HOW DO DIFFERENT TIME SPANS AFFECT THE PREDICTION ACCURACY OF BUSINESS FAILURE?

X. Càmara-Turull[a], M.A. Fernández Izquierdo[b],
M.T. Sorrosal Forradellas[c]

The prediction of business failure has been widely studied by many authors. Most of the studies focused on improve the results by applying new methodologies or by using more suitable financial information. This study aims to analyze the impact of the input data timeframe on the prediction accuracy of business failure. Using an artificial neural network, the self-organizing maps (SOM), we compare the results obtained by using 9, 6 and 3 years of input data. We concluded that the 3-year case provides a better global results despite of the 6-year case presents the lowest error type I.

Keywords: *business failure, timeframe data, self-organizing maps*

JEL Classification: *G33, C80, C45*

1. INTRODUCTION

The global financial crisis that started in 2007 in the United States of America, and the consequent increase of bankruptcy firms since then in many countries, has revived the interest in predicting companies' failure.

Even though business failure has been widely tackled in financial literature, most authors focused on applying new methodologies and selecting the most suitable financial ratios and other macro-economic variables to improve prediction.

However, there is much less attention on study the time dimension in business failure prediction. This can be approached through two perspectives: (1) in terms of number of years taken as inputs for prediction; and (2) with regard to the bankruptcy time horizon, *i.e.* the number of years between the end of the analyzed period and the bankruptcy.

[a] Department of Business Management. Faculty of Business and Economics. Universitat Rovira i Virgili. Av. Universitat, 1, E-43204 Reus, Spain. Email: xavier.camara@urv.cat
[b] Department of Finance and Accounting. Faculty of Law and Economics. Universitat Jaume I, Av. Sos Baynat, s/n, E-12071 Castelló, Spain. E-mail: afernand@uji.es
[c] Department of Business Management. Faculty of Business and Economics. Universitat Rovira i Virgili. Av. Universitat, 1, E-43204 Reus, Spain. Email: mariateresa.sorrosal@urv.cat

In relation to the first approach, most studies use data from 1 to 4 years previous to the bankruptcy time horizon (Bellovary *et al.*, 2007) but other authors (Blum, 1974) demonstrated that using financial data of 8 years improves the prediction accuracy.

On the other hand, as du Jardin (2015) has pointed out, it is possible to obtain positive prediction results using a bankruptcy time horizon of only one year, but these results are getting significantly worse when a medium term prediction is used. Therefore, a bankruptcy horizon of 1 year is the common approach (Bellovary *et al.*, 2007). Nonetheless, other authors as Dwyer (1992) built methods with high success percentage using a 3-year bankruptcy horizon. In a similar way, du Jardin and Séverin (2011) also have proved that Self-Organizing Maps (SOM) can increase the period of bankruptcy prediction up to 3-years without losing predictive accuracy in business failure.

This study focuses on analyzing the first approach: the impact of the input data timeframe on the prediction accuracy of business failure when SOM models are applied in classifying healthy and bankrupt companies.

The paper is aimed to compare the error rate in identifying healthy and bankrupt companies from the Spanish chemical industry, through SOM methodology, using 3 years previous to the bankruptcy time horizon (2005-2007), by using 6 years (2002-2007) and 9 years (1999-2007). That will enable us to analyze the consequences of longer or shorter timeframe data sample, and to observe if the inclusion of former information provides added value or distorts the obtained results.

The remainder of this paper is organized as follows: Section 2 briefly reviews prior studies on business failure. In Section 3 we describe the data and methodology used. In Section 4 the empirical results are discussed. In the final section, we present the conclusions of the study.

2. LITERATURE REVIEW

Creditors, owners or managers of a company have always been concerned about business failure prediction. Back in the early part of the 20th century, some researchers tried to respond this problem by individual ratio analysis and subsequent comparison between bankrupt and healthy companies (FitzPatrick, 1932; Merwin, 1942; Smith and Winakor, 1935).

Nevertheless, was in the mid 1960's when the studies of Beaver (1966) by applying univariate techniques and Altman (1968) by using multi variant discriminant analysis awake the interest on this topic. Altman formulated two scenarios, taking data both from 1 and 2 years previously to bankruptcy.

In the 1970's and 1980's a large number of works were made applying different statistic techniques: Libby (1975) combined the principal components analysis (PCA) with discriminant analysis; Ohlson (1980) proposed a logit model to predict business failure; Zmijewski (1984) was pioneer on using a probit model by applying profitability related ratios only; or West (1985) that combined factorial analysis with a logit model. Most of papers in this period were designed using data that are measured over a period t to achieve a prediction over a period
t+1.

In 1990's there was an exponential increase in number of works and techniques studying business failure. Especially, new techniques based on artificial intelligence (AI) and data mining started to be used, such as artificial neural networks (Chen, *et al.,* 2013; Chen and du Jardin, 2009; du Jardin and Séverin, 2012; Odom and sharda, 1990; Serrano-Cinca, 1996; Séverin, 2010; Tam, 1991); support vector machines (Bose and Pal, 2006; Ding, *et al.,* 2008; Min and Lee, 2005; Shin, *et al.*, 2005; Van Gestel, *et al.*, 2010); evolution algorithms (Etemadi, *et al.*, 2009; Martens, *et al.*, 2010; Mokhatab Rafiei, *et al.*, 2011; Shin and Lee, 2002; Varetto, 1998); case-based reasoning (Borrajo, *et al.*, 2011; Li, *et al.*, 2011; Li and Sun, 2009, 2010); rough sets (Bose, 2006; Dimitras, et al., 1999; Mckee, 2000); decision trees (Frydman, *et al.*, 1985; Gepp, *et al.*, 2010; Li, *et al.*, 2010; Mckee and Greenstein, 2000; Olson, *et al.*, 2012); Bayesian (kernel) classifiers (Gestel *et al.*, 2006; Sarkar and Sriram, 2001; Sun and Shenoy, 2007; Wu, 2011); data envelopment analysis (Cielen, *et al.*, 2004; Pendharkar, 2002; Premachandra, *et al.*, 2011); or fuzzy logic (Scherger, *et al.*, 2014; Terceño and Vigier, 2011; Vigier and Terceño, 2008, 2012).

