Bankruptcy prediction and neural networks: The contribution of variable selection methods

du Jardin, Philippe

Edhec Business School

17 September 2008

du Jardin, P., 2008, Bankruptcy prediction and neural networks: The contribution of variable selection methods, proceedings of the Second European Symposium on Time Series Prediction (Estsp 2008), Helsinki University of Technology, 17-19 September, Porvoo, Finland.

Bankruptcy prediction and neural networks: the contribution of variable selection methods

Philippe du Jardin

Edhec Business School – Information Technology Department
393, Promenade des Anglais – BP 3116 – 06202 Nice Cedex 3 – France

Abstract. Of the methods used to build bankruptcy prediction models in the last twenty years, neural networks are among the most challenging. Despite the characteristics of neural networks, most of the research done until now has not taken them into consideration for building financial failure models, nor for selecting the variables to be included in the models. The aim of our research is to establish that to improve the prediction accuracy of the models, variable selection techniques developed specifically for neural networks may well offer a useful alternative to conventional methods.

1 Introduction

The history of bankruptcy prediction models may be divided into two main periods. The first starts with Altman's [1] and Ohlson's [2] models. During this period, which goes from the late 1960's to the late 1980's, research relied largely on discriminant analysis and logistic regression as methods of building the most accurate models. But as much research has since shown, these methods suffer from major drawbacks and the real input-output variables dependency (i.e., the dependency between financial ratios as explanatory variables and the probability of failure) may be neither linear nor logistic; in other words, it has gradually become clear that other methods should be studied and used to create bankruptcy models.[3]

The second period begins in the late 1980's, when many authors, in attempts to overcome the limitations described above, undertook research to assess the ability of non-parametric methods to accurately predict the risk of bankruptcy or the risk of financial failure. It was also during this period that non-linear techniques such as neural networks emerged in this field of research and demonstrated their frequent ability to outperform most existing techniques, whether parametric or not.

But, whatever the method, when the goal of the research is to seek an effective means of improving the accuracy of a prediction, the variables to be included in the models are commonly selected either because they are among those commonly used in the field of financial analysis, such a set being historically validated through univariate statistical tests (most of the time, t of F test–in which case there is no guarantee that this historical reference is sufficient to create the best models), or because selection is the result of automated processes, often optimized for linear methods (and in this case, there is no guarantee that such processes are relevant in any situation). For instance, it is not particularly relevant to use a variance-based criterion or a likelihood-based criterion to select a set of variables for a model with a method to which these parameters are not well suited, especially with non-linear methods. The subsets which could be estimated in such a way may be under-optimized because the criterion used to assess their legitimacy does not make sense in a non-linear context.

Thus, we have seen the influence of different variable selection processes on the accuracy of a model and studied the fitness of the most widely used methods for designing bankruptcy models and several well known variable selection techniques. The content of the paper is organized as follows. In section 2 we describe the methods traditionally used to identify variables when the aim of a research is to build the most reliable bankruptcy prediction models. In section 3, we describe the methods and sample used in our experiments. Then, in section 4, we present and discuss the empirical results, and, in section 5, we summarize the main findings of the study.

2 Literature review

A reading of the major articles published over the past 50 years shows that, when developing business failure models, researchers usually use a two-step procedure to choose the « best » variables to be included in their models. Whereas a large set of variables is first identified based on general considerations (financial, empirical, and so on), only a few are finally chosen based on statistical issue.

The first group, often made up of a few tens of variables, is most often identified without using any automatic process but is arbitrarily chosen based on the popularity of variables in literature or on their predictive ability as assessed in previous studies. This « historical » set was built up on the strength of the seminal work done by researchers who, in the 1930's, first assessed the usefulness of financial ratios as a means of predicting corporate failure and by those who contributed to an understanding of the role played by multivariate statistical methods in the field of bankruptcy prediction. Among these latter researchers are Altman [1], Odom and Sharda [4], Zmijewski [5] and Zavgren [6]. All of this work may be viewed as the initial step towards the elaboration of a comprehensive set of essential bankruptcy predictors, which has been complemented over the years by other variables, whether they are accounting-based measures of the financial health of a firm or not (statistical variables calculated with financial data, variables measuring the evolution of financial indicators, non-financial variables describing a quantitative or qualitative characteristic of a company, market variables as a means to explain and quantify the way financial markets may evaluate the performance of companies through the price or the return they place on firm equity).

