. Although this model is not as popular as the *logit* one in this area of research, there are several studies that have used this methodology (Zmijewski, 1984; Lennox, 1999) with similar outcomes as with other techniques.

The evolution of ICT has created adequate conditions for the development and application of other techniques that, despite their limitations, do not demand certain conditions, unlike statistical techniques.

Among artificial intelligence techniques, the most used ones have been neural networks and the induction of rules and decision trees. During the last few years, studies that use the theory of Rough Sets have emerged.

Several neural networks models have been used throughout the last decade in studies related to corporate solvency prediction. Among them we can highlight Bell et al. (1990) and Koh & Tan (1999), using the multilayer perceptron model; Coats & Fant (1993) and Lacher et al. (1995), using the *Cascor* method (cascade correlation) as a learning algorithm; and Serrano & Martin (1993), using a multilayer *perceptron* net and Kohonen's self-organizing maps.

Zdzislaw Pawlak originally created the theory of Rough Sets in the beginning of the 80's, but it was only during the following decade that it was applied to insolvency or corporate failure. The first models appeared during the 90's (Slowinski & Zopounidis, 1995; Dimitras et al., 1999).

3. Empirical Analysis

3.1 Definition of Bankruptcy Used

In the studies on corporate bankruptcy prediction, the companies are usually divided into two categories: bankrupt and non-bankrupt. In our study we identified a company as bankrupt in case of an insolvency process or if the courts had declared the company insolvent.

3.2 Variable Selection

The lack of a theoretical basis, that explains corporate bankruptcy and that serves as a guideline in the selection process of the explaining guidelines, has allowed the use of multiple criteria when choosing the explaining guidelines. As in previous studies, the selection of the independent variables—in this study, economical-financial ratios—was based on their level of use and on the level of meaning obtained in several studies published in specialized literature.

The list of the 49 selected ratios that were created with the information obtained in the companies' Balance Sheet and income statement, and which make up our sample, is in Table 1.

Table 1. List of selected ratios

	Ratio
X1	Total assets / Total liabilities
X2	(Current assets - Inventory) / Current liabilities
X3	(Current assets - Current liabilities) / Total liabilities
X4	(Current assets - Current Liabilities) / Sales
X5	Current assets / Total assets
X6	Current assets / Current liabilities
X7	Current assets / Total Liabilities
X8	Fixed assets / Current Assets
X9	Depreciation expense / Operating gains
X10	Equity / Total assets
X11	Equity / Net-Fixed assets
X12	Equity / Total liabilities
X13	Cash-flow / Current liabilities
X14	Cash-flow / Total liabilities
X15	Financing charge / Operating gains
X16	Financing charge / Operating profit (or loss)
X17	Financing charge / Total bank loans

X18	Cash / Total assets
X19	Cash / Current liabilities
X20	(Cash + Bills receivable)/ Current liabilities
X21	Bills payable / Total assets
X22	Inventory / Total assets
X23	Inventory / Sales
X24	Working capital / Total assets
X25	Staff costs/ Fixed assets
X26	Staff costs / Operating gains
X27	Net-Fixed assets / Total assets
X28	Total Liabilities / Equity
X29	Operating gains / Current assets
X30	Operating gains / Total assets
X31	Operating gains / Operating costs
X32	Operating gains / Fixed assets
X33	(Net profit before tax + Depreciation expense + Provisions)/ Financing charge
X34	Extraordinary result / Net profit
X35	Net profit /Total assets
X36	Net profit / Equity
X37	Net profit / Total liabilities
X38	Net profit / Operating gains
X39	Net profit/ Sales
X40	Operating profit (or loss)/ Total assets
X41	Operating profit (or loss)/ Financing charge
X42	Operating profit (or loss)/ Operating gains
X43	Operating profit (or loss)/ Sales
X44	Sales / Total assets
X45	Sales / Current assets
X46	Sales / Cash
X47	Sales / Inventory
X48	(Net profit before tax + Financing charge) / Sales
X49	Net profit before tax /(Net profit before tax + Financing charge)