Lately, many researchers have worked on hybrid methods based on two or three algorithms. The most popular models are integrating artificial neural networks, case-based reasoning and support vector machine with other techniques (Sun *et al.*, 2014).

Despite of the introduction of new techniques, most of studies use a timeframe data from 1 to 3 years previous to the bankrupt. Some authors have analyzed the time range of prediction. Studies as du Jardin and Séverin (2011, 2012) o du Jardin (2015) analyze how bankrupt horizons of 1, 2 or 3 years affect the classification error rates applying SOM and using financial data from the 7 years previous to the bankrupt time horizon. Gepp and Kumar (2008) applied survival analysis along with multiple discriminant analysis and logit analysis, aiming to compare the results obtained from 1 to 10 years bankrupt horizons. Few authors have analyzed the impact of the number of years used as input data in the prediction accuracy. Among them, Berg (2007) applied different methodologies comparing business failure prediction results using financial information of 1, 2 or 3 years previous to the bankrupt.

3. DATA AND METHODOLOGY

3.1. SAMPLE DESCRIPTION

The sample contained 180 companies randomly chosen from Spanish chemical industry listed in the SABI[1] database with a period of 9 years (1999-2007). From these 180 companies, 60 of them got bankrupted on the following 3 years (2008, 2009 or 2010).

The bankruptcy timeframe has been selected because it coincides with the beginning of the financial crisis. There are more bankrupt companies from 2007 which allows us to measure, in a more realistic way, the global error rate committed by the prediction model. Using a small number of bankrupt firms, any misclassifying company may have a great impact on distorting the results. This sensitivity decreases by increasing the bankrupt companies in the sample.

On the other hand, a sample of one single sector has been taken so that the financial ratio values represent the concrete financial situation of the company, not considering the specific factors of the sector it belongs.

Some test of coherence has been applied to the initial data, following the accounting principles. Those companies that do not fulfill the coherence test have been eliminated from the sample, as well as those ones which have insufficient information for calculating the financial ratios applied in this study.

The final sample was reduced to 159 companies, 110 that have not been bankrupt in the following 3 years after the study (healthy companies), and 49 that have been bankrupt in this period, (bankrupt companies). The reasons were:
i) 14 companies were ruled out due to a lack of data for the entire series and
ii) 7 companies did not pass the consistency tests applied for this period.

3.2. DESCRIPTION OF VARIABLES

The business failure literature provides a great variety in the number of factors used as variables. This number of factors ranges between 1 and 57 (du Jardin, 2009). Among most used factors according to Bellovary, *et al.* (2007), in this study 18 financial ratios have been selected, considering the possibility of elaborate them with the available information and avoiding duplicities in the information given by these ratios.

[1] SABI: Sistema de Análisis de Balances Ibéricos (Iberian Balance Sheets Analysis System) [online database] (2015). Madrid: Informa D&B. Available by license: http://www.informa.es/es/soluciones-financieras/sabi.

Table 1 shows the ratios used in the paper. These ratios analyze aspects such as profitability, solvency, liquidity, structure, turnover or size of the company.

Table 1. Factors (ratios) used in the study

1. Net profit / Total assets
2. Current ratio: Current assets / Current liabilities
3. Return on Shareholder Funds: Profit before taxes / Shareholders funds
4. Return on total assets: Profit before taxes / Total assets
5. Gearing ratio: Noncurrent liabilities / Shareholder funds
6. Interest cover: EBIT / Interests
7. Quick ratio: (Current assets – Stocks) / Current liabilities
8. Solvency ratio: Shareholder funds / Total assets
9. Return on capital employed: EBIT / (Shareholders funds + Noncurrent liabilities)
10. Profit margin: Profit before taxes / operating income
11. ROA: EBIT / Total assets
12. Leverage: Debt / Shareholders funds
13. Sales / Total assets
14. Cash flow / Sales
15. Cash flow / Total assets
16. Fixed assets / Total assets
17. Working capital ratio: Working capital/Total assets
18. Natural logarithm of total assets

3.3. THE USE OF SELF-ORGANIZING MAPS

One of the methodologies that has demonstrated its efficiency in business failure, also used in this study, are the Self-Organizing Maps (SOM). It is an artificial neural network introduced by (Kohonen, 1982, 1988). In contrast to many other neural networks based on supervised learning, SOM is based on unsupervised learning, that makes it one of the best known algorithms in artificial neural networks. (Back *et al.*, 1998; Chen, 2012; Chen *et al.*, 2013; du Jardin and Séverin, 2011, 2012; du Jardin, 2015; Kiviluoto, 1998; López Iturriaga and Sanz, 2015; Serrano-Cinca, 1996).

SOM consists of two neural layers. The input layer has as many neurons as variables are considered in the problem. Its function is to capture the input data and transfer them to the output layer. Each neuron of the input layer is connected to each neuron of the output layer through synaptic weights (w_{ij} corresponds to the weight that connects the neuron i in the input layer with the neuron j in the output layer).

The output layer is a two dimensional map. The input data are located at the map according to the similarity between variables.

By reducing the dimensions of the analysis an easiest and natural visualization of data is provided. Next, we briefly describe how it works.

The input vector k is presented to the network. In our study, k represents one company in a concrete year. Thus, according to the analysis carried out, there will be 477 input vectors (159 companies per 3 years), 954 vectors (159 companies per 6 years) or 1.431 (when considering 9 years). Each vector X_k is defined by 18 financial ratios of the company, which are the vector components. Thus, x_{ik} is the value of ratio i in the company-year identified as k, being $i = 1, \ldots, 18$.

Afterwards, as similarity measure, the Euclidean distance between the input vector and the weights vector that joins the input neurons with each of the output neurons is calculated (Equation (1)). The neuron in the output layer that minimizes this distance will be the "winning neuron".

$$d_j^k(t) = \sqrt{\sum_{i=1}^{N} \left(x_i^k(t) - w_{ij}^k(t)\right)^2} \tag{1}$$

Initially, $t = 0$, the synaptic weights take random values.