The second group, on the other hand, is most often selected through a computer-based procedure designed to mine the former group for the best set of variables, depending on an evaluation criterion to define *a priori*.

The evaluation criterion is very often a criterion that does not depend on the method used to develop models. This independence means that the inductive algorithm is not used to assess the value of a set of variables. However, whatever the criterion considered, it may not be without some influence on this algorithm. Indeed, a probabilistic distance, such as a Mahalanobis distance, or a distance calculated through a transformation of an intra- or inter-group covariance matrix, such as a Wilks Lambda, may be considered the criterion most suited to select variables to be used with a discriminant analysis, while a likelihood criterion may be most suited to select variables for a logistic regression. Nevertheless, the use of these criteria with methods for which they are not entirely optimized or suited is common practice. Indeed, many authors use such criteria to build neural network models [7, 8, 9, 10, 11, 12]. But Leray and Gallinari [13] have stated that since many parametric variable selection methods rely on the hypothesis that input-output variable dependence is linear or that input variable redundancy is well measured by the linear correlation of these variables, such methods are clearly ill-suited to non-linear methods, and hence to neural networks.

Moreover, many of those who have developed neural models have identified their final sets of variables simply on account of their popularity in the financial literature. If one analyzes the linking these studies, it is clear that the criteria used to assess the legitimacy of most of these variables make sense only in a linear context [4, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Very little research has used either a genetic algorithm [23, 24, 25, 26, 27] or a method suitable for non-linear techniques to take into account the characteristics of neural networks, and in each case, with only few variables, small samples, and without attempting comparisons of several methods or criteria [28, 29, 30, 31]; the significance of these experiments is thus reduced. Many authors also strongly recommend comparing the results obtained with different classification or regression techniques, but do not apply the same reasoning to the selection methods that will choose the variables relied on by these techniques.

The point is to show that many authors of bankruptcy models use variable selection methods without considering the very characteristics of the modelling techniques. It is for this reason that the aim of our research is to use « modelling method-variable selection technique » pair analysis to examine the influence of this common practice and to analyze the influence of the latter on the former, in terms of prediction accuracy. Only one study [23] has compared a pair of sets of variables optimized for a discriminant analysis (stepwise method and F test), a logistic regression (stepwise method, Rao's score test to add variables and a likelihood ratio test to discard variables) and a neural network (genetic algorithm), but just to analyze the differences between the models in terms of accuracy over different prediction timeframes (one, two or three years).

3 Methods and samples

3.1 Modelling techniques

Modelling methods are chosen for their popularity in the financial literature. Of the more than 50 regression or discriminant techniques, three predominate: discriminant analysis, logistic regression and a special type of neural network, known as multilayer perceptron, trained with a steepest descent method. A ten-cross validation technique is used to define all neural network parameters (topology, learning rate, momentum term, weight decay) and those that lead to the best out-of-sample error are then selected for our experiments. The final architecture is then composed of a single hidden layer with four nodes, an output node, a bias node in each layer and two weight decay parameters (one for the hidden layer and one for the output); a hyperbolic tangent is used as an activation function. We set the learning rate to 0.4, the momentum term to 0.4, and the weight decay to 10^{-4} and 10^{-3}. The learning process was stopped after 1,000 iterations, as no change could be observed in the error rate.

3.2 Variable selection procedures

The variable selection techniques we choose are those most commonly used in the literature. First, we have chosen a technique that relies on a forward search procedure to explore a (sub)space of possible variable combinations, a Fisher F test to interrupt the search, and a Wilks Lambda to compare variable subsets and determine the « best » one. This technique was complemented with two others: a forward stepwise search and a backward stepwise search, with a likelihood statistic as an evaluation criterion of the solutions and a Khi^2 as a stopping criterion. We then select three of the most commonly used [13] methods especially designed for neural networks, two of them evaluating the variables without using the inductive algorithm (filter methods) and one using the algorithm as an evaluation function (wrapper method). The first is a zero-order technique, which uses the evaluation criteria designed by Yacoub and Bennani [32] and the second is a first-order method that uses the first derivatives of network parameters with respect to variables as an evaluation criterion. The last one relies on the evaluation of an out-of-sample error calculated with the neural network. We do not choose a second-order method, based on second derivatives of network parameters, so as to investigate an equivalent number of points of comparison. With all these criteria, we use only a backward search procedure, rather than a forward or a sequential search, and the network is retrained after each variable removal. The zero and first order criterion were calculated as follows. With a network composed of n inputs, one hidden layer with h neurons and one output, where w_{ji} is the weight between input i and neuron j in the hidden layer, and w_j the weight between neuron j and the output, the relevance or the saliency S of a variable i may be defined as:

$$S_i = \sum_{j=1}^{h} \left(\frac{|w_{ji}|}{\sum_{k=1}^{n} |w_{jk}|} \frac{|w_j|}{\sum_{j'=1}^{h} |w_{j'}|} \right) \tag{1}$$

$$S_i = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{\partial y_i}{\partial x_{ji}} \right| \tag{2}$$

where x_i is a variable, y the output of the network calculated with only one neuron and N the sample size.

3.3 Variable selection procedure

To select variables, 1,000 random bootstrap samples were drawn from the original dataset. Each bootstrap sample involved selection. To identify important variables, those that were included in more than 70% of the selection results are included in the final models. To avoid discarding potentially relevant but highly correlated variables, variable pairs in which one or both variables are included in more than 90% of the bootstrap selections are considered pairs containing a relevant variable. Then, for each identified pair, the variable that occurs in most of the selection results is ultimately chosen. Once these selections are done, the entire process is repeated to choose the final subsets.

3.4 Model development

We used the following procedure to develop the models. The sample was randomly divided into two sub-samples: a learning sample A of 450 companies and a test sample T of fifty companies. 25 bootstrap samples are drawn from A and, for each selected set of variables, used to estimate as many models as bootstrap samples. Finally, the resulting models are used to classify the observations of sample T thanks to a majority voting scheme. These steps were repeated 100 times and the out-of-sample error is first estimated, along with a test sample, and then re-estimated using the 25 x 100 models, along with a validation sample of 520 companies.

3.5 Samples

The datasets (learning, test and validation sets) are drawn from a French database, Diane, which provides financial data on more than 2 million French companies. The learning and test samples consist of 250 bankrupt and 250 non-bankrupt retail firms which have assets of less than 750,000 €. Annual reports from 2002 are taken from this database to calculate a set of financial ratios, and we add one variable (shareholders' funds) from 2001. The validation set consists of companies belonging to the same sector and the same asset size category (260 bankrupt and 260 non-bankrupt firms), but the data are from 2003, with one variable (shareholders' funds) from 2002.

3.6 Variables

We have selected a set of 41 initial financial ratios that can be broken up into six categories that best describe company financial profiles: liquidity, solvency, financial structure, profitability, efficiency and turnover.

4 Results

4.1 Selected variables and individual discrimination power

Table 1 ranks the variables by frequency of appearance in the six sets of variables, and table 2 shows the same ranking but only for variables that are identified with the criteria optimized for a neural network. This ranking is compared in table 3, where the variables are ranked by their discrimination ability, as assessed by an F test. In this table, we have added their rank as it appears in the previous table. The first half of table 3 (line 1 to line 21) shows the variables for which the F test reveals the highest discrimination power. This part of the table also contains 13 of the 14 variables selected with the neural network. This result indicates that there is a relationship between a parametric measure of discrimination and all the others we used in this study and which are non-parametric. However, this relationship is fairly rough because the two rankings are quite different. For instance, as table 3 shows, the six variables that are most frequently selected with a neural network (EBITDA/total assets, change in equity position, shareholders' funds/total assets, (cash + marketable securities)/total assets, EBIT/total assets, and cash/current liabilities) are ranked 4th, 20th, 12th, 3rd and 13th respectively. By contrast, variables with high discrimination ability, such as EBITDA/total sales, cash/total assets, current liabilities/total assets, or cash/total debt, are not selected with any selection techniques.