3.3 Sample Selection Procedure

The sample was created using information obtained in the SABI database (Financial company information for companies in Spain and Portugal) and by Judicial Managers. In order to be part of the sample, the companies had to present 3 consecutive financial years, between 2003 and 2009. A sample of 420 non-bankrupt companies and 125 bankrupt companies was created, these belonging to the textile and clothing industries. This sample was then subdivided in two more samples, a test one, made up of 80 randomly selected bankrupt companies and of 160 healthy or non-bankrupt companies, with data corresponding to the same economic exercises and of approximately the same size, and another validation sample, with the remaining companies (260 non-bankrupt companies and 45 bankrupt companies).

3.4 Techniques Used and Results Obtained

In this study we have designed five predictive models using three statistical techniques (Discriminant Analysis, *Logit* and *Probit*) and two artificial intelligence techniques (Neural Networks and Rough Sets).

The statistical software used was SPSS 18.0 for Windows. To design the neural network model, we used SAS Enterprise Miner software, and for Rough Sets we used non-commercial software created by Diaz (2002).

We will now present the main data of the designed models, as well as the results obtained.

Discriminant Analysis

In Table 2 we can see the significance of the obtained discriminant function.

Table 2. Significance of the discriminant function

Discriminant function	Wilks' Lambda	Chi-square	p-value
$Z = -12.572 + 0.870 * X_{10} + 13.234 * X_{31} + 2.529 * X_{37} - 2.612 * X_{48} - 0.222 * X_{49}$	0.204	372.684	0.000

The centroids (average overall index of each group) are -2.799 for the group of bankrupt firms, and 1.382 for the group of healthy firms. An analysis of each company's scores allowed us to conclude that the point that allowed for the least error was -0.64. To classify each company from the test sample, as well as the companies in the validation sample, we used the mentioned value as a cut point. The results obtained are in Table 6.

Logit

The estimate of the *logit* model was also based on data from the test sample referring to the year prior to the failure, made up of 80 bankrupt companies and 160 non-bankrupt companies, using the *stepwise forward wald* procedure. The ratios included in the selected model are X_{14} (*Cash-Flow/Total liabilities*) and X_{35} (*Net profit/Total assets*) that can be seen in Table 3.

Table 3. Estimated parameters and significance level

Parameters	Coefficients	Wald Statistics	p-value
X_{14}	13.021	2.795	0.019
X_{35}	46.913	16.800	0.000
Constante	2.829	12.367	0.000

To perform the cut point, we used the value of 0.5.

Probit

To design the *probit* model we have also used SPSS 18.0 for Windows. During the variable selection we used a manual procedure for variable input and output, until we reached a globally significant model.

The variables selected by the model were: X_{14} (*Cash-Flow / Total liabilities*) and X_{31} (*Operating gains / Operating costs*). Their estimated coefficients can be found in table 4.

Table 4. Indicators estimated by the *probit* model

Parameters	Coefficients	Standard Error	p-value
X_{14}	6.447	2.653	0.015
X_{31}	18.063	4.916	0.000
Constante	-16.699	4.690	0.000

To verify the quality of the estimated values' adjustment we used the Pearson test. This is a Chi-squared test for the null hypothesis stating that there are no differences between the observed values and the predicted values. The obtained Chi-square value is 19.693 with a value of $p=1$, thus allowing us not to reject the null hypothesis. To apply the *probit* model to a specific situation (a company), we use the following equation: $Y=N(Z)$, where:

Y = probability the company will continue working,

$N(.)$ =function of the normal cumulative distribution function and

Z = the theoretical index or the value of Z in the normal distribution table.

The estimated predicative model presents the following coefficients for the Z index:

$$Z = -16.699 + 6.447 * X_{14} + 18.063 * X_{31}$$

For a certain company the value of Z is calculated by multiplying the ratio values by the corresponding coefficients of the previous function. In order to exemplify the functioning of this model, let's take as an example two companies' data, which had different outcomes.