In order to reduce the distance between the input vector and the weights vector associated to the winning neuron, the synaptic weights are updated according to the following rule:

$$w_{ij}^k(t+1) = w_{ij}^k(t) + \alpha(t)h_j^k(t)\left[x_i^k(t) - w_{ij}^k(t)\right] \tag{2}$$

Where α(t) is the learning rate, a function that takes values from 0 to 1 and it decreases as the number of iterations increases. The neighborhood function (h(t)) allows to update the winning neuron's weights and its neighborhood (neurons close to the winning neuron at a radio that decreases with the number of iterations). This process facilitates the activation of the same winning neuron or other neuron close to it when similar input vectors are presented.

This process is repeated for the rest of input vectors and a maximum number of iterations T previously defined.

3.4. CLASSIFYING OUTPUT NEURONS

For each time span of input data used (3, 6 and 9 years) we will obtain a map in which the companies of the sample will be located according to their financial data similarity. Since we already know if these companies have been bankrupt or not within the 3 years after the input data timeframe, we will be able to evaluate, for each map, the error rate of misclassifying companies in "healthy" or "bankrupt". Each neuron in the output layer will be labeled as follows:

Notice E_p a set of companies projected on neuron p in the output layer where (p=1,…,P). The number of neurons in the output layer is different in each analysis, given that, as described previously, it depends on the timeframe taken.

Each neuron p is associated to a percentage of default (Q) defined as the ratio of bankrupt companies in each neuron in the output layer:

$$Q_p = \frac{\left|\left\{k \in E_p \middle| \text{Class}(k) = \text{bankrupt}\right\}\right|}{\left|E_p\right|} \quad (3)$$

being $\left|E_p\right|$ the number of companies located at neuron p. The bigger value Q_p, the higher bankruptcy probability will be for a company located at this neuron. In this manner we can label neurons according the following criterion:

$$\text{Label}_p = \begin{cases} \text{bankrupt} & \text{if } Q_p \geq 0,5 \\ \text{healthy} & \text{if } Q_p < 0,5 \end{cases} \quad (4)$$

On the basis of this classification, we calculate the prediction error rate exposed at the following section.

4. RESULTS COMPARING HORIZONS

This study aims to determine whether the SOM input data timeframe affects its capability in classifying healthy and bankrupt companies correctly. To do this, we repeat the analysis three times using 9, 6 and 3 years of input data. The SOM has been obtained using the toolbox for Matlab designed at Helsinki University of Technology (Vesanto, *et al.* 1999). The *toolbox* determines automatically the optimal map size according to the 18 rates used, the number of companies and the number of years studied (patterns).

The SOM capacity to place similar input vectors in close neurons in the output layer is useful to determine the common features of those companies located at each zone in the map. In consequence, a company with an unknown future solvency can be classified into “healthy” or “bankrupt” considering its location in the map.

Next, we introduce the resulting map for each input data timeframe:

4.1. SOM BY USING 9 YEARS OF INPUT DATA

Figure 1 presents the results by applying the percentage of default defined in Equation (3) to the neurons of the output layer of the SOM trained out using financial data for the period 1999-2007. If the percentage of default of one neuron is equal or greater than 50% then we consider it is a “bankrupt” cell -in red- whereas if the percentage of default is lower than 50%, we consider it is a ”healthy” cell -in blue-. There are blank cells which represent the map units where no company has been located.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0,2	0	0,08
0	0	0	0	0	0,33	0	0,14	0	0
0	0	0,07	0	0,5	0,2	0,14	0	0	0
0,38	0,4	0	0,25	0,25	0,17	0	0	0,33	0
0,17	0,33	0	0	0	0	0	0,5	0	0,43
0,57	1	0,13	0,1	0	0	0	0,17	0	0,17
0,3	0	0,07	0	0	0	0,14	0	1	0,1
0,21	0	0	0	0	0	0,67	0,38	0,25	0,2
0,7	0,67	0	0	0,38	0,2	0,67	0,63	0,5	0,44
1	0,14	1	0,5	0	0,25	0	0	0,2	0
0,5	0,5	1		0,33	0,27	0,29	0,17	0,17	0,21
0,61	0,33	1	0,67	0,43	0,33	0,2	0	0,11	0,67
1	1	0,4	0	0	0	0,29	0	0,063	0,2
1	1	1	0	0,14	0,44	0,29	0,4	0	0
1	1	0	0,6	0,67	0,14	0,71	0,25	0,67	0
1	1	0,25	0	0,83	0,33	0,714	1	0,75	0
1	0,89	0,63	1	0,29	0	1	1	0,57	0
1	1	0,75	0,64	0	0	0,333	1	1	0,38
1	0,82	0,5	0,38	0,25	0	0,88	0,75	0,78	0,33

Figure 1. Percentage of default in the map using 9 years of input data

4.2. SOM BY USING 6 YEARS OF INPUT DATA

Figure 2 shows the results obtained by applying the same procedure to the SOM trained using financial data for a 6-year timeframe (2002-2007). Again, in red color are the "bankrupt" cells and in blue are the "healthy" ones.

0	0	0	0	0	0	0	0	0,13
0	0	0	0	0	0	0	0	0
0,08	0	1	0,14	0	0	0	0	0
0	0,6	0	0	0	0	0,25	0	0
0	1	0	0	0	0	0	0,67	0
	0,6	0	0	0,2	0	0,14	0,14	0,33
0,2	0	0	0	0	0,29	0,6	0,33	0
0	0,17	0	0,38	0,33	0,2	0	0,1	0,29
0	0	0	0,5		0,54	0,29	0,33	0,5
1	0	0,4	0	1	1	0	0,07	0,25
1	0	0	0	0,57	0,8	0,2	0	0
1	0,25	0,5	0,78	1	0	0	0	0,35
1	0,5	0	0	0,33	0	0	1	0
1	1	0,27	0	0,17	0,2	0,43	0,67	0
1	1	0,43	0,57	0	0,33	0,2	0,7	0
1	0,75	0	0,5	0	0	1	0,5	0,25
1	0,67	0,33	0,2	0,5	0,8	0,8	1	0,83

Figure 2. Percentage of default in the map using 6 years of input data

4.3. SOM BY USING 3 YEARS OF INPUT DATA

Finally, Figure 3 represents the resultant map using 3 years of financial data (2005-2007), with the same color scale used before (red for “bankrupt” cells and blue for “healthy” cells).

The goal of this study is to test if the input data time span affects the prediction capability of business failure of the SOM model proposed. For that purpose we compare the error/success percentage for each of the three situations for our sample.