	Number of selections	Rank of appearance in the 6 models					
EBITDA/Total Assets	6	4	4	5	5	6	6
Shareholder's Funds /Total Assets	5	1	1	2	3	7	
Change in Equity Position	5	1	3	3	4	7	
(Cash + Marketable Securities)/Total Assets	4	2	4	4	7		
EBIT/Total Assets	3	2	4	5			
Total Debt/Shareholders' Funds	2	1	2				
Cash/Total Debt	2	3	3				
Cash/Current Liabilities	2	3	5				
EBIT/Total Sales	2	5	6				
Cash/Total Sales	2	7	8				
Net Income/Total Assets	1	1					
Cash/Total Assets	1	1					
Current Assets/Current Liabilities	1	2					
Profit before Tax/Shareholders' Funds	1	2					
(Cash + Marketable Securities)/Total Sales	1	5					
Operating Cash Flow/Total Sales	1	6					
Total Liabilities/Total Assets	1	6					
Accounts Receivable/Total Sales	1	8					

Table 1: Ranking of the variables

Rank		Number of selections
1	EBITDA/Total Assets	3
1	Change in Equity Position	3
3	Shareholder's Funds/Total Assets	2
3	(Cash + Marketable Securities)/Total Assets	2
3	EBIT/Total Assets	2
3	Cash/Current Liabilities	2
7	Current Assets/Current Liabilities	1
7	Accounts Receivable/Total Sales	1
7	Operating Cash Flow/Total Sales	1
7	EBIT/Total Sales	1
7	Net Income/Total Assets	1
7	Cash/Total Sales	1
7	Total Debt/Shareholders' Funds	1
7	Total Debt/Total Assets	1

Table 2: Ranking of the variables selected with a neural network

As a consequence, it appears that using a t or an F test for a selection or pre-selection of the inputs of a neural network is unreliable, as these tests may lead to the choice of useless variables as well as to the removal of variables of great interest.

Such might well have been the case here, with the change in equity position, for which the F test is quite low, even though this variable is in fact relevant according to the neural network. Indeed, selection with a Wilks Lambda removes this variable. But when the value of an F test falls below a certain level, the only other variable selected is accounts receivable/total sales, which is selected only once.

		F	p-val.	Rank[1]
1	EBIT/Total Sales	220,15	0,000	7
2	EBITDA/Total Sales	219,49	0,000	
3	EBIT/Total Assets	218,96	0,000	3
4	EBITDA/Total Assets	213,91	0,000	1
5	Net Income/Total Assets	210,01	0,000	7
6	Shareholder's Funds/Total Assets	207,59	0,000	3
7	Total Debt/Total Assets	202,20	0,000	7
8	Total Debt/Shareholders' Funds	201,14	0,000	7
9	Cash/Total Assets	195,01	0,000	
10	Cash/Total Sales	179,60	0,000	7
11	Current Liabilities/Total Assets	179,32	0,000	
12	(Cash + Marketable Securities)/Total Assets	171,62	0,000	3
13	Cash/Current Liabilities	168,19	0,000	3
14	Cash/Total Debt	150,50	0,000	
15	(Cash + Marketable Securities)/Total Sales	145,63	0,000	
16	Current Assets/Current Liabilities	133,77	0,000	7
17	Quick Ratio	131,30	0,000	
18	Accounts Payable/Total Sales	85,95	0,000	
19	Value Added/Total Sales	68,37	0,000	
20	Change in Equity Position	44,29	0,000	1
21	Operating Cash Flow/Total Sales	28,57	0,000	7
22	Net Operating Working Capital/Total Assets	27,21	0,000	
23	Net Operating Working Capital/Total Sales	21,10	0,000	
24	Operating Cash Flow/Total Assets	19,40	0,000	
25	Long Term Debt/Total Assets	19,32	0,000	
26	Inventory/Total Sales	16,00	0,000	
27	Accounts Receivable/Total Sales	13,38	0,000	7
28	Gross Trading Profit/Total Sales	10,53	0,001	
29	Profit before Tax/Shareholders' Funds	8,97	0,003	
30	Quick Assets/Total Assets	7,13	0,008	
31	Current Assets/Total Sales	4,83	0,028	
32	Financial Expenses/Total Sales	4,04	0,045	
33	Quick Assets/Total Assets	3,47	0,063	
34	Change in Other Debts	2,20	0,139	
35	Total Sales/Shareholders' Funds	2,16	0,142	
36	Labor Expenses/Total Sales	0,62	0,431	

37	Net Income/Shareholders' Funds	0,20	0,651
38	Financial Debt/Cash Flow	0,18	0,669
39	Long Term Debt/Shareholders' Funds	0,17	0,681
40	EBITDA/Permanent Assets	0,11	0,743
41	Total Sales/Total Assets	0,02	0,878