Company A (non-bankrupt) presented a value of 0.231421971 for the X_{14} ratio and a value of 0.927256383 for the X_{31} ratio. Company B (bankrupt) presented a value of 0.068887624 for the X_{14} ratio and a value of

0.904680241 for the X31 ratio. When replacing the values of each ratio in the previous equation shown, the value of Z for each company is the following:

Company A: $Z = 1.54$

Company B: $Z = -0.80$

The value of Z of 1.542 is converted into an estimated probability using the normal cumulative distribution table, thus obtaining an estimated probability of about 94% that company A will continue working. Company B presents an estimated probability of continuing to be active of around 21%, thus being classified as bankrupt, as the cut point used is 0.5.

Neural Networks

The automatic learning technique we are going to apply consists of the development of a neural network, whose goal is to identify companies that present a high probability of failure.

We started the study with all the variables on the input layer, having the model correctly classified all the healthy companies. However, *Type Error 1* was too high (42.5%). To reduce the number of variables in the input layer, we used one of the criteria available in the program, in this case *Chi-square*.

The network that obtained the best results had a *Multilayer Perceptron* architecture with *Back-propagation* learning and a hidden layer. The input layer had 13 variables (neurons), the intermediate or hidden layer had 2 variables and the output layer had 1. A "bias" tendency neuron was introduced, connected to all the neurons of the hidden layer and to the output layer neuron, as we can see in Figure 1.

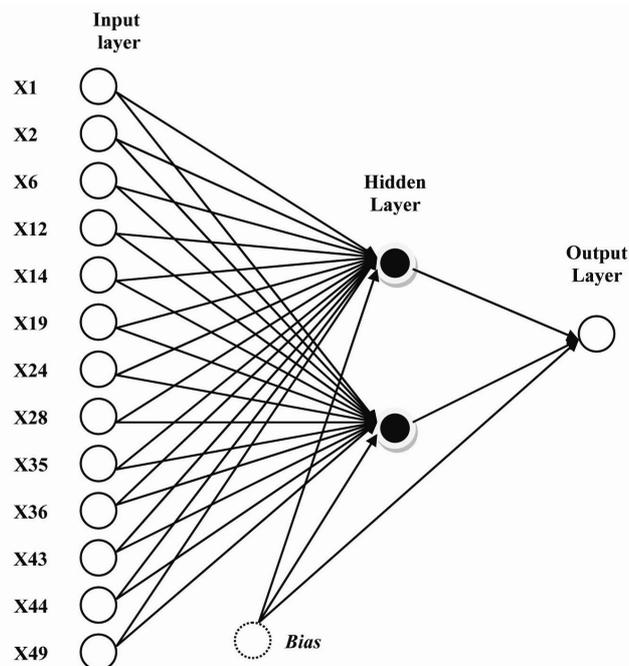


Figure 1. Architecture for the three-layer back-propagation neural network

Rough Sets

In order to perform an empirical analysis using the Rough Sets Theory, we used software developed by Díaz (2002), as mentioned earlier. One of the most important operations during the pre-processing phase is to perform the discretization of attribute values (ratios). The discretization was divided into 5 categories with *equal frequency binning*. The results obtained for the ratio values of the companies can be seen in Table 5 (Note 1).