0,2	0,2	0	0	0	0	0	0
0,33		0	0		0	0	0,14
0	0	0	0	0	0	0	0
0	0,33	0	0,25	0	0	0	0
0,14	0	0	0	0	0,75	0,33	0
	0		0,25	0,4	0,33		0,14
1	0	0	0	0	0	0	0
1	1	0,4	0,25	0	0,25	0	0,33
1	1	0,75	0,33		0	0	0
0,5	0,29	0,25	0,13	0	0		0,36
0,5	0	0,83	0	1	0,33	0,5	0,4
0	0,2		0,4	1	0,14	0	0
1	0	0	1	1	1	0	0
1	0,8	0,75	1	1	0,86	0	0,78

Figure 3. Percentage of default in the map using 3 years of input data

In classifying a company as “healthy” or “bankrupt” two error types might occur:

- Error type I: when the SOM model misclassify a “bankrupt” company as a “healthy”. In our model, that would mean the location of a “bankrupt” company in a blue colored cell (“healthy zone”).
- Error type II: when model misclassify a “healthy” company as a “bankrupt” one. In our model, error type II occurs when the SOM model locates a “healthy” company in a cell colored in red (“bankrupt” zone).

Error percentages for each type are shown in Table 2 and so is the total error resulting for each of the three time spans considered.

The results obtained suggest that as increasing the number of years of input data included in the study so does the total error percentage. In particular, when we use data from the three previous years to the bankruptcy time horizon this percentage is down to 12,58%, because of we obtained a 14,95% using financial data from the previous 9 years. Analyzing separately each type of error, we observe that error type II- that classifies a healthy company as a bankrupt one- improves as we reduce the number of years included in the study. On the other hand, error type I –that classifies a bankrupt company as a healthy one- gets its best accuracy percentage in the 6 years study: a 92,35% (7,65% of error) compared to the 91,40% (8,60% of error) that is obtained in the 3-year study and the 90,99% (9,01%) in the 9 years study.

Table 2. Error type percentages depending on the input data time span

	SOM 3 years	SOM 6 years	SOM 9 years
Error type I	8,60%	7,65%	9,01%
Error type II	3,98%	5,77%	5,94%
Total Error	12,58%	13,42%	14,95%

5. CONCLUSIONS

Most of the studies about business failure prediction are focused on the correct selection of factors and applying more sophisticated techniques for improving the prediction accuracy. However, few studies focused their attention on the time span of the analysis.

This paper analyses whether the input data time span affects the error percentages in classifying companies as healthy or bankrupt, applying self-organizing maps (SOM). 18 financial ratios among most used factors according to literature are used in this study. We have analyzed 159 companies from the Spanish chemical industry, and have included financial data within 3 years (2005-2007), 6 years (2002-2007) and 9 years (1999-2007). In order to train different SOM and evaluate how the input data time span affects its capability of grouping companies correctly in two groups: “healthy” companies- those who have not been bankrupt in the period 2008-2010- and “bankrupt” companies- those that were bankrupt within this period.

Once SOM are obtained for the three different scenarios, we observed that as increasing the number of years included in the study so does the percentage of total error committed by the SOM model. Despite the error type II (classifying a healthy company as a bankrupt one) increases as increasing the timeframe data, in error type I (classifying bankrupt company as a healthy one) the best result is obtained by the SOM with 6-year data.

It is generally accepted that cost of incurring the error type I is worse than the cost of incurring error type II. This is due to the losses that it might generate investing in a high risk company that has been considered as a healthy one.

However, error type II must not be ignored, because considering a healthy company as a bankrupt one it might negatively affect its reputation, relationship with clients, commercial suppliers or financial suppliers, increasing costs and/or decreasing incomes.

We therefore conclude that the 3-year timeframe scenario provides better prediction accuracy despite of its error type I which is slightly higher than the 6-year timeframe scenario as it shows the lowest total error.

REFERENCES

ALTMAN, E.I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy". *The Journal of Finance*, Vol. 23, No. 4, p. 589-609.

BACK, B.; SERE, K.; VANHARANTA, H. (1998). "Analyzing Financial Performance with Self-Organizing Maps". In *Neural Networks Proceedings*. *IEEE World Congress on Computational Intelligence,* Vol. 1, p. 266–270.

BEAVER, W.H. (1966). "Financial Ratios As Predictors of Failure". *Journal of Accounting Research*, Vol. 4, p. 71-111.

BELLOVARY, J.L.; GIACOMINO, D.E.; AKERS, M.D. (2007). "A Review of Bankruptcy Prediction Studies□: 1930-Present". *Journal of Financial Education*, Vol. 33(Winter), p. 1-42.

BERG, D. (2007). "Bankruptcy Prediction by Generalized Additive Models". *Applied Stochastic Models in Business and Industry*, Vol. 23, p.129–143.

BLUM, M. (1974). "Failing Company Discriminant Analysis". *Journal of Accounting Research*, Vol. 12, No. 1, p. 1-25.

BORRAJO, M.L.; BARUQUE, B.; CORCHADO, E.; BAJO, J.; CORCHADO, J.M. (2011). "Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises". *International Journal of Neural Systems*, Vol. 21, No. 4, p. 277-296.

BOSE, I. (2006). "Deciding the financial health of dot-coms using rough sets". *Information & Management*, Vol. 43, No. 7, p. 835-846.

BOSE, I.; PAL, R. (2006). "Predicting the survival or failure of click-and-mortar corporations: A knowledge discovery approach". *European Journal of Operational Research*, Vol. 174, No. 2, p. 959-982.

CHEN, M.Y. (2012). "Visualization and dynamic evaluation model of corporate financial structure with self-organizing map and support vector regression". *Applied Soft Computing*, Vol. 12, No. 8, p. 2274-2288.

CHEN, N.; RIBEIRO, B.; VIEIRA, A.; CHEN, A. (2013). "Clustering and visualization of bankruptcy trajectory using self-organizing map". *Expert Systems with Applications*, Vol. 40, No. 1, p. 385-393.

CHEN, W.S.; DU, Y.K. (2009). "Using neural networks and data mining techniques for the financial distress prediction model". *Expert Systems with Applications*, Vol. 36, No. 2, p. 4075-4086.