[1] Rank of the variables in table 2

Table 3: Rank of the variables according to an F test

4.2 Model Accuracy

Several techniques are used to assess the prediction accuracy of the models. To define several points of comparison, we have first analyzed to what extent the two groups (i.e., bankrupt vs. non-bankrupt) could be discriminated using variables drawn at random. For each bootstrap sample, we have evaluated the accuracy of discriminant analysis models, logistic regression models and neural network models. This is a powerful way of measuring the distance between a hazard and a deterministic process, and estimating the economy of the latter. Indeed, if the discrepancy is small, we can expect that this process is useless, and the more it increases, the higher its added value. We then calculate the accuracy of models built with the 41 initial variables. This measure can be used to evaluate the performance of pruning strategies, and hence to analyze the relationship between a dimensionality reduction process and model accuracy. In a third and last step, we calculate the performance of the models built with the six final sets of variables and the three selected classification techniques: discriminant analysis, regression analysis and neural network. The aim of this final step is to discover the way modelling techniques may be influenced by a selection procedure and to identify points of compatibility. For instance, is there any difference between two neural models, one built with variables selected by a Wilks Lambda criterion, and the other by a zero or first-order criterion? And what about a logistic model compared to a discriminant model using the same set of variables?

4.2.1 Model accuracy with variables drawn at random

To assess to what extent our samples can be discriminated, we have drawn 50 sets of variables at random and calculated the correct classification rates with bootstrap samples. Table 4 shows the overall results.

	DA	LR	NN
Non-bankrupt	83.22%	82.39%	83.99%
Bankrupt	73.62%	76.88%	78.45%
Total	78.42%	79.64%	81.22%

DA: Discriminant analysis – LR: Logistic regression – NN: Neural network

Table 4: Model accuracy with variables drawn at random

These results demonstrate that it is not easy to discriminate the groups, since the correct classification rate is roughly equal to 80%. However, this rate is not bad if we take into account the fact that the variables are drawn at random, which reveals that the initial 41 predictors demonstrate a good discriminatory ability when applied to our samples.

4.2.2 *Model accuracy with all variables*

Is there any gap in terms of accuracy between a set of randomly selected variables and a set including all variables? The results are shown in table 5. When all variables are taken into consideration, the correct classification rate increases slightly, but the main drawback of this model is its great complexity. In tables 4 and 5, the neural network offers better results than the two other methods.

	DA	LR	NN
Non-bankrupt	93.56%	91.18%	93.60%
Bankrupt	77.72%	81.76%	86.94%
Total	85.64%	86.47%	90.27%

DA: Discriminant analysis – LR: Logistic regression – NN: Neural network

Table 5: Model accuracy with all variables calculated on test samples

4.2.3 *Model accuracy as shown by pairs "modelling method–selection technique"*

We then analyze the relationship between modelling techniques and variable selection methods. The aim is to investigate whether there are any pairs that perform better than others and to study especially the behaviour of a neural network while using sets of variables that were optimized for other methods.

We first measure the accuracy of different combinations « modelling method–selection technique », but only for those for which the evaluation criterion suits the classification technique. We have compared the results of the following six pairs of methods: discriminant analysis–Wilks Lambda, logistic regression–likelihood criterion (with two search procedures), and neural network–zero-order, first-order, and error criteria. As tables 6 and 7 show, the neural network outperforms discriminant analysis and to a lesser extent logistic regression. Indeed, the best result – 93.85% – is achieved with a neural network on the validation samples, followed by that for logistic regression with 90.77% and discriminant analysis with 85.19%.

	DA Wilks Step.	LR Lik. B Step.	LR Lik. F Step.	NN Error B	NN 0 Order B	RN NN 1st Order B
Non-bankrupt	91.20%	93.60%	89.56%	92.78%	91.96%	92.82%
Bankrupt	83.20%	90.42%	88.84%	95.28%	95.22%	92.82%
Total	87.20%	92.01%	89.20%	94.03%	93.59%	92.82%

DA: Discriminant analysis – LR: Logistic regression – NN: Neural network
Lik.: Likelihood – B: Backward – F: Forward – Step.: Stepwise

Table 6: Model accuracy for different pairs « modelling technique–variable selection method » calculated on test samples

	DA Wilks Step.	LR Lik. B Step.	LR Lik. F Step.	NN Error B	NN 0 Order B	RN NN 1st Order B
Non-bankrupt	89.62%	91.15%	88.85%	93.08%	92.69%	91.15%
Bankrupt	80.77%	90.38%	88.46%	94.62%	91.92%	88.85%
Total	85.19%	90.77%	88.65%	93.85%	92.31%	90.00%