Table 5. Discretization intervals of the attributes

Atributte	0	1	2	3	4
X1]*, 1.18248[[1.18248, 1.36442[[1.36442, 1.53683[[1.53683, 1.85887[[1.85887, *[
X2]*, 0.47699[[0.47699, 0.69996[[0.69996, 0.92569[[0.92569, 1.34922[[1.34922, *[
X3]*, -0.09372[[-0.09372, 0.10828[[0.10828, 0.29041[[0.29041, 0.57612[[0.57612, *[
X4]*, -0.05932[[-0.05932, 0.05292[[0.05292, 0.13964[[0.13964, 0.30047[[0.30047, *[
X5]*, 0.41664[[0.41664, 0.57106[[0.57106, 0.68512[[0.68512, 0.82800[[0.82800, *[
X6]*, 0.86943[[0.86943, 1.15045[[1.15045, 1.42227[[1.42227, 1.99607[[1.99607, *[
X7]*, 0.58187[[0.58187, 0.80693[[0.80693, 1.01154[[1.01154, 1.28637[[1.28637, *[
X8]*, 0.20818[[0.20818, 0.46092[[0.46092, 0.75194[[0.75194, 1.40584[[1.40584, *[
X9]*, 0.02052[[0.02052, 0.03634[[0.03634, 0.05882[[0.05882, 0.10230[[0.10230, *[
X10]*, 0.15333[[0.15333, 0.26622[[0.26622, 0.34815[[0.34815, 0.46188[[0.46188, *[
(...)					

The algorithm for the attribute selection is based on the MDL (*Minimum Description Length*) principle. This principle eliminates the variable on each step of the algorithm, which simultaneously maximizes the gains, due to the decrease in complexity of the resulting model and minimizes the precision loss that it has (Díaz, 2002).

The inference method used had, as selection criteria, the rule satisfied by a new observation, aiming at achieving the highest number of observations within the training group, supporting it, that is to say, assuring the purpose is met.

During the training phase the maximum permissible error was 2.5% (beta = 0.025), yielding 38 reducts based on the following set of P attributes:

$$P = \{X3, X10, X22, X27, X32, X41, X42, X48, X49 \}$$

The reduct number 32 corresponds to the one that presented the best overall performance, having obtained 32 rules generated from {X3, X10, X32, X41, X48, X49}. We will now present the first two rules.

Rules generated from {3,10,32,41,48,49}

Reduct number 32. (Number of rules generated: 32)

$$r_0: [(X48 = 3)] \Rightarrow [(Decision = 1)]$$

CoverP={0,23,35,40,43,44,45,47,63,67,70,74,75,78,79,85,88,91,92,101,103,105,110,116,117,118, 119,123,125,126,127,129,133,136,141,152,154,155,156,162,167,169,176,181,183,184,187,191,193,348}

Error = 0.0200

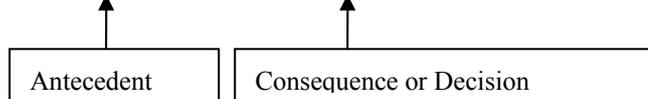
$$r_1: [(X41 = 4)] \Rightarrow [(Decision = 1)]$$

CoverP={0,1,2,5,10,11,12,14,32,33,39,41,44,53,54,57,64,69,70,73,78,89,92,93,97,100,101,104,105,113,121,123, 129,133,140,144,145,157,160,161,162,164,165,169,173,175,176,180,184,185,186, 187}

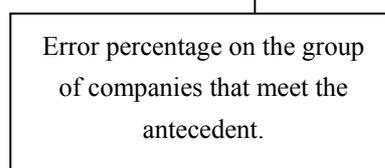
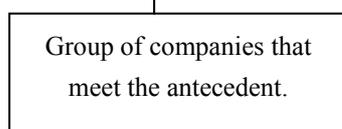
Error = 0.0000

In order to exemplify the interpretation of each rule, we will use the first one, coded r_0:

$$r_0: [(X48 = 3)] \Rightarrow [(Decision = 1)]$$



CoverP={0,23,35,40,43,44,45,47,63,67,70,74,75,78,79,85,88,91,92,101,1103,105,110, 116,117,119, } Error = 0.0200



Obtained Results

In Table 6 we can see the summary of the results obtained by the five models for each sample.

