Failure prediction models for Tunisian companies: Development and comparison between multivariate discriminant analysis, logistic regression, multilayer perceptron and support vector machines

Abstract

In this study, we try to develop a model for predicting corporate default based on a multivariate discriminant analysis (ADM), logistic regression (LR), multilayer perceptron (MLP) and a support vector machine (SVM). The four models are applied to the Tunisian cases. Our sample consists of 212 companies in the various industries (106 'healthy' companies and 106 "distressed" companies) over the period 2005-2010. The results of the use of a battery of 87 ratios showed that 16 ratios could build the ADM model, 12 ratios for the LR method and 9 ratios for the MLP method. Similarly, the results obtained show that liquidity and solvency have more weight than profitability and management in predicting the distress. Despite the slight superiority of the results provided by the MLP model, the results provided by the four models are good either in terms of correct percentage of classification or in terms of stability of discriminating power over time and space.

Keywords: distressed firms, forecasting model, multivariate discriminant analysis, logistic regression, neural network, multilayer perceptron, support vector machine

JEL: G17 - G33 - G34

1. Introduction :

The diagnosis of default risk has experienced significant development using both classical statistical methods as methods from artificial intelligence that analyze the financial situation from a given set of ratios. In the present work, we will estimate and compare the discriminating power of the MDA, the LR, the MLP and the SVM models. The first and the second are a classic statistical method, while the third and the fourth belongs to the methods from artificial intelligence.

The principle is relatively simple. With the financial characteristics described using ratios and a sample of companies that cover both "healthy" companies and firms "failing", the objective is to determine the best combination of ratios to differentiate the two business groups. Based on this combination, we will estimate the percentage of correct classification of each method. To achieve this goal, this article will address in the first section, the methodology through the constitution samples, presentation and justification of the four selected models. The estimate of the discriminating power of the four models in the time and space will be of the second section. The third section will compare the results provided by the different methods.

2. The methodology

In this work, we will use the four models for the purpose of forecasting corporate failures, and then test their validity in time and in space. However, it is above all, to address the composition of samples, the selection of variables, presenting the models and demonstrate their usefulness.

2.1 The constitution of samples

The choice of the sample posed us serious problems. Indeed, the implementation of the models assumes the existence of two business groups « healthy » and « distressed ». The selection of the reference population leads to a choice between two alternatives:

- Constitute a sample the widest possible, which includes companies from different industries, size, geographical location and economic environments.
- Choose a reference population to ensure homogeneity of the sample, at the expense of its size.

In practice, and according to most studies [Beaver (1966), Altman (1968), Edmister (1972)], we adopted the option of a larger sample affecting several sectors. Our sample consists of 212 Tunisian companies in the various sectors (which will be discussed below), (106 "healthy" companies and 106 "distressed" companies) over the period 2005-2010.

The "healthy" companies were selected from the Tunisian stock exchange and among statutory accountants. While "distressed" companies come from the office of assistance to companies in difficulty, which sits at the Ministry of Industry. The selection of firms in difficulty was based on the following criteria:

- Be suspension of payments for at least six months
- Have very serious social problems,
- Must be identified by statutory auditors, National Social Security Fund or fiscal institutions

From this basic sample, and referring to the approach of Platt and Platt, (1991); Altman et al, (1994); Bardos (1998a) and Varetto (1998), it was possible to set up two sub-samples:

- A first, called "Initial" sample consisting of 152 companies, 76 "healthy" and 76 "distressed". We'll take the last three years of the same companies to form three sub-samples we call "Initial one year prior to distress," "Initial two years before distress" and "Initial three years prior to distress." these sub-samples used to develop the model and to test its validity in time.
- A second sample, called "Control" sample, composed of 60 other companies, 30 "healthy" and 30 "distressed". From the last three years of these companies, we will establish three sub-samples that we call "control one year prior to distress," "Control two years prior to distress" and "Control three years prior to distress." These sub-samples are designed to test the validity of the model in space.

Companies belonging to both sample of "healthy" and the "distressed" companies are distributed between the different sectors as follows:

Table 1: The distribution of the companies between the different sectors

Sectors \ Companies	Healthy	Distressed
Textile , Clothing and Leather Industries	28	23
Food-processing industry	23	19
Various industries	19	19
Industries of Building materials, Ceramic and Glass	13	18
Mechanical engineering industries, Metallic, Metallurgical and Electric	11	13
Services (hotel)	8	9
Chemical industries	4	5
Total	106	106

2.2 The choice of default indicators:

In the absence of a theory of business distress, the choice of indicators is completely subjective. Indeed, it is based on experience and intuition of the one who develops the model. Generally, this choice often results from previous choices; this is to say the choice of all first authors of reference. In order that our work be as exhaustive as possible, we chose 87 ratios contained in the works of Ramser and Foster (1931), Fitzpatrick (1932), Winakor and Smith (1935), Merwin (1942), Beaver (1966), Altman (1968), Deakin (1972) Edmister (1972), Blum (1974); Altman et al (1977), Taffler (1983) and Zmijewski (1984).

2.3 Overview and principle of the Multivariate Discriminant Analysis model

2.3.1 Literature review

The objective of the multivariate discriminant analysis is to determine a function called Z-score, which is none other than the linear combination of explanatory variables retained. This combination must be able to distinguish at best the two groups through the identification of the level of risk of each company. The linear discriminant analysis requires the observance of two assumptions that of the multi-normality and that of the homoscedasticity. The first assumes that the accounting variables used follow a normal law; the second requires the equality of matrices variance-covariance for the two categories of failing and healthy firms. To circumvent the problem of homoscedasticity, some authors have made use of quadratic discriminant analyzes, which require only the hypothesis of multi-normality of ratios (Lachenbruch and al, 1973; Marks and Dunn, 1974; Rose and Giroux, 1984). Only we found that they are always less efficient than the linear analysis and this mainly for two reasons. First, the absence of multi-normality ratios is much more harmful to the effectiveness of the quadratic analysis than to those of the linear analysis (Lachenbruch, 1975); secondly, even in the case of non-respect of the hypothesis of multi-normality, quadratic discriminant analysis is efficient only if it is applied to a sample of large size.

2.3.2 MDA model principle

Developed by Altman (1968), multivariate discriminant analysis (MDA) assumes the existence of two groups of firms each with its own indicators of its financial situation. For these two groups then we can determine a discriminant function that is sharing in the better the set of firms in two separate groups. This discriminant function is a linear combination of the most relevant indicators, to differentiate the two groups we associate a score Zj has each company j.

$$Z_j = \alpha_0 + \alpha_1 x_{1j} + \alpha_2 x_{2j} + \cdots + \alpha_n x_{nj} + c$$

Avec:

x_{nj} : The value taken by the indicator x_n of the enterprise j

α_i : The numerical adjustment coefficients.

c : A constant

The classification in one or the other of the groups is done by comparing the value of the score Zj with a critical value Z*. We must however, during the drafting of the discriminant function maximize the intergroup variance and minimize the intra-group variance.

During this discrimination, there may be two types of errors:

The error of first species: classify a failing firm with sound.

The error of second species: classify a healthy firm with failing.

The cost associated with the error of first species is very different from the cost associated with the error of second species. In effect, the first cost is the one that will bear a creditor in the event of failure of its debtor. While the second cost corresponds to the opportunity cost, that is to say, the gap between the pay that a creditor might have been able to collect on the loan refused and the rate of return offered by the use of these funds.

The proportion of correct classification allows you to judge the quality of the discriminant function.

2.4 Overview and principle of the logistic model

2.4.1 Literature review

Logistic regression, viewed as a generalization of linear discriminant analysis, has been introduced by Day & Kerridge (1967), Cox (1970), and developed by Anderson (1972, 1982), Martin (1977), Olshon (1980) who was the pioneer in the use of logistic regression in the domain of prediction of business distress. Among the major works that have used this method we can cite Mensah (1984), Albert & Lesaffre (1986), Aziz &*al* (1988), Bardos (1989), Burgstahler & *al* (1989), Flagg & *al* (1991), Platt & Platt (1991), Zopounidis (1995), Bardos et Zhu (1997), Mossman & *al* (1998) and more recently Altman & *al* (2005), Jones & Hensher (2004, 2007, 2007a), Zeitun & *al* (2007), Li & *al* (2011), Ahn & al (2011), Tserng & al (2011), kim & Kang (2012), Serrano-cenca & al (2013) et Wang & al (2014), Yu & al (2014).

As in multiple linear regression, it is relates to estimate parameters of model, to measure its adequacy (quality of adjustment) and to deduce the significance and the interpretation of the estimated parameters. Logistic regression is an econometric technique with a dichotomous dependent variable y_i, representing the state of the company that takes:

- The value 1 if the company is "distressed"
- The value 0 if the firm is "healthy".

This type of regression allows to determinate the probability that a firm is classified in the group of « healthy » or the group of « distressed ».

In general, from a sample of base and a set of ratios, we will proceed as follows:

- Check the distribution normality of selected ratios by eliminating those not responding to the corresponding test.
- Examine the individual discriminating power of these ratios by classifying them by categories.
- Evaluate the existing correlations between the ratios by eliminating those that are redundant.
- Observe the discriminating power of different combinations and select by iteration the combination that offers the best correct percentage of classification with the lowest cost of the first kind, that is the one that provides the best value:

  intergroup dispersion / intragroup dispersion.

2.4.2 Logistic model principle

We have:

y_1, y_2,......y_n : random variables, called dependent variables, each taking the value 1 or 0, values that correspond to groups G1 and G2 to discriminate.

x_1, x_2,........x_J : the components of a multi-dimensional vector X = (x_1, x_2,........x_J) and that represent random variables called explanatory or independent variables.

(β) = (β_0, β_1,..........β_J): are the unknown coefficients of the model to be estimated.

The idea is to build a model linking π(x) = p[Y=1/ X] (the probability that Y = 1 given X).

With :

$$\text{probability of default } [\pi(x)] = P(Y=1/\,X=x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+.......+\beta_K x_K)}}$$

and

$$\text{probability of non-default } [1-\pi(x)\ ] = P(Y=0/\,X=x) = 1-\frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+.......+\beta_K x_K)}}$$

2.5 Overview and principle of the MLP model

2.5.1 Literature review

Warren McCulloch and Walter Pitts (1943) were the pioneers in the field of neural networks by presenting the "formal neuron" as the first attempt to imitate the functioning of the human brain. In 1949, Hebb presents the first rule of learning neural networks, a move which allowed, later, Rosenblatt (1958) to propose the first algorithm of learning making the adjustment of the parameters of a neuron possible.
After publishing their book "perceptron" in which Papert (1968) shows the limits of monolayer neural networks, connectionism resumed in the 1980s after a long period of hibernation. Indeed, the work of Hopfield (1982), who proposed associative neural networks, induced an interesting renaissance of neural networks.
Rumelhart, Hinton, and Williams (1986) published their work on the error-retroagitation algorithm that optimizes the parameters of a multi-layered neural network. From this date, research on neural networks has expanded greatly and has been integrated in all areas.
The use of artificial neural networks (ANNs) in failure prediction dates back to 1990. Indeed Odom et al. (1990) were the pioneers in the field.
According to Odom et al. (1990), Raghupathi (1991), Salchenberger (1992), Tam (1997) and Altman (1994), the multilayer perceptron with gradient retro-extension algorithm (RPG) learning remains the reference in failure anticipation.
The use of a learning algorithm other than the RPG technique in the context of the implementation of a multilayer perceptron stems from one of the limits of this type of network, namely its blocking on the local minima.
In the area of failure prediction, the multilayer perceptron (PMC) represents the reference network [Poddig (1995)]. However, there are other networks of artificial neurons other than the PMC such as the radial base function networks (FBR) and self-organizing maps of Kohonen.
The operating principle of the multilayer perceptron is as follows:

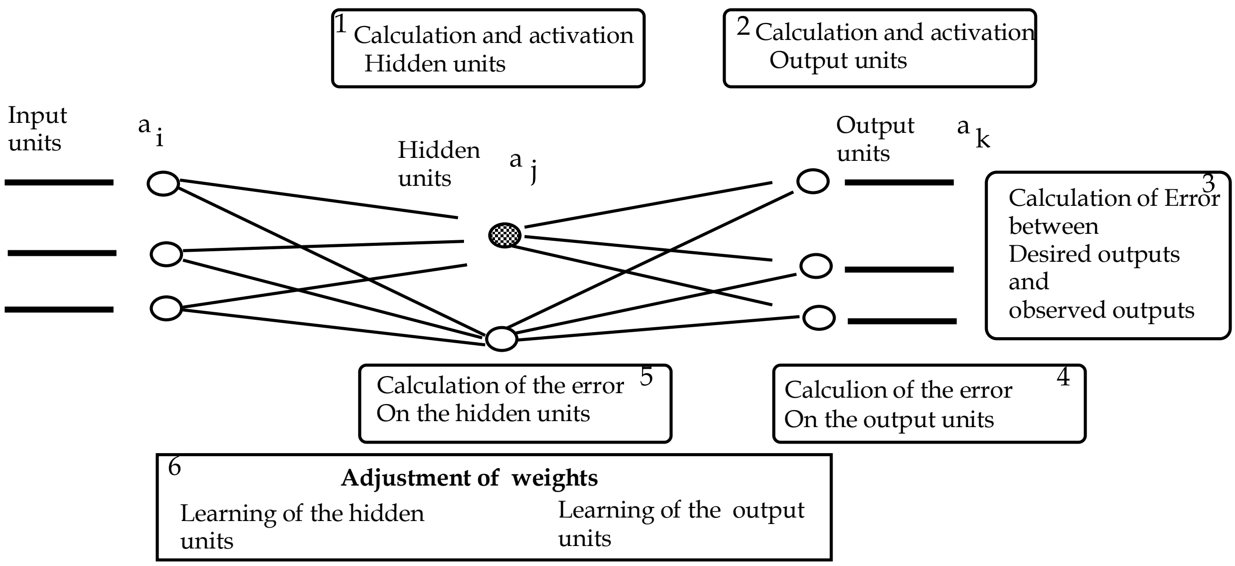

Fig 1 : The multilayer perceptron: learning by backpropagation of the error

2.5.2 MLP model principle

The neural networks allowing for estimating a function f such that $f : x \to y$ with $x^T = [x_1, x_2......x_E] \in IR^E$ if $y \in IR^s$ are regressions.

If $y \in [c_1, c_2......c_s]$ then it is classification. In this case, we should have as many output neurons as classes.

The desired outputs are of the form:

$y^{d^T} = [0,0,1.....0,1.....,0]$

When estimating the function f we must identify the connection weights between neurons.

Now let us recall before all the principle of the formal neuron [Me Cultoch and Pitts, 1943].

Let E inputs x_i et y outputs. The sum of inputs x_i weighted by w_i is equal to α

With : $$\alpha = \sum_{i=1}^{E} w_i x_i + b = \sum_{i=0}^{E} w_i x_i \quad \text{avec } x_0 = 1$$

Let φ an activation function that can be linear where we have:

$$y = \varphi(x) = \varphi\left(\sum_{i=0}^{E} w_i x_i\right)$$

If φ is linear the separator is a hyperplane.

If φ is not linear the separator is a hyperbola of dimension E.

We distinguish different activation functions that determine the activation threshold of a neuron.

- Identity function: $\varphi(x) = x$
- Heaviside function: $\varphi(x) = 0$ if $x < 0$ et $\varphi(x) = 1$ if $x \geq 0$
- Sigmoid function: $\varphi(x) = \frac{1}{1+e^x}$
- Hyperbolic tangent function : $\varphi(x) = \frac{e^x - e^x}{e^x + e^{-x}} = \frac{e^{2x} - 1}{2^{2x} + 1}$
- Average function (Gaussian = normal)

The available data are as follows:

We have a base of N couples $\{x(n); y^d(n)\}$.

With :

$x(n)$: Observations on the independent variables

$y^d(n)$: The desired outputs in value for example n

$X \in IR^{ExN}$

$$X = \{X(n)\} = \left\{\begin{bmatrix} x_1(n) \\ x_2(n) \\ \\ x_E(n) \end{bmatrix}\right\} = \begin{bmatrix} x_1(1)\ x_1(2)......x_1(N) \\ x_1(2) \qquad\qquad x_2(N) \\ \\ x_E(1) \qquad\qquad x_E(N) \end{bmatrix}$$

$$y^d \in \{y^d(n)\} = \left\{\begin{bmatrix} y_1^d(n) \\ \\ \\ y_s^d(n) \end{bmatrix}\right\} = \begin{bmatrix} y_1^d(1)\ y_1^d(2)......y_1^d(N) \\ y_2^d(1) \qquad\qquad \\ \\ y_s^d(1) \qquad\qquad y_s^d(N) \end{bmatrix}$$

With E : the number of input variables

S : the number of neurons in the output layer

We will assume a multilayer network with inputs (E inputs), a hidden layer with j neurons and an output layer of S neurons.

Are: $W^1 \in IR^{J\,x\,E}$ the matrix of connection weights between inputs $\left(X(n) \in IR^E\right)$ and the J neurons of the hidden layer.

$W^2 \in IR^{S\,x\,J}$ the matrix of connection weights between the J neurons of the hidden layer and the S neurons of the output layer.

So :

$$W^1 = \{W_{ji}\} = \begin{bmatrix} W_{11}\ W_{12} & & W_{1E} \\ & W_{21} & \\ & & \\ W_{J1} & & W_{JE} \end{bmatrix} \qquad (1)$$

$$W^2 = \{W_{sj}\} = \begin{bmatrix} W_{11}\ W_{12} & & W_{1J} \\ & W_{21} & \\ & & \\ W_{S1} & & W_{SJ} \end{bmatrix} \qquad (2)$$

x_1		1		1	
x_2		2		2	
	➔		➔		➔ $y(n) \in IR^S$
x_E		J		S	
$X(N) \in IR^E$		$W^1 = \{W_{ji}\}$		$W^2 = \{W_{sj}\}$	
Inputs		hidden layer composed of J neurons		output layer composed of S neurons	

Let φ^1 and φ^2 two nonlinear activation functions.

φ^1 of the sigmoid type relating to the connections of the hidden layer and φ^2 of the soft max type relative to the connections of the output layer.

Let $Z(n) \in IR^J$ an intermediate variable.

α_j^1 : the weighted sum of the connections between all the E inputs and the jth neurons in the hidden layer.

α_s^2 : the weighted sum of the connections between the J hidden neurons and the sth output neuron.

We then:

$$\alpha_j^1 = \sum_e W_{je} X_e \quad \text{and} \quad Z_j = \varphi^1\left[\left(\alpha_j^1\right)\right] \qquad (3)$$

$$Z_j = \varphi^1\left[\sum_e W_{je} X_e\right] \qquad (4)$$

Once we have finished with the modeling of the passage from the input neurons to the hidden neurons, we will approach the second half of the process, which relates to the passage of the hidden neurons to the output neurons. Indeed:

$$\alpha_s^2 = \sum_j W_{sj} Z_j \quad \text{and} \quad y_s = \varphi^2\left[\left(\alpha_s^2\right)\right] \qquad (5)$$

$$= \varphi^2 \left[\sum_j W_{sj} Z_j \right] \quad (6)$$

then $$y_s = \varphi^2 \left[\sum_j W_{sj} \varphi^1 \left(\sum_e W_{je} X_e \right) \right] \quad (7)$$

2.6. Overview and principle of the SVM model

2.6.1. Literature review

The support vector machine (SVM), were developed in the 1990's based on theoretical work of Vladimir Vapnik (1960-1990), mainly the theory of Vapnik-Chervonenkis [Validimir Vapnik and Alexey Chervonenkis (1968)] or statistical learning theory which was then applied in several works such as those of Gunn (1998), Taylor and al (2004), Wei and al (2007). The advantages of this new technique namely its ability to work with large data, the presence of theoretical basis and relevance of empirical results have allowed it to be quickly adopted and applied in many fields such as bio- IT, research information and finance.
The method has its reputation and popularity following the work of Boser et al (1992) and Vapnik (1995). It solves problems of nonlinear discrimination by equating the classification problem to a problem of quadratic optimization.
To the wide margin separators (SVM) are a set of supervised learning techniques to solve problems of discrimination. In the case of discrimination of a dichotomous variable, they seek to determine the maximum margin hyper plane able to classify or separate correctly the data while being far away as possible from all observations.
Indeed, the method is based on two main ideas. The first relates to the concept of maximum margin which is none other than the distance between the separation border and the nearest samples called Support Vector. The second relates to the notion of the kernel function which, in the case where the data are not linearly separable, to expand the representation space of the input data until the infinite to increase the likelihood of find a possible linear separation (separating hyper-plane).

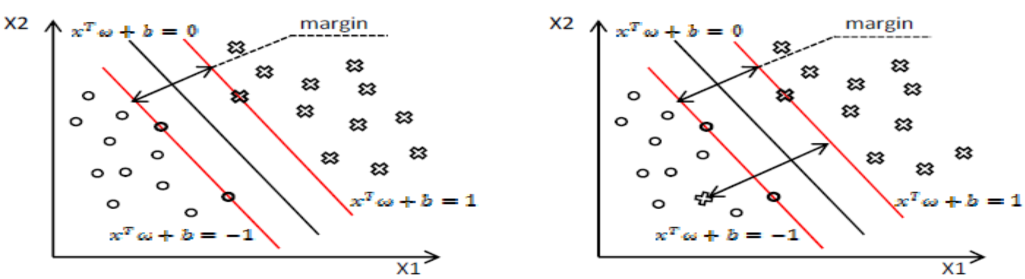

Fig 2 : linearly separable hyperplane (left) and non-linearly separable (right).

2.6.2. SVM model principle

SVM will be used in this case to resolve the problem of discrimination between "healthy" companies and firms in "difficulty". The resolution of this classification problem requires the construction of a function h which allows to match each input vector x to an output y, hence
y = h (x)
If we limit ourselves to a binary discrimination then $y \in \{-1,1\}$ the input vector x being in a space X with a scalar product, one can take $X = IR^N$.

In the case where the discrimination is linear, the function is obtained by the linear combination of the input vector $x = (x_1, x_2, x_3 ... x_N)^T$ weighted by a weight vector $w = (w_1, w_2, w_3 ... w_N)$ hence $h(x) = w^T.x + w_0$.
If $y = h(x) \geq 0$, then x (in our case the company) is classified in the group of "distressed" companies.
If $y = h(x) \pi 1$ then x is classified in the group of "healthy" companies.
The decision border $h(x) = 0$ is a separating hyperplane. As part of SVM, the objective of the supervised learning algorithm is to learn the function h (x) for a sample or a training set

$$\{(x_1, l_1), (x_2 l_2), ... (x_p, l_p)\} \subset IR^N \times \{-1,1\}$$

P: the size of the training set
N: the dimension of the input vectors
l_p : Are the labels
Vapnik (1982) showed the existence of a single optimal hyperplane whose equation is:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* l_k (x.x_k) + w_0$$

With α_k^* : optimal Lagrange multipliers
In the case where the discrimination problems are not linearly separable, the foregoing formulation is no longer valid. To remedy this, the SVM using a kernel function which comprises in changing the original space into a higher dimensional space or even infinite to increase the probability of having a linear separator. The formulation is appling to the input vectors X a transformation ϕ. So, the arrival space ϕ (x) is called feature space. rédescription
The only difference here compared to the linear discrimination instead of the initial space, we will use the redescription space to search the hyper plane whose equation is:

$$h(x) = w^T \phi(x) + w_0$$

Under constraint $l_k h(x_k) \phi 0$
The expression of the equation for the optimal hyper plane is:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* l_k K(x_k, x) + w_0$$

K: the kernel function defined by

$$K(x_i, x_j) = \phi(x_i)^T . \phi(x_j)$$

Thus, the data is linearly or nonlinearly separable; the SVM method allows determining the optimal hyper plane able to discriminate between the two categories of companies.

3. Estimation of the Multivariate Discriminant Analysis model parameters:

This part will be devoted to the estimation of the discriminating power of A.D.M. Both in time through its application on the initial sample two and three years before the failure and in space by applying it on the three control sub-samples.
First, we will use 87 explanatory variables (see Appendix 1). To determine the weighting coefficient of each exogenous variable in our discriminant function, we used a software frequently applied in the analysis of the data, the software S.P.S.S.
Applying this software to our sample, we obtained the following results: (see appendix 2)
If we take into account the significance (see Appendix 2) and the redundancy (variance-covariance matrix) of the explanatory variables of the model for a degree of significance of

1%, we must retain only the 16 ratios that will constitute the explanatory variables of the model to be estimated. The estimate by M.D.A. gives us the following results:

Table 2 : Eigen values

Function	Eigen values	% of variance	% cumulated	Canonical correlation
1	8,669[a]	100,0	100,0	,947

a. The first 1 canonical discriminant functions were used for the analysis.

Table 3 : Coefficients of canonical discriminant functions

	Function
	1
R6	2,891
R7	-9,988
R15	5,942
R16	-,023
R19	-3,389
R26	1,855
R33	-,927
R40	8,230
R58	-2,510
R61	-,027
R73	-,631
R78	-,210
R79	8,369
R83	-,493
R84	-4,234
R85	,024
(Constant)	,225

Non-standardized coefficients

The last 16 ratios will represent the explanatory variables of our final model:

Z = 2,8907 R6 - 9,9883 R7 + 5,9415 R15 - 0,0225 R16 - 3,3888 R19 + 1,8554 R26 - 0,9273 R33 + 8,23 R40 - 2,5098 R58 - 0,0274 R61 - 0,6312 R73 - 0,2096 R78 + 8,3685 R79 - 0,4930 R83 - 4,2335 R84 + 0,0242 R85 + 0,2247

With:

Table 4 : The Ratios Retained by the M.D.A. Method

Ratios	Formulas
R_6	Permanent Capital / Total Balance Sheet
R_7	Current assets / Total assets
R_{15}	Equity / Total assets
R_{16}	Working capital / Cash flow from operations
R_{19}	Short-Term Debt / Total Liabilities
R_{26}	Amortization of Capital Assets / Gross Fixed Assets
R_{33}	current assets (excluding stocks) / current liabilities
R_{40}	current assets (excluding stock) / Total assets
R_{58}	receivables / Total assets
R_{61}	Medium and long-term debt / Cash flow

R_{73}	Net income / Turnover
R_{78}	Size Ln (Total assets)
R_{79}	Total Liabilities / Total Assets
R_{83}	Value Added / Total Liabilities
R_{84}	Total Fixed Asset / Total assets
R_{85}	Working capital / Cash-flow

In the prediction equation retained by the discriminant analysis, we note the presence of several ratios that have been selected as explanatory variables in previous studies. Indeed,

Table 5: The Presence of Several Explanatory Ratios in Previous Studies

Ratio	Authors
R_6	Conan and Holder (1979) ; Holder and al (1984)
R_7	Deakin (1972) ; Taffler (1982) ; Holder and al (1984)
R_{15}	Le crédit commercial de France (1995)]
R_{19}	Beaver (1966) ; Plat and Plat (1991)
R_{26}	Altman and al (1984) ; le modèle du C.E.S.A. (1974)
R_{33}	Deakin (1972) ; Edmister (1972) ; Houghton (1984) ; Burgstahler and al (1989) ; Michalopoulas and al (1993)
R_{40}	Conan and Holder (1979)]
R_{61}	Conan and Holder (1979) ; Bardos (1984)
R_{79}	Deakin (1972) ; Rose and Giroux (1984) ; Burgstahler and al (1989) ; Michalopoulas and al (1993) ; Altman and al (1994)

The presence of these ratios in the models makes it possible to cover all aspects of the company, its solvency, its liquidity level, its financial autonomy, its financial structure, the degree of maturity of these debts and the degree of aging of this equipment.
The global significance test used in the MDA regression is the chi-square with k degrees of freedom (K is the number of explanatory variables in our case k = 16). If the critical probability is lower than the level of significance we have set, we can consider that the model is globally significant. In our model, the likelihood ratio statistic (chi-square) is equal to 322,187; the associated critical probability is zero. The model is thus globally very significant, there is indeed a relationship between the explanatory variables and the variable to be explained.

Table 6 : Lambda of Wilks

functions test	Wilk’s Lambda	Chi-square	ddl	Signification
1	,103	322,187	16	,000

Once the overall significance of the chosen model is demonstrated, our work now consists in verifying the discriminatory capacity and the stability of the results presented by the M.D.A. both in time using the initial samples one year, two years and three years before the failure than in the space using the control samples.

3.1 Estimation of the model discriminatory power one year before distress:

The estimation of the MDA model on the original sample, one year prior distress, shows that in the "healthy" firms group, the model classifies all "healthy" firms in their original group correctly.

In the distressed companies group, that interests us the most, we find no firm misclassified, so the model classifies successfully both companies "healthy" as "distressed".

Table 7 : Estimates of initial sample one years before distress:

Classification table[b]

Observations			Predicted		
			Selected observations[a]		
			Y		Percentage correct
			0	1	
Stapes 1	Y	0	76	0	100,0
		1	0	76	100,0
	global Percentage				100,0

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1

As far as the error Type I cost is much higher than that of an error type II [about 1 to 20 in Altman and al (1977)], then it seems more appropriate to judge the quality of the model on the base of the correct percentages of classification, in general, and of the error type I rate that it induces, in a particular way. These results "appear" as a whole interesting because they have the advantage of providing a combination of ratios based on which one can make a diagnostic of the company.

We say "appear interesting" because we should not judge the model before testing the performance over time (testing the model on the same companies but for different periods of time, two years and three before distress) and in space (testing the model on a control sample consisting of companies other than those in the sample of origin).

3.2 Validation of the model discriminatory power over time:

3.2.1 For the same companies two years before distress

The validation of model on exercises that come two years before distress gives the results in in the following table.

Table 8: Estimates of initial sample two years before distress :
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
			0	1	correct Percentage	0	1	Correct percentage
Stape 1	Y	0	76	0	100,0	76	0	100,0
		1	0	76	100,0	5	71	93,4
	global Percentage				100,0			96,7

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model correctly classifies all « healthy » firms in their original group. In the « distressed » firms group, there are five firms misclassified, so the firms are considered as "healthy" when they are actually distressed. The model retains thus its discriminatory power, since the percentage of correct classification varies by only 3.3% from 100% to 96.7%, the error type I increases from 0 to 6.58%, while the error type II remains zero.

3.2.2 For the same companies three years before distress:

By distancing yet the period between the date of the estimates and the date of the failure of a period for an additional year, the application of the multivariate discriminant analysis provides the results presented in the table 9

Table 9: Estimates of initial sample three years before distress :
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		correct percentage	Y		
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	75	1	98,7
		1	0	76	100,0	6	70	92,1
	global Percentage				100,0			95,4

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In passing from one year to three years before the failure, the method loses more of its accuracy. In fact, the percentage of correct classification increased from 100 per cent to 95.4%. The error of first species (error type I) jumped from 0 per cent to 7.89 %. In effect, the method class 6 companies as "sound", then they really are "faulty".
The error of second species increased from 0 % to 1.32 %. Actually, the discriminant analysis multivariate range a single company in the group of "failed" when it is really "healthy".

Table 10: Results of estimation in the time

	1 year before distress	**2 years before distress**	**3 years before distress**
% of correct classification	100 %	96. 71 %	95.4 %
% of classification error	0 %	3. 29 %	4.6 %
% of error type I	0 %	6. 58 %	7.89 %
% of error type II	0 %	0 %	1.32 %

Indeed, we notice that for the model used, the percentage of the error Type I varied only by 7.89% between the first and third years before distress. Furthermore, we find that the correct percentage of classification decreased only by 4.6% (it goes from 100% to 95.4%).
For our model, the most interesting element, in addition to its high correct percentage of classification, it is the weakness of the error Type I whose cost is higher. Concerning the error type II, we see that it remains ≤ 1.32%.

3.3 Validation of the model discriminatory power in space:

To test the discriminatory power of the model in space, we use a control sample consisting of two new groups. The first contains the distressed firms while the second contains "healthy" companies, each list 30 firms. The model will be tested on companies other than those that were originated. The application of our MDA model on these samples gives us the estimates presented in the table 11

Table 11: Estimates of initial and control samples one year before distress :
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		Correct percentage	Y		Correct percentage
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	29	1	96,7
		1	0	76	100,0	3	27	90,0

global Percentage			100,0			93,3

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model classifies only one firm in the « distressed » group when she is « healthy ». In the « distressed » group, there are also three misclassified firms so they are considered by the model « healthy » when they are actually distressed.
This model has a remarkable accuracy by classifying 93.34% of the control sample correctly. The error Type I is around 10% while the error type II is 3.33%.
Studying companies' exercises of control sample in case of two years before distress, we get the results announced at the table 12.

Table 12: Estimates of control sample two years before distress:
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		Percentage correct	Y		Percentage correct
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	29	1	96,7
		1	0	76	100,0	2	28	93,3
	global Percentage				100,0			95

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model classifies 29 firms correctly so we conclude an error type II equal to 3.33%. While in the group of distressed companies, there is two firm misclassified, giving us an error Type I of about 6.67%.
The increase of the efficiency of the MDA function, in this validation test (it passed from 93.3% to 95%), is due to the fact that the two samples of distressed firms (the initial sample and the control one) are randomly selected from a pool of 106failed firms. Moreover, as the samples are both small, the distributions of firms by size and industry differ considerably and this affects the efficiency of the function.
If we further increase the time between the prediction date and the advent of distress, using the same control sample but for three years before distress, we obtain the results reported in the following table.

Table 13: Estimates of control sample three years before distress:
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		Percentage correct	Y		Percentage correct
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	27	3	90,0
		1	0	76	100,0	2	28	93,3
	global Percentage				100,0			91,7

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

There are five misclassified companies. Two are considered as "healthy" when they are actually distressed and three are considered as distressed when they are really "healthy".
If we summarize, we get the following table:

Table 14: Results of estimation in the time and space

	Initial sample			Control sample		
	1year	2 years	3 years	1year	2 years	3 years
% of correct classification	100 %	96. 71 %	95.4 %	93,34%	95%	91,67%
% of classification error	0 %	3. 29 %	4.6 %	6,66%	5%	8,33%
Error type I	0 %	6. 58 %	7.89 %	10%	6,67%	6,67%
Error type II	0 %	0 %	1.32 %	3,33%	3,33%	10%

In effect, from the summary table above, using the initial sample for a maturity of one year prior to the failure, our model presents a rate of correct classification of 100 %. Such a result is consistent with that found by Frydman, Altman & Kao (1985) and Izán (1984) but remains well above those achieved by Yu et al (2014), Serrano-canca and al (2013), Myoung-Jong Kim, Dae-Ki Kang (2012) and Rafiei and al (2011). The same for the coming years two to three years before the failure, the method presents rates of correct classification, respectively, of the order of 96.71 per cent and 95.4 per cent largely superior to those made by Blum (1974), Altman (1968), Moyer (1977), Altman et al (1977), Frydman et al (1985), Dimitras and al (1987), Altman et al (1994), Back and al (1996), Charitou and al (2004) and Wu et al (2007) (the appendix 3 table 15).

By applying our model on a sample test, its percentage of correct classification remains beyond 90 %, outperformance as well the results obtained by Deakin (1972), Taffler (1982), Rose and Giroux (1984), Flagg and al (1991) and Brabazon and Keenan (2004) (see the appendix 3 table 15 and 16).

4. Estimation of the Logistic model parameters:

From the three subsamples which we called "Initial one year prior to distress," "Initial two years before distress" and "Initial three years before distress," each consist of the same 152 firms (76 "distressed" and 76 "healthy") but for different years (each sample is interested in the same year for all companies), and a set of 87 ratios (Appendix 1), we will try to formulate a logistic model, estimate its coefficients, calculate the probability of default in posteriori and develop a decision rule.

To perform the estimation, we used the "SPSS" software.

In a first step, it was assumed a model with 87 explanatory variables. The estimated model has provided us with results rather critical because the error rate is 50%:

Table 17 : Classification Table[a,b]

Observed			Predicted Y: 0	Predicted Y: 1	Percentage Correct
Step 0	Y	0	0	76	,0
		1	0	76	100,0
	Overall Percentage				50,0

Constant is included in the model.
The cut value is ,500

Such an error rate is explained by the importance of correlations between the explanatory variables: collinearity problem, correlation matrix and variance-covariance. Thing that leads us to take great care in selecting all ratios. Indeed, the number of ratios should not be too high for the study to be performed (Rose and Giroux (1984) identified more than 130 different ratios). In addition, the phenomenon of redundancy between ratios must be avoided: from the analysis of the correlation matrix, we observed a strong correlation between some explanatory variables; there is a great redundancy (the same information is provided by several ratios).

To solve this problem of collinearity, we opted for the "Feedward" method. It consists in introducing into the model, each time, the most correlated explanatory variable with the dependent variable until the matrix becomes not reversal. During this operation, we must be careful and retain only the independent variables that are significant at the 5% and can improve the $\overline{R^2}$ and we will ensure that all aspects of the situation of the company are covered.

Once this is done, based on 87 ratios initially taken, we are left with only 12 ratios, which will constitute the explanatory variables of the model to be estimated.

The estimate by the Logit model gives the following results:

Table 18 : Variables in the Equation

		B	S.E.	Wald	Df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	R_5	14,088	15960,342	,000***	1	,999	1312882,320	,000	.
	R_6	-131,311	43256,749	,000***	1	,998	,000	,000	.
	R_7	-272,144	40875,140	,000***	1	,995	,000	,000	.
	R_{15}	10,482	20133,088	,000***	1	1,000	35663,913	,000	.
	R_{19}	-23,350	13228,722	,000***	1	,999	,000	,000	.
	R_{26}	66,129	15652,150	,000***	1	,997	5,243E28	,000	.
	R_{28}	178,682	40767,715	,000***	1	,997	3,988E77	,000	.
	R_{33}	-13,401	6809,594	,000***	1	,998	,000	,000	.
	R_{40}	87,654	29863,406	,000***	1	,998	1,169E38	,000	.
	R_{61}	-,502	319,246	,000***	1	,999	,606	,000	3,348E271
	R_{74}	-15,515	25788,736	,000***	1	1,000	,000	,000	.
	R_{79}	52,925	14977,442	,000***	1	,997	9,663E22	,000	.
	Constant	126,426	38236,323	,000***	1	,997	8,052E54		

a. Variable(s) entered on step 1: R5, R6, R7, R15, R19, R26, R28, R33, R40, R61, R74, R79.

A careful analysis of the Wald test shows that all the variables used by the model are significant at a rate of 5 %.

The last twelve ratios represent the explanatory variables in our final model:

$$Z = 14{,}057\ R_5 - 131{,}311\ R_6 - 272{,}144\ R_7 + 10{,}482\ R_{15} - 23{,}350\ R_{19} + 66{,}129\ R_{26} + 178{,}682\ R_{28} - 13{,}401\ R_{33} + 87{,}654\ R_{40} - 0{,}501\ R_{61} - 15{,}515\ R_{74} + 52{,}925\ R_{79} + 126{,}426$$

Table 19 : The ratios retained by the method of logistic regression

Ratios	Formules
R_5	Cash and cash equivalents / current liabilities
R_6	Permanent Capital / Total Balance Sheet
R_7	Current assets / Total assets
R_{15}	Equity / Total Assets
R_{19}	Short-Term Debt / Total Liabilities
R_{26}	Amortization of Capital Assets / Gross Fixed Assets
R_{28}	Working Capital / Total Assets
R_{33}	current assets (excluding stocks) / current liabilities
R_{40}	current assets (excluding stock) / Total assets
R_{61}	Medium and long-term debt / Cash flow
R_{74}	Net Income / Total liabilities
R_{79}	Total Liabilities / Total Assets

In the equation used by logistic regression forecasting, we notice the presence of several ratios that have been selected as explanatory variables in previous studies.

Table 20 : the presence of several explanatory ratios in previous studies

Ratio	Authors
R_6	Conan & Holder (1979) ; Holder & al (1984)
R_7	Deakin (1972) ; Taffler (1982) ; Holder & al (1984)
R_{15}	Le crédit commercial de France (1995)]
R_{19}	Beaver (1966) ; Plat & Plat (1991)
R_{26}	Altman & al (1974) ; le modèle du C.E.S.A. (1974)
R_{33}	Deakin (1972) ; Edmister (1972) ; Houghton (1984) ; Burgstahler & al (1989) ; Michalopoulas & al (1993)
R_{40}	Conan & Holder (1979)]
R_{61}	Conan & Holder (1979) ; Bardos (1984)
R_{79}	Deakin (1972) ; Rose & Giroux (1984) ; Burgstahler & al (1989) ; Michalopoulas & al (1993) ; Altman & al (1994)

The overall significance test used in the logistic regression is the chi-square with k degrees of freedom (k is the number of explanatory variables in our case k = 12). If the critical probability is less than the significance level that one is fixed, we can consider that the model is globally significant. In our model, the statistical likelihood ratio (chi-square) is equal to 210.717; the critical probability associated is zero. The model is generally very significant, there is indeed a relationship between the explanatory variables and the variable to be explained.

Table 21 : Omnibus Tests of Model Coefficients

		Chi-square	Df	Sig.
Step 1	Step	210,717	12	,000
	Block	210,717	12	,000
	Model	210,717	12	,000

Similarly, decrease in value - 2 log likelihood from one stage to another also indicates the same result, that the introduction of new variables improves the model. In our case, this value down from 210.717 to zero.

Table 22 : Itération History[a,b,c]

Iteration		-2 Log likelihood	Coefficients
			Constant
Step 0	1	210,717	,000

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 210,717

c. Estimation terminated at iteration number 1 because parameter estimates changed by less than ,001.

Table 23 : Model Summary

Step	-2 Log likelihood	Cox &Snell R Square	Nagelkerke R Square
	,000[a]	,750	1,000

a. Estimation terminated at iteration number 20 because maximum iterations has been reached.

Cox & Snell R Square and Nagelkerke R Square tests help to determinate the percentage of the binary dependent variable that is explained by the explanatory variables retained confirmed the significativity of our model. Indeed, the Nagelkerke R Square test is an adjusted version of the Cox & Snell R Square one and therefore closer to reality. So, for our model, we notice that 100% of the variation in the dichotomous variable could be explained by the explanatory variables used and retained.

Once the overall significance of the model used is demonstrated, it remains to be seen whether the explanatory variables are significant. The Wald test in the logistic regression (see table above) demonstrates that, the twelve explanatory variables, retained in our model, are significant at 5 %.
The Hosmer and Lemeshow test divided into deciles based on predicted probabilities, and then computes a chi-square from observed and expected frequencies. The value p = 100% here is calculated from the chi-square distribution with 6 degrees of freedom, it indicates that the logistic model used is excellent.

Table 24 : Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	,000	6	1,000

After checking the overall significance of the model and the significance of the explanatory variables, our job is now to verify the performance and stability of the Logit model retained both in time, by applying it to the initial samples a year, two and three years prior to distress and in space using control samples one year, two years and three years before distress.

4.1 Estimation of the model discriminatory power one year before distress:

The estimation of the Logit model on the original sample, one year prior distress, shows that in the "healthy" firms group, the model classifies all "healthy" firms in their original group correctly.
In the distressed companies group, that interests us the most, we find no firm misclassified, so the model classifies successfully both companies "healthy" as "distressed" (table 25).

Table 25 : Estimates of initial sample one years before distress:

Classification table[b]

			Predicted		
			Selected observations[a]		
			Y		
Observations			0	1	Percentage correct
Stape 1	Y	0	76	0	100,0
		1	0	76	100,0
	global Percentage				100,0

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1

As far as the error Type I cost is much higher than that of an error type II [about 1 to 20 in Altman and al (1977)], then it seems more appropriate to judge the quality of the model on the base of the correct percentages of classification, in general, and of the error type I rate that it induces, in a particular way. These results "appear" as a whole interesting because they have the advantage of providing a combination of ratios based on which one can make a diagnostic of the company.
We say "appear interesting" because we should not judge the model before testing the performance over time (testing the model on the same companies but for different periods of time, two years and three before distress) and in space (testing the model on a control sample consisting of companies other than those in the sample of origin).

4.2 Validation of the model discriminatory power over time:

4.2.1 For the same companies two years before distress

The validation of model on exercises that come two years before distress gives the results in in the following table.

Table 26 : Estimates of initial sample two years before distress :
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
			0	1	Percentage correct	0	1	Percentage correct
Stape 1	Y	0	76	0	100,0	76	0	100,0
		1	0	76	100,0	1	75	98,7
	global Percentage				100,0			99,3

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model correctly classifies all « healthy » firms in their original group. In the « distressed » firms group, there are five firms misclassified, so the firms are considered as "healthy" when they are actually distressed. The model retains thus its discriminatory power, since the percentage of correct classification varies by only 0.66% from 100% to 99.34%, the error type I increases from 0 to 1.32%, while the error type II remains zero.

4.2.2 For the same companies three years before distress:

We will proceed in the same way as before, the same firms but for three years before distress, we get the results presented in the table 27.

Table 27: Estimates of initial sample three years before distress:
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
			0	1	Percentage correct	0	1	Percentage correct
Stape 1	Y	0	76	0	100,0	76	0	100,0
		1	0	76	100,0	5	71	93,4
	global Percentage				100,0			96,7

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the group of « failed » firms, we find that the model classifies four firms in the group of « healthy » one, while they are « distressed » which produces an error type I of about 5.26%. In the group of « healthy » companies, all companies are correctly classified and we have a percentage of error Type II equal to zero.
The forecasting ability of selected ratios, showed a satisfactory stability over time, since the overall error rate only increased from 0% to 3.29% % over the last three years preceding the distress, particularly some stability is noted for the classification of « healthy » companies .The following table will present a summary of changes in correct percentages of classifications and in errors of type I and II in time.

Table 28: Results of estimation in the time

	1 year before distress	2 years before distress	3 years before distress
% of correct classification	100 %	99. 34 %	96.71 %
% of classement error	0 %	0. 66 %	3.29 %
% of error type I	0 %	1. 32 %	6.58 %
% of error type II	0 %	0 %	0 %

Indeed, we notice that for the model used, the percentage of the error Type I varied only by 6.58% between the first and third years before distress. Furthermore, we find that the correct percentage of classification decreased only by 3.29% (it goes from 100% to 96.71%).
For our model, the most interesting element, in addition to its high correct percentage of classification, it is the weakness of the error Type I whose cost is higher. Concerning the error type II, we see that it remains zero.

4.3 Validation of the model discriminatory power in space:

To test the discriminatory power of the model in space, we use a control sample consisting of two new groups. The first contains the distressed firms while the second contains "healthy" companies, each list 30 firms. The model will be tested on companies other than those that were originated. The application of our Logit model on these samples gives us the estimates presented in the table 29

Table 29: Estimates of initial and control samples one year before distress :
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		Percentage correct	Y		Percentage correct
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	30	0	100,0
		1	0	76	100,0	3	27	90,0
	global Percentage				100,0			95,0

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model classifies two firms in the « distressed » group when they are « healthy ». In the « distressed » group, there are also misclassified firms so they are considered by the model « healthy » when they are actually distressed.
This model has a remarkable accuracy by classifying 95% of the control sample correctly. The error Type I is around 10% while the error type II is zero.
Studying companies' exercises of control sample in case of two years before distress, we get the results announced at the table 30.

Table 30: Estimates of control sample two years before distress:
Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		Percentage correct	Y		Percentage correct
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	30	0	100,0
		1	0	76	100,0	1	29	96,7
	global Percentage				100,0			98,3

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, we find that the model classifies all firms correctly so we conclude an error type II equal to zero. While in the group of distressed companies, there is a single firm misclassified, giving us an error Type I of about 3.33%.
The increase of the efficiency of the Logit function, in this validation test (it passed from 5% to 98.33%), is because that the two samples of distressed firms (the initial sample and the

control one) are randomly selected from a pool of 106failed firms. Moreover, as the samples are both small, the distributions of firms by size and industry differ considerably and this affects the efficiency of the function.
If we further increase the time between the prediction date and the advent of distress, using the same control sample but for three years before distress, we obtain the results reported in the following table.

Table 31: Estimates of control sample three years before distress:
Classification table[c]

			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
Observations			0	1	Percentage correct	0	1	Percentage correct
Stape 1	Y	0	76	0	100,0	30	0	100,0
		1	0	76	100,0	2	28	93,3
	global Percentage				100,0			96,7

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

In the « healthy » companies group, all firms are correctly classified. However, in the « distressed » firms group, there are two misclassified companies so they are considered as "healthy" when they are actually distressed.
If we summarize, we get the following table:

Table 32: Results of estimation in the time and space

	Initial sample			**Control sample**		
	1year	**2 years**	**3 years**	**1 year**	**2 years**	**3 years**
% of correct classification	100%	99,34 %	96,71 %	95 %	98,33 %	96,67 %
% of classement error	0%	0,66 %	3,29 %	5 %	1,67 %	3,33 %
Error type I	0%	1,32 %	6,58 %	10 %	3,33%	6,67 %
Error type II	0%	0 %	0 %	0 %	0 %	0 %

We notice that the percentage of correct classification, in the initial sample, varies from 100% to 96.71% (a change of 3.29%). It is a result that remains well above those achieved by Ohlson (1980) and Olson et al (2012). Note that Ohlson was the pioneer in the use of logistic regression in the prediction of business distresss. For the control sample, that percentage increased from 98,33% to 95%, a negative variation of 3.33%. Overall, the results provided by our model outperforms those presented by Wilcox (1973), Zavgren (1985), Flagg and al (1991), Barniv and Mcdonald (1992), Back and al (1996), Charalambous and al (2000), Charitou and al (2004), Wu and al (2007), Ahn and al (2011), Tserng and al (2011), Serrano-cenca and al (2013) and Wang and al (2014). (see appendix 4 table 33)

5. Estimation of the MLP model discriminatory power:

Among the set of ratios or economic variables that were examined in the literature, we found that not all of them explain in the same way the dependent variable. Indeed, some provide little information, others are tainted with noise, and others are redundant and therefore likely to give more weight to a particular dimension of the studied phenomenon. Selection therefore means finding a relevant subset, consisting of as many independent elements as possible and in sufficient number to adequately explain the process to be modelled. This amounts to testing the ability of the selected model and its variables to discriminate between healthy and failing firms.

Neural methods retain only nine of the 87 ratios of the initial battery (see appendix 5 table 34). Indeed, taking into account the significance and redundancy of the model's independent variables for a percentage of error (equal to 1%), we have retained only 9 ratios.
The purpose of this section is twofold. First, we estimate the discriminatory power of the MLP method and second to check this discriminatory power in time and in space.
Before presenting the results of the estimation, we should pay particular attention to two dimensions. The first dimension is the array of information on the network and its architecture that allow us to check first that the specifications are correct and then to extract the specificities of the network summarized in the following points:

- the number of units in the input stratum corresponds to the number of independent variables (see fig 3).
- likewise, a unit of a specific result is created for each class of healthy and failing companies for a total of 2 units in the output stratum.
- the automatic selection of the architecture chose a single hidden layer consisting of 5 units in addition to a biased one. Indeed, the architecture of the multilayer perceptron retained confirms