DA: Discriminant analysis – LR: Logistic regression – NN: Neural network
Lik.: Likelihood – B: Backward – F: Forward – Step.: Stepwise

Table 7: Model accuracy for different pairs « modelling technique–variable selection method » calculated on validation samples

We then analyze the results obtained when a modelling technique is used with a selection procedure for which the fit is not deemed acceptable. Table 8 displays the results obtained with the set of variables selected with a Wilks Lambda and those selected with a likelihood criterion, and table 9 gives the results calculated with the three sets of variables optimized for a neural network.

Table 8 shows that a variable selection process based on a variance criterion (i.e., Wilks Lambda) leads to bad results; the adequate classification rate of 87.20% achieved with discriminant analysis is slightly lower with the two other methods. The criterion used here relies on assumptions that dovetail with those on which discriminant analysis is founded. It is little wonder then that variables that cannot satisfactorily classify a high percentage of firms with discriminant analysis are unable to provide good results with other methods; the models built with logistic regression and the neural network produce nearly equal results. Therefore, this criterion is clearly ill-suited to non-linear techniques.

	Wilks Lambda Stepwise			Likelihood Backward Stepwise			Likelihood Forward Stepwise		
	AD	RL	RN	AD	RL	RN	AD	RL	RN
Non-bankrupt	91.20%	88.06%	90.02%	87.28%	93.60%	89.68%	87.98%	89.56%	88.08%
Bankrupt	83.20%	79.18%	77.20%	84.84%	90.42%	92.74%	82.42%	88.84%	91.14%
Total	87.20%	83.62%	83.61%	86.06%	92.01%	91.21%	85.20%	89.20%	89.61%

DA : Discriminant analysis – LR : Logistic regression – NN : Neural network

Table 8: Model accuracy according to modelling techniques and two variable selection criteria (Wilks Lambda–Likelihood) calculated on test samples

The sets of variables that are selected with a likelihood criterion lead to less accurate results with discriminant analysis than with logistic regression – 86.06% – as opposed 92.01% with a backward search, and 85.20% as opposed to 89.20% with a forward search. However, with a neural network, the results of these two sets are fairly good – 91.21% and 89.61% – similar to the results obtained with logistic regression. As it happens, the network leads to better results in one case out of two. With the likelihood criterion, logistic and neural models lead to broadly similar results, but this is no longer the case with neural network-based criteria. The error criterion achieved an accuracy of 94.03% compared with 90.00% for logistic regression, and only 84.39% for discriminant analysis. The discrepancy between the results of the three methods is nearly the same with a zero-order criterion, with respective figures for correct classification of 93.59%, 88.01% and 83.60%, but with a first-order criterion there is a decrease, with figures of 92.82 %, 89.19 % and 84.45 %.

	Error Backward			0 Order Backward			1st Order Backward		
	AD	RL	RN	AD	RL	RN	AD	RL	RN
Non-bankrupt	83.38%	90.44%	92.78%	83.20%	86.38%	91.96%	87.06%	88.16%	92.82%
Bankrupt	85.38%	89.56%	95.28%	84.00%	89.64%	95.22%	81.84%	90.22%	92.82%
Total	84.28%	90.00%	94.03%	83.60%	88.01%	93.59%	84.45%	89.19%	92.82%

DA: Discriminant analysis – LR: Logistic regression – NN: Neural network

Table 9: Model accuracy according to modelling techniques and three variable selection criteria (Error, Zero and First-Order) calculated on test samples

Therefore, the neural network leads to far better results than other methods, especially with an error criterion, which is not really surprising, since this criterion is both the evaluation criterion of the variable relevance and the measure of this relevance. This is a very characteristic feature of wrappers, because the inductive algorithm is used directly during variable selection. This result is then consistent with what we might expect. The zero-order criterion's outperformance of a first-order criterion can be put down primarily to chance, as there is no evidence that the former is better than the latter.