Table 6. Summary of the classification accuracy percentages of all models

Technique	Test/ Training Sample		Holdout Sample		Two years prior to bankruptcy		Three years prior to bankruptcy	
	B (Note 2)	NB (Note 3)	B	NB	B	NB	B	NB
	Discriminant Analisis	92.50	99.38	87.30	98.60	72.20	97.30	67.60
<i>Logit</i>	95.00	98.10	92.06	98.95	83.70	96.90	52.00	98.20
<i>Probit</i>	95.00	98.75	92.06	99.65	77.80	98.00	55.80	98.40
Neural Networks	97.50	99.40	68.25	100.00	61.75	98.00	51.40	97.53
Rough Sets	97.50	99.40	96.80	98.20	85.30	96.90	62.10	96.90

The findings show that all models provide a classification level of accuracy higher than 92.5 %, with the type II error always less than 2 %. We note that this type of error is always very low in any of years studied, indeed it never exceeds 3.1%. These results are superior to those obtained in similar comparative analysis studies (Frydman et al., 1985; Koh Tan, 1999; Lennox, 1999; Min & Lee, 2005; Lee et al., 2005).

The model obtained with the neural network was the only one who could correctly classify all non-failed companies in the validation sample, but it was also the one that had the worst performance in the classification of the failed firms of the same sample. The overall level of accuracy of this model reaches values close to those obtained by Tam & Kiang (1992) and Coats & Fant (1993). Even so, it continues to have the worst performance in the classification of non-failed sample firms in the second and third years prior to failure.

The overall results for *logit* are similar to those obtained by Ohlson (1980) for the first and second years, but superior to those obtained by Zavgren (1985) and Ward (1994). The results of *probit* model are superior to those obtained by Zmijewski (1984) and Gentry et al. (1987). The discriminant analysis model was the one that had the best performance in the classification of firms in the sample for the third year prior to failure. This level of accuracy is higher than that obtained by Blum (1974), Aly et al. (1992), Sung et al. (1999), and however lower than Deakin (1972), particularly the sample for the third year prior to failure. Within the models of the artificial intelligence the results obtained with Rough Sets model were generally higher than the neural network model results. In statistical models the percentages of precision were more balanced for the different samples.

4. Conclusion

Predicting corporate failure is a research topic well explored in literature. However, it is not an easy task to combine and contrast results from different studies that have been done on the topic. In the face of absence of a general *business failure theory*, researchers tend to adopt its own concept of business failure and their studies are very heterogeneous in terms of variables, models and sample selection techniques.

Nevertheless, there are significant methodological improvements that should be recognized. Moreover, the accumulation of knowledge and empirical evidence that have been achieved will surely contribute to the formulation of a general *business failure theory*.

The main goals of this work were to design corporate bankruptcy prediction models using different techniques and verify each model's predictive performance. Based on the results we can conclude that:

- With sample testing or training data, all models provided a level of accuracy in the classification of companies exceeding 90%, with *type error II* always presenting a value inferior to 2% (this type of error occurs when you consider a healthy company as a failure);
- Type II error is always very low in all the years, with percentages that never go beyond 3.1%;
- *Type error I* (classifying a bankrupt company as a healthy one) increase with time, registering rates higher than 20%, when prior data is two years and reaching percentages of 40% the corresponding data three years prior to failure.
- With the validation sample the Neural Networks model was able to correctly classify all the non-bankrupt companies, although it was the model that registered the lowest accuracy levels with the bankrupt companies.
- We found that the ratios included in each model differ according to the technique used and that none of the

indicators were simultaneously selected by all the five models.

The accuracy of each model is highly dependent on the quality of the data used as the basis for its design. It is possible that the data provided by some companies does not reflect the real financial and economical situation, namely in the group of bankrupt companies, which may have influenced the models' performance, especially regarding the second and third years prior to failure, for over 80% of the companies did not have audited accounts nor were they obliged to do so.

We believe that by using a significant sample with audited accounts and also by incorporating qualitative variables, it is possible to design models with good predictive accuracy that may be used as an additional tool in decision making.

Our conclusions have been obtained starting from a sample of firms in the Portuguese textile sector. An extension of this work would be to test its robustness in other countries (for example, Spain or Italy) and/or for other industries such as footwear, were Portugal has been very well succeeded.

Acknowledgements

We wish to thank the two anonymous referees for helpful comments and suggestions.