CIELEN, A.; PEETERS, L.; VANHOOF, K. (2004). "Bankruptcy prediction using a data envelopment analysis". *European Journal of Operational Research*, Vol. 154, No. 2, p. 526-532.

DIMITRAS, A.I.; SLOWINSKI, R.; SUSMAGA, R.; ZOPOUNIDIS, C. (1999). "Business failure prediction using rough sets". *European Journal of Operational Research*, Vol. 114, No. 2, p. 263-280.

DING, Y.; SONG, X.; ZEN, Y. (2008). "Forecasting financial condition of Chinese listed companies based on support vector machine". *Expert Systems with Applications*, Vol. 34, No. 4, p. 3081-3089.

DU JARDIN, P. (2009). "Bankruptcy prediction models: How to choose the most relevant variables?". *Bankers, Markets & Investors*, Vol. 98, p. 39-46.

DU JARDIN, P. (2015). "Bankruptcy prediction using terminal failure processes". *European Journal of Operational Research*, Vol. 242, No. 1, p. 286–303.

DU JARDIN, P.; SÉVERIN, É. (2011). "Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model". *Decision Support Systems*, Vol. 51, No. 3, p. 701-711.

DU JARDIN, P.; SÉVERIN, É. (2012). "Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time". *European Journal of Operational*

Research, Vol. 221, No. 2, p. 378-396.

DWYER, M.M.D. (1992, January 1). *A comparison of statistical techniques and artificial neural network models in corporate bankruptcy prediction*. University of Wisconsin at Madison. Retrieved from http://dl.acm.org/citation.cfm?id=166736

ETEMADI, H.; ANVARY ROSTAMY, A.A.; DEHKORDI, H.F. (2009). "A genetic programming model for bankruptcy prediction: Empirical evidence from Iran". *Expert Systems with Applications*, Vol. 36, No. 2, p. 3199-3207.

FITZPATRICK, P.J. (1932). "A Comparison of the Ratios of Successful Industrial Enterprises With Those of Failed Companies". *The Certified Public Accountant*, Vol. 12 (October, November and December), p. 598-605, 656-662 and 727-731 respectively.

FRYDMAN, H.; ALTMAN, E.I.; KAO, D.L. (1985). "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress". *The Journal of Finance*, Vol. 40, No. 1, p. 269-291.

GEPP, A.; KUMAR, K.; BHATTACHARYA, S. (2010). "Business failure prediction using decision trees". *Journal of Forecasting*, Vol. 29, p. 536-555.

GESTEL, T. VAN; BAESENS, B.; SUYKENS, J.A.K.; VAN DEN POEL, D.; BAESTAENS, D.E.; WILLEKENS, M. (2006). "Bayesian kernel based classification for financial distress detection". *European Journal of Operational Research*, Vol. 172, No. 3, p. 979-1003.

KIVILUOTO, K. (1998). "Predicting bankruptcies with the self-organizing map". *Neurocomputing*, Vol. 21, No. 1-3, p. 191–201.

KOHONEN, T.K. (1982). "Self-organized formation of topologically correct feature maps". *Biological Cybernetics*, Vol. 43, No. 1, p. 59-69.

KOHONEN, T.K. (1988). *Self-Organization and Associative Memory*. *Springer-Verlag*. Berlin.

LI, H.; ADELI, H.; SUN, J.; HAN, J.G. (2011). "Hybridizing principles of TOPSIS with case-based reasoning for business failure prediction". *Computers & Operations Research*, Vol. 38, No. 2, p. 409-419.

LI, H.; SUN, J. (2009). "Hybridizing principles of the Electre method with case-based reasoning for data mining: Electre-CBR-I and Electre-CBR-II". *European Journal of Operational Research*, Vol. 197, No. 1, p. 214-224.

LI, H.; SUN, J. (2010). "Business failure prediction using hybrid2 case-based reasoning (H2CBR)". *Computers & Operations Research*, Vol. 37, No. 1, p. 137-151.

LI, H.; SUN, J.; WU, J. (2010). "Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods". *Expert Systems with Applications*, Vol. 37, No. 8, p. 5895-5904.

LIBBY, R. (1975). "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence". *Journal of Accounting Research*, Vol. 13, No. 1, p. 150-161.

LÓPEZ ITURRIAGA, F. J.; SANZ, I. P. (2015). "Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks". *Expert Systems with Applications*, Vol. 42, No 6, p. 2857–2869.

MARTENS, D.; VAN GESTEL, T.; DE BACKER, M.; HAESEN, R.; VANTHIENEN, J.; BAESENS, B. (2010). "Credit rating prediction using Ant Colony Optimization". *Journal of the Operational Research Society*, Vol. 61, No. 4, p. 561-573.

MCKEE, T.E. (2000). "Developing a Bankruptcy Prediction Model via Rough Sets Theory". *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 9 (February), p. 159–173.

MCKEE, T.E.; GREENSTEIN, M. (2000). "Predicting Bankruptcy Using Recursive Partitioning and a Realistically Proportioned Data Set". *Journal of Forecasting*, Vol. 19, No. 3, p. 219-230.

MERWIN, C.L. (1942). *Financing Small Corporations in Five Manufacturing Industries, 1926-36*. National Bureau of Economic Research, Inc.

MIN, J.H.; LEE, Y.C. (2005). "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters". *Expert Systems with Applications*, Vol. 28, No. 4, p. 603-614.

MOKHATAB RAFIEI, F.; MANZARI, S.M.; BOSTANIAN, S. (2011). "Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis: Iranian evidence". *Expert Systems with Applications*, Vol. 38, No. 8, p. 10210-10217.

ODOM, M.D.; SHARDA, R. (1990). "A neural network model for bankruptcy prediction", in *IJCNN International Joint Conference on Neural Networks* p. 163–168, Vol.2.

OHLSON, J.A. (1980). "Financial Ratios and Probabilistic Prediction of Bankruptcy". *Journal of Accounting Research*, Vol. 18, No. 1, p. 109-131.

OLSON, D.L.; DELEN, D.; MENG, Y. (2012). "Comparative analysis of data mining methods for bankruptcy prediction". *Decision Support Systems*, Vol. 52, No. 2, p. 464-473.