Neural models, when developed with appropriate variables, are thus much more reliable than logistic or discriminant models. Nevertheless, logistic models seem to better fit the data than discriminant models, whatever the variables used. In addition, with an error criterion, a logistic model produces 90.00% accuracy, whereas the neural model achieves 94.03%, leaving the logistic model – at 84.38% – in the dust.
The accuracy of a model is in part the result of the intrinsic characteristics of the modelling technique and in part that of the fit between this technique and the variable selection procedure involved in its design. In the field of bankruptcy prediction, all the experiments that have been done with large samples show that both financial ratios and a probability of bankruptcy behave in a non-linear manner. It is precisely for this reason that, as long as this non-linearity cannot be taken into account, it is hardly possible to develop accurate models. Although using a selection criterion that fits logistic regression to design a neural model may be relevant, the choice of a criterion that fits discriminant analysis for the same purpose should not be recommended. It is necessary, at the very least, to consider other solutions.

5 Conclusion

We have demonstrated that a neural network-based model for predicting bankruptcy performs significantly better when designed with appropriate variable selection techniques rather than other types, and particularly those commonly used in the financial literature. Unlike the former, the latter are fast and easy to use, which may account for their under-use. However, a few studies have looked into other techniques, mainly genetic algorithms. So the reasons for the failure of neural network-based variable selection methods to be adopted more widely must be found elsewhere, perhaps in the absence of cross-disciplinary approaches to this particular field. Neural network algorithms are in exactly the same situation: while many types are commonly used in many scientific disciplines, only one is systematically used in the field of corporate finance. And variable selection techniques face the same issue: they come from a field of knowledge that has little to do with corporate finance. Of course, all these results should be confirmed by additional studies in a variety of other settings, such as other samples, types of firms, sectors, and so on, but they point to the need to use relevant variable selection techniques to develop neural models. As it happens, the most recent research papers continue to rely on traditional methods: variables are still selected because they were selected in earlier [33] or as a result of their popularity in the field of financial analysis [34].

We have also demonstrated that there is a relationship between the discrimination ability of a variable, as measured with a t test or an F test, and its ability to be selected by an automatic procedure that relies on other measures, but we have also found a discrepancy in this relationship, which indicates that such statistical tests should not be used alone if the purpose of the selection is to create a neural model.

As a consequence, we may use them – but with extreme caution – to build non-linear models, and if we intend to do so, we would do well to use them in conjunction with other techniques.

References

[1] E. I. Altman, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, 23:589-609, 1968.

[2] J. A. Ohlson, Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, 18:109-131, 1980.

[3] E. K. Laitinen and T. Laitinen, Bankruptcy Prediction: Application of the Taylor's Expansion in Logistic Regression, *International Review of Financial Analysis*, 9:327-349, 2000.

[4] M. C. Odom and R. Sharda, A Neural Network Model for Bankruptcy Prediction, *proceedings of the IEEE International Joint Conference on Neural Networks* (IJCNN 2001), pages 163-168, San Diego (California), 1990.

[5] M. E. Zmijewski, Methodological Issues Related to the Estimation of Financial Distress Prediction Models, *Journal of Accounting Research*, 22:59-82, 1984.

[6] C. V. Zavgren, Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis, *Journal of Business Finance and Accounting*, 12:19-45, 1985.

[7] R. L. Wilson and R. Sharda, Bankruptcy Prediction Using Neural Networks, *Decision Support System*, 11:545-557, 1994.

[8] K. C. Lee, I. Han and Y. Kwon, Hybrid Neural Network Models for Bankruptcy Predictions, *Decision Support Systems*, 18:63-72, 1996.

[9] N. Kumar, R. Krovi and B. Rajagopalan, Financial Decision Support with Hybrid Genetic and Neural Based Modeling Tools, *European Journal of Operational Research*, 103:339-349, 1997.

[10] Z. R. Yang and R. G. Harrison, Analysing Company Performance Using Templates, *Intelligent Data Analysis*, 6:2002.

[11] A. Charitou, E. Neophytou and C. Charalambous, Predicting Corporate Failure: Empirical Evidence for the UK, *European Accounting Review*, 13:465-497, 2004.

[12] P. P. M. Pompe and J. Bilderbeek, The Prediction of Bankruptcy of Small- and Medium-Sized Industrial Firms, *Journal of Business Venturing*, 20:847-868, 2005.

[13] P. Leray and P. Gallinari, Feature Selection with Neural Networks, *Behaviormetrika*, 26:145-166, 1998.

[14] K. Y. Tam and M. Y. Kiang, Managerial Applications of Neural Networks: The Case of Bank Failure Predictions, *Management Science*, 38:926-947, 1992.