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., Haldeman, R. G., & Narayanan P. (1977). Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations. *Journal of Banking and Finance*, 1(1), 29–54.
- Aly, I. M., Barlow, H. A., & Jones, R. W. (1992). The Usefulness of SFAS N.º 82 (Current Cost) Information in Discriminating Business Failure: An Empirical Study. *Journal of Accounting, Auditing and Finance*, 7(2), 217–229.
- Associação Têxtil e Vestuário de Portugal. (2013). A indústria têxtil e vestuário Portuguesa. ATP. Retrieved from <http://formacao.aeportugal.pt/docs/aep-formacao-documentos/>
- Bank of Portugal. (2012). Análise setorial da indústria dos têxteis e vestuário. Estudos da Central de Balanços, Banco de Portugal, Eurosistema.
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 4(3), 71–111.
- Bell, T. B., Ribar, G. S., & Verchio, J. R. (1990). *Neural Nets versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures*. Deloitte & Touche, University of Kansas Symposium of Auditing Problems, Kansas City, 29–53.
- Blum, M. (1974). Failing Company Discriminant Analysis. *Journal of Accounting Research*, 12(1), 1–25.
- Charitou, A., Neophytou, E., & Charalambous, C. (2004). Predicting Corporate Failure: Empirical Evidence for the UK. *European Accounting Review*, 13(3), 465–497. <http://dx.doi.org/10.1080/0963818042000216811>
- Coats, P. K., & Fant, L. F. (1993). Recognizing Financial Distress Patterns Using a Neural Network Tool. *Financial Management*, 22(3), 142–155.
- Díaz Gómez, F. (2002). *Aprendizaje y Generación Automática de Conocimiento: Construcción de Redes Bayesianas mediante Rough Sets*. Tesis Doctoral, Universidad de Vigo, Vigo.
- Dimitras, A. I., Slowinski, R., Susmaga, R., & Zopounidis, C. (1999). Business Failure Prediction Using Rough Sets. *European Journal of Operational Research*, 114(2), 263–280. [http://dx.doi.org/10.1016/S0377-2217\(98\)00255-0](http://dx.doi.org/10.1016/S0377-2217(98)00255-0)
- Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, 90(3), 487–513. [http://dx.doi.org/10.1016/0377-2217\(95\)00070-4](http://dx.doi.org/10.1016/0377-2217(95)00070-4)
- Elam, R. (1975). The Effect of Lease Data on the Predictive Ability of Financial Ratios. *The Accounting Review*, 50(1), 25–43.
- European comission. (2013). Textiles and clothing industry. Enterprise and industry. Retrieved from <http://ec.europa.eu/enterprise/sectors/textiles/>
- Galindo, R. D., & Drozdowskyi, A. (2002). Un Modelo de Predicción de Fracaso Empresarial. *Partida Doble*, 134, 60–67.