PENDHARKAR, P.C. (2002). "A potential use of data envelopment analysis for the inverse classification problem". *Omega*, Vol. 30, No. 3, p. 243-248.

PREMACHANDRA, I. M.; CHEN, Y.; WATSON, J. (2011). "DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment". *Omega*, Vol. 39, No. 6, p. 620-626.

SARKAR, S.; SRIRAM, R.S. (2001). "Bayesian Models for Early Warning of Bank Failures". *Management Science*, Vol. 47, No. 11, p. 1457-1475.

SCHERGER, V.; VIGIER, H.P.; BARBERÀ-MARINÉ, M.G. (2014). "Finding Business Failure Reasons Through a Fuzzy Model of Diagnosis". *Fuzzy Economic Review*, Vol. XIX, No. 1, p. 45-62.

SERRANO-CINCA, C. (1996). "Self organizing neural networks for financial diagnosis". *Decision Support Systems*, Vol. 17, No. 3, p. 227-238.

SÉVERIN, É. (2010). "Self organizing maps in corporate finance: Quantitative and qualitative analysis of debt and leasing". *Neurocomputing*, Vol. 73, No. 10-12, p. 2061-2067.

SHIN, K.S.; LEE, T.S.; KIM, H. (2005). "An application of support vector machines in bankruptcy prediction model". *Expert Systems with Applications*, Vol. 28, No. 1, p. 127-135.

SHIN, K.S.; LEE, Y.J. (2002). "A genetic algorithm application in bankruptcy prediction modeling". *Expert Systems with Applications*, Vol. 23, No. 3, p. 321-328.

SMITH, R.F.; WINAKOR, A.H. (1935). *Changes in the financial structure of unsuccessful industrial corporations*. Bureau of Business Research, Bulletin No. 51. University of Illinois Press.

SUN, J.; LI, H.; HUANG, Q.H.; HE, K.Y. (2014). "Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches". *Knowledge-Based Systems*, Vol. 57, p. 41-56.

SUN, L.; SHENOY, P.P. (2007). "Using Bayesian networks for bankruptcy prediction: Some methodological issues". *European Journal of Operational Research*, Vol. 180, No. 2, p. 738-753.

TAM, K.Y. (1991). "Neural network models and the prediction of bank bankruptcy". *Omega*, Vol. 19, No. 5, p. 429-445.

TERCEÑO, A.; VIGIER, H.P. (2011). "Economic-financial forecasting model of businesses using fuzzy relations". *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 45, No. 1, p. 215-232.

VAN GESTEL, T.; BAESENS, B.; MARTENS, D. (2010). "From linear to non-linear kernel based classifiers for bankruptcy prediction". *Neurocomputing*, Vol. 73, No. 16-18, p. 2955-2970.

VARETTO, F. (1998). "Genetic algorithms applications in the analysis of insolvency risk". *Journal of Banking & Finance*, Vol. 22, No. 10-11, p. 1421-1439.

VESANTO, J.; HIMBERG, J.; ALHONIEMI, E.; PARHANKANGAS, J. (1999). "Self−Organizing Map in Matlab□: the SOM Toolbox". In *Matlab DSP Conference 1999*, p. 35-40. Espoo, Finland.

VIGIER, H.P.; TERCEÑO, A. (2008). "A model for the prediction of “diseases” of firms by means of fuzzy relations". *Fuzzy Sets and Systems*, Vol. 159, No. 17, p. 2299-2316.

VIGIER, H.P.; TERCEÑO, A. (2012). "Analysis of the inconsistency problem in the model for predicting “diseases” of firms". *Fuzzy Economic Review*, V84ol. 17, No. 1, p. 73-88.

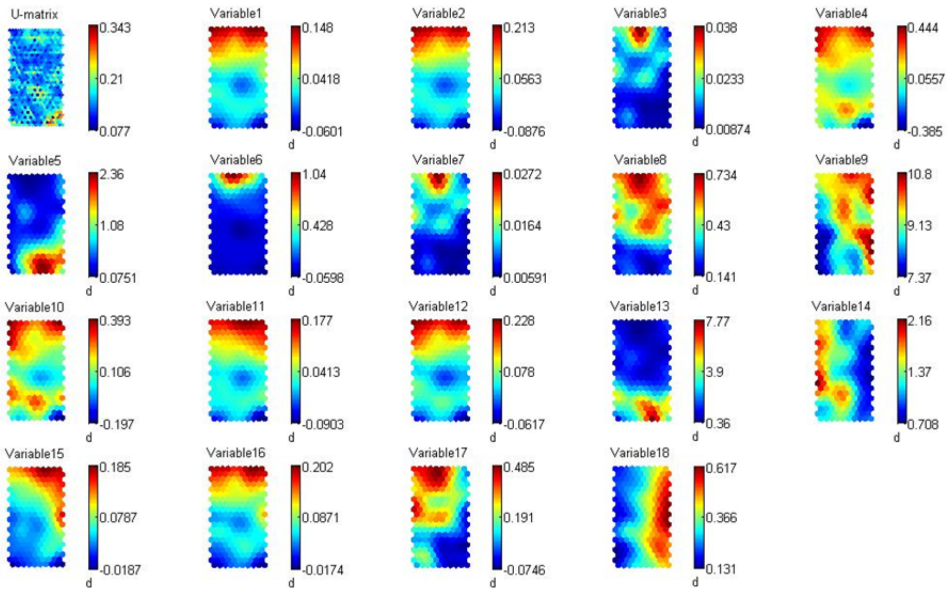
WEST, R.C. (1985). "A factor-analytic approach to bank condition". *Journal of Banking & Finance*, Vol. 9, No. 2, p. 253-266.

WU, W.W. (2011). "Improving Classification Accuracy and Causal Knowledge for Better Credit Decisions". *International Journal of Neural Systems*, Vol. 21, No. 4, p. 297-309.

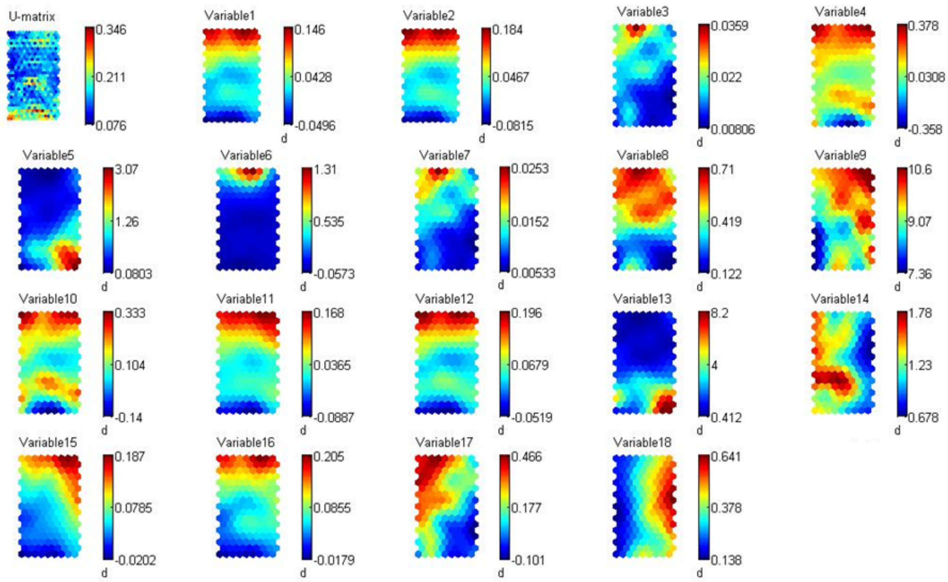
ZMIJEWSKI, M.E. (1984). "Methodological Issues Related to the Estimation of Financial Distress Prediction Models". *Journal of Accounting Research*, Vol. 22, p. 59-82.

ANNEX A: SOM COMPONENT PLANES

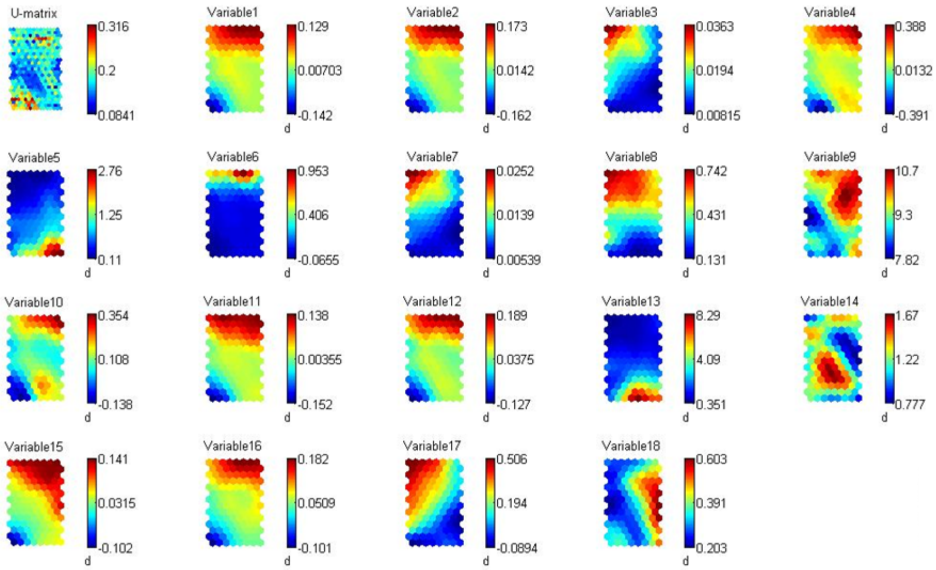
SOM component plane of 9 years of input data



SOM component plane of 6 years of input data



SOM component plane of 3 years of input data



Legend:

Variable 1	Net Profit/Total Assets	Variable 10	Return on Capital Employed
Variable 2	Return on Total Assets	Variable 11	Profit Margin
Variable 3	Current Ratio	Variable 12	ROA
Variable 4	Return on Shareholder Funds	Variable 13	Leverage
Variable 5	Gearing	Variable 14	Sales/Total Assets
Variable 6	Interest cover	Variable 15	cash flow/Sales
Variable 7	Quick ratio	Variable 16	cash flow/Total Assets
Variable 8	Solvency Ratio	Variable 17	Working capital ratio
Variable 9	ln(Total Assets)	Variable 18	Fixed Asset/Total Assets

ANNEX B: COMPANY LIST

Healthy companies			
E1	AC MARCA SA	E35	FLEXICEL SA
E2	AGRAR FERTILIZANTES SA	E36	FLINT GROUP IBERIA SA
E3	ALCALIBER SA	E37	FORESTAL DEL ATLANTICO SA
E4	APLICACION Y SUMINISTROS TEXTILES SA	E38	FUCHS LUBRICANTES SAU
E5	AROM SA	E39	GAT FERTILIQUIDOS SA
E6	BADRINAS SA	E40	GIVAUDAN IBERICA SA
E7	BECA GRAFIC SOCIEDAD ANONIMA	E41	GRUPO DRV PHYTOLAB SL.
E8	BEISSIER SA	E42	GUEROLA SA
E9	BERIOSKA SL	E43	HISPANAGAR SA
E10	BIOKIT SAU	E44	HOUGHTON IBERICA SA
E11	BIOSYSTEMS SA	E45	HUBERGROUP SPAIN SA.
E12	CASALS CARDONA INDUSTRIAL SA	E46	IBERICA DE REVESTIMIENTOS SL
E13	CHIMIGRAF IBERICA SL	E47	IGEPAK SA
E14	COLORES CERAMICOS SA	E48	INDUSTRIA JABONERA LINA SAU.
E15	COMERCIAL PRODUCTOS DE LIMPIEZA SA	E49	INDUSTRIAL QUIMICA DEL NALON, SA
E16	COMPAÑIA DE SUMINISTROS CADES SA	E50	INDUSTRIAS CATALA SA
E17	COMPUESTOS Y GRANZAS SA	E51	INDUSTRIAS DEL ACETATO DE CELULOSA, SA
E18	CROMOGENIA UNITS SA	E52	INDUSTRIAS PROA SA
E19	DERETIL SA	E53	INDUSTRIAS QUIMICAS DE BADAJOZ SOCIEDAD ANONIMA
E20	DERMOFARM SA	E54	INDUSTRIAS QUIMICAS DEL ADHESIVO SA
E21	DERYPOL SA	E55	INDUSTRIAS QUIMICAS I V M SA
E22	DIAGNOSTIC GRIFOLS SA	E56	INDUSTRIAS QUIMICAS SATECMA SA
E23	DISTILLER, SA	E57	INSTITUTO GRIFOLS SA
E24	DURA EUROPE SA.	E58	JABER SA
E25	ELECTROQUIMICA DE HERNANI SA	E59	JABONES PARDO SA
E26	ELECTROQUIMICA DEL NOROESTE SA	E60	JOVI SA
E27	ELECTROQUIMICA DEL SERPIS SA	E61	LABORATORIO ALDO UNION, S.L.
E28	ESLAVA PLASTICOS SA	E62	LABORATORIO BONIQUET SAU
E29	ESMALDUR SA	E63	LABORATORIO JAER SA
E30	ESMALTES, SOCIEDAD ANONIMA	E64	LABORATORIOS ALTER SA
E31	ESTEVE QUIMICA SA	E65	LABORATORIOS CINFA SA
E32	EUPINCA SA	E66	LABORATORIOS ERN SA
E33	FABRICACION Y APLICACION DE PINTURAS ESPECIALES SA	E67	LABORATORIOS FARMACEUTICOS ROVI SA
E34	FARMHISPANIA S.A.	E68	LABORATORIOS GRIFOLS SA

E69	LABORATORIOS INIBSA SA	E90	PRODUCTOS CONCENTROL SOCIEDAD ANONIMA
E70	LABORATORIOS KIN SA	E91	PRODUCTOS Q P SOCIEDAD ANONIMA.
E71	LABORATORIOS MAYMO SA	E92	PRODUCTOS QUIMICOS MOPASA SL
E72	LABORATORIOS NORMON SA	E93	PROTEIN SOCIEDAD ANONIMA
E73	LABORATORIOS OVEJERO, SA	E94	QUIDE, SOCIEDAD ANONIMA
E74	LABORATORIOS RAYT SA	E95	QUIMICA DEL CINCA SL
E75	LABORATORIOS SALVAT SA	E96	QUIMICAS ORO SA
E76	LABORATORIOS SYVA SA	E97	SA DABEER
E77	LABORATORIOS VERKOS SA	E98	SEID SA
E78	LAINCO SA	E99	SIPCAM INAGRA SA
E79	MONTEFIBRE HISPANIA SA	E100	SOCIEDAD ANONIMA REVERTE PRODUCTOS MINERALES
E80	NORTE-EUROCAO SL.	E101	SOCIEDAD ANONIMA ROBAMA
E81	PINTURAS BLATEM SL	E102	SOCIEDAD ESPAÑOLA DE COLORANTES NATURALES Y AFINES SA
E82	PINTURAS HEMPEL SA	E103	SPUMATEX SL
E83	PINTURAS MACY SOCIEDAD ANONIMA	E104	SUN CHEMICAL SA
E84	PIROTECNIA IGUAL SA	E105	SYNTHESIA ESPAÑOLA SA
E85	PLASTICOS COMPUESTOS SA	E106	TAKASAGO INTERNATIONAL CHEMICALS EUROPE SA
E86	PLYMOUTH RUBBER EUROPA SA	E107	UBE CHEMICAL EUROPE SA
E87	POLIDUX SOCIEDAD ANONIMA	E108	UNILEVER ESPAÑA SA
E88	PRAXAIR ESPAÑA SL	E109	UNION QUIMICO FARMACEUTICA SA
E89	PRODUCTOS CITROSOL SOCIEDAD ANONIMA	E110	ZELNOVA SA

Bankrupt companies			
Q1	AGROQUIVIR SA	Q26	LABORATORIOS PEREZ GIMENEZ, SA
Q2	ARALUR SL	Q27	MASSO Y CAROL SA
Q3	CARPINTERIA RUEDA SL	Q28	MERIDIONAL DE LIMPIEZAS SA
Q4	CATRAL EXPORT SOCIEDAD LIMITADA	Q29	MOLDURAS VICENTE Y ANA SOCIEDAD LIMITADA
Q5	CERERIA MAS SA	Q30	NORQUIMIA SA
Q6	COLAMINA SA	Q31	NUÑEZ Y CAÑADAS SA
Q7	COMERCIAL INDUSTRIAL DE LA SERENA SL	Q32	OMICRON QUIMICA SA
Q8	COSMETICA TECNICA SA	Q33	PINTURAS SANTANA SA
Q9	ESTELRICH Y ESTELRICH SL	Q34	PINTURAS TARRACOL SOCIEDAD LIMITADA
Q10	FUEGOS ARTIFICIALES ANTONIO CABALLER SA	Q35	PINTURAS VALSOL SOCIEDAD ANONIMA
Q11	GAIRALT SA	Q36	PROCOAT SL
Q12	GIP QUIMICA ESPAÑOLA SA	Q37	PRODUCTOS LAZARO S A
Q13	GONCUVI SL	Q38	PRODUCTOS QUETZAL SOCIEDAD ANONIMA
Q14	HEREDEROS DE SALVADOR SEGURA SA	Q39	PRODUCTOS SOLRAC SA
Q15	HERMANOS SEOANE SL	Q40	QUIMICA APLICADA CANARIA SA
Q16	INADYP SL	Q41	QUIMIGAMA SL
Q17	INDUSTRIAS EGA SA	Q42	R VIÑALS SOLER, SL
Q18	INDUSTRIAS ESPEJO SA	Q43	RECUBRIMIENTOS TECNICOS SA
Q19	INDUSTRIAS GMB SA	Q44	ROBERTO CARLOS EISER SL
Q20	INDUSTRIAS PARSAN SL	Q45	TRAMAGRAF SA
Q21	INDUSTRIAS QUIMICAS TEXTILES, SA	Q46	TRANSFORMADOS Y FORMULACIONES QUIMICAS SL
Q22	INTEC FABRICACIONES SA	Q47	URGAZE SA
Q23	INTER LARAK SL	Q48	VICENTE MARIN SOCIEDAD LIMITADA
Q24	KRAUTO IBERICA SA	Q49	VISOPLAST SL
Q25	LA SEDA DE BARCELONA SA		