[15] N. Wilson, K. S. Chong and M. J. Peel, Neural Network Simulation and the Prediction of Corporate Outcomes: Some Empirical Findings, *International Journal of the Economics of Business*, 2:31-50, 1995.

[16] J. E. Boritz and D. B. Kennedy, Effectiveness of Neural Network Types for Prediction of Business Failure, *Expert Systems with Applications*, 9:503-512, 1995.

[17] R.C. Lacher, P. K. Coats, S. C. Sharma and L. F. Fant, A Neural Network for Classifying the Financial Health of a Firm, *European Journal of Operational Research*, 85:53-65, 1995.

[18] C. Serrano-Cinca, Feedforward Neural Networks in the Classification of Financial Information, *European Journal of Finance*, 3:183-202, 1997.

[19] T. Laitinen and M. Kankaanpaa, Comparative Analysis of Failure Prediction Methods: The Finnish Case, *European Accounting Review*, 8:67-92, 1999.

[20] Z. R. Yang, M. B. Platt and H. D. Platt, Probabilistic Neural Networks in Bankruptcy Prediction, *Journal of Business Research*, 44:67-74, 1999.

[21] P. C. Pendharkar, A Threshold-Varying Artificial Neural Network Approach for Classification and its Application to Bankruptcy Prediction Problem, *Computers and Operations Research*, 32:2561-2582, 2005.

[22] C. H. Wu, G. H. Tzeng, Y. J. Goo and W. C. Fang, A Real-Valued Genetic Algorithm to Optimize the Parameters of Support Vector Machine for Predicting Bankruptcy, *Expert Systems with Applications*, 32:397-408, 2007.

[23] B. Back, T. Laitinen and K. Sere, Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis and Genetic Algorithms, Technical Report, Turku Centre for Computer Science, 1996.

[24] J. Wallrafen, P. Protzel and H. Popp, Genetically Optimized Neural Network Classifiers for Bankruptcy Prediction – An Empirical Study, *proceedings of the 29th Hawaii International Conference on System Sciences* (HICS 1996), pages 419-426, Maui (Hawaii), 1996.

[25] K. Kiviluoto, Predicting Bankruptcies with the Self-Organizing Map, *Neurocomputing*, 21:191-220, 1998.

[26] R. S. Sexton, R. S. Sriram and H. Etheridge, Improving Decision Effectiveness of Artificial Neural Networks: A Modified Genetic Algorithm Approach, *Decision Sciences*, 34:421-442, 2003.

[27] A. Brabazon and P. B. Keenan, A Hybrid Genetic Model for the Prediction of Corporate Failure, *Computational Management Science*, 1:293-310, 2004.

[28] E. W. Tyree and J. A. Long, Bankruptcy Prediction Models: Probabilistic Neural Networks versus Discriminant Analysis and Backpropagation Neural Networks, Working Paper, Department of Business Computing, City University, 1996.

[29] C. Charalambous, A. Charitou and F. Kaourou, Application of Feature Extractive Algorithm to Bankruptcy Prediction, *proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks* (IIE-IJCNN 2000), pages 303-308, July, Como (Italy), 2000.

[30] T. K. Sen, P. Ghandforoush and C. T. Stivason, Improving Prediction of Neural Networks: A Study of Two Financial Prediction Tasks, *Journal of Applied Mathematics and Decision Sciences*, 8:219-233, 2004.

[31] I. Bose and R. Pal, Predicting the Survival or Failure of Click-and-Mortar Corporations: A Knowledge Discovery Approach, *European Journal of Operational Research*, 174:959-982, 2006.

[32] M. Yacoub and Y Bennani, HVS: A Heuristic for Variable Selection in Multilayer Artificial Neural Network Classifier*, proceedings of the International Conference on Artificial Neural Networks and Intelligent Engineering* (ICANNII 1997), pages 527-532, Saint-Louis (Missouri) 1997.

[33] S. Jones and D. A. Hensher, Predicting Firm Financial Distress: A Mixed Logit Model, *Accounting Review*, 79: 1011-1038, 2007.

[34] Z. Zhu, H. He, J. A. Starzyk and C. Tseng, Self-Organizing Learning Array and its Application to Economic and Financial Problems, *Information Sciences*, 177:1180-1192, 2007.