- Gentry, J. A., Newbold, P., & Whitford, D. T. (1985). Classifying Bankruptcy Firms with Funds Flow Components. *Journal of Accounting Research*, 23(1), 146–160.
- Gentry, J. A., Newbold, P., & Whitford, D. T. (1987). Funds Flow Components, Financial Ratios, and Bankruptcy. *Journal of Business Finance & Accounting*, 14(4), 595–606. <http://dx.doi.org/10.1111/j.1468-5957.1987.tb00114.x>
- Gereffi, G. (1999). International trade and industrial upgrading in the Apparel commodity chain. *Journal of International Economics*, 48, 37–70.
- Grice, J. S., & Dugan, M. T. (2001). The Limitations of Bankruptcy Prediction Models: Some Cautions for the Researcher. *Review of Quantitative Finance and Accounting*, 17(2), 151–166.
- Hair, J. F., Anderson, R. E., Tatham R. L., & Black, W. C. (1998). *Multivariate Data Análisis* (5th ed.). Prentice Hall.
- Keasey, K., & Watson, R. (1987). Non-Financial Symptoms and the Prediction of Small Company Failure: A Test of Argenti's Hypotheses. *Journal of Business Finance & Accounting*, 14(3), 335–354.
- Koh, H. C., & Tan, S. S. (1999). A Neural Network Approach to the Prediction of Going Concern Status. *Accounting and Business Research*, 29(3), 211–216. <http://dx.doi.org/10.1080/00014788.1999.9729581>
- Lacher, R. C., Coats, P. K., Sharma, S. C., & Fant, L. F. (1995). A Neural Network for Classifying the Financial Health of a Firm. *European Journal of Operational Research*, 85(1), 53–65. [http://dx.doi.org/10.1016/0377-2217\(93\)E0274-2](http://dx.doi.org/10.1016/0377-2217(93)E0274-2)
- Lee, K., Booth, D., & Alam, P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Systems With Applications*, 29(1), 1–16. <http://dx.doi.org/10.1016/j.eswa.2005.01.004>
- Lennox, C. (1999). Identifying Failing Companies: A Re-evaluation of the Logit, Probit and DA Approaches. *Journal of Economics and Business*, 51, 347–364. [http://dx.doi.org/10.1016/S0148-6195\(99\)00009-0](http://dx.doi.org/10.1016/S0148-6195(99)00009-0)
- Lizarraga, F. D. (1998). Modelos de Previsión del Fracaso Empresarial: Funciona entre nuestras Empresas el Modelo de Altman de 1968? *Revista de Contabilidad*, 1(1), 137–164.
- Martin, D. (1977). Early warning of bank failure: a logit regression approach. *Journal of Banking and Finance*, 1(3), 249–276.
- 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>
- Moyer, R. C. (1977). Forecasting Financial Failure: A Re-Examination. *Financial Management*, 6(1), 11–17.
- Nordás, H. K. (2004). *The Global Textile and Clothing Industry post the Agreement on Textiles and Clothing*. World Trade Organization Publications.
- Norton, C. L., & Smith, R. E. (1979). A Comparison of General Price Level and Historical Cost Financial Statements in the Prediction of Bankruptcy. *The Accounting Review*, 54(1), 72–86.
- Ohlson, J. S. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131.
- Pawlak, Z. (1982). Rough Sets. *International Journal of Computer and Information Sciences*, 11, 341–356. [http://dx.doi.org/10.1016/S0377-2217\(96\)00382-7](http://dx.doi.org/10.1016/S0377-2217(96)00382-7)
- Rodríguez López, M., & Díaz Gómez, F. (2005). La Teoría de los “Rough Sets” y la Predicción del Fracaso Empresarial. Diseño de un Modelo para las PYMES. XIII Congreso AECA, septiembre, Oviedo.
- Serrano Cinca, C., & Martin Del Brio, B. (1993). Predicción de la Quiebra Bancaria Mediante el Empleo de Redes Neuronales Artificiales. *Revista Española de Financiación y Contabilidad*, 22(74), 153–176.
- Slowinski, R., & Zopounidis, C. (1995). Application of the Rough Set Approach to Evaluation of Bankruptcy Risk. *Intelligent Systems in Accounting, Finance and Management*, 4(1), 27–41.
- Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy. *Journal of Management Information Systems*, 16(1), 63–85.
- Tam, K., & Kiang, M. (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science*, 38(7), 926–947.

- Ward, T. J. (1994). Cash Flow Information and the Prediction of Financially Distressed Mining, Oil and Gas Firms: A Comparative Study. *Journal of Applied Business Research*, 10(3), 78–86.
- Zavgren, C. V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis. *Journal of Business Finance & Accounting*, 12(1), 19–45. <http://dx.doi.org/10.1111/j.1468-5957.1985.tb00077.x>
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22, Supplement, 59–82.

Notes

Note 1. In order to save space we only mentioned the discretization intervals for the first 10 attributes. Complete data will be made available if asked.

Note 2. Bankrupt.

Note 3. Healthy or Non-Bankrupt.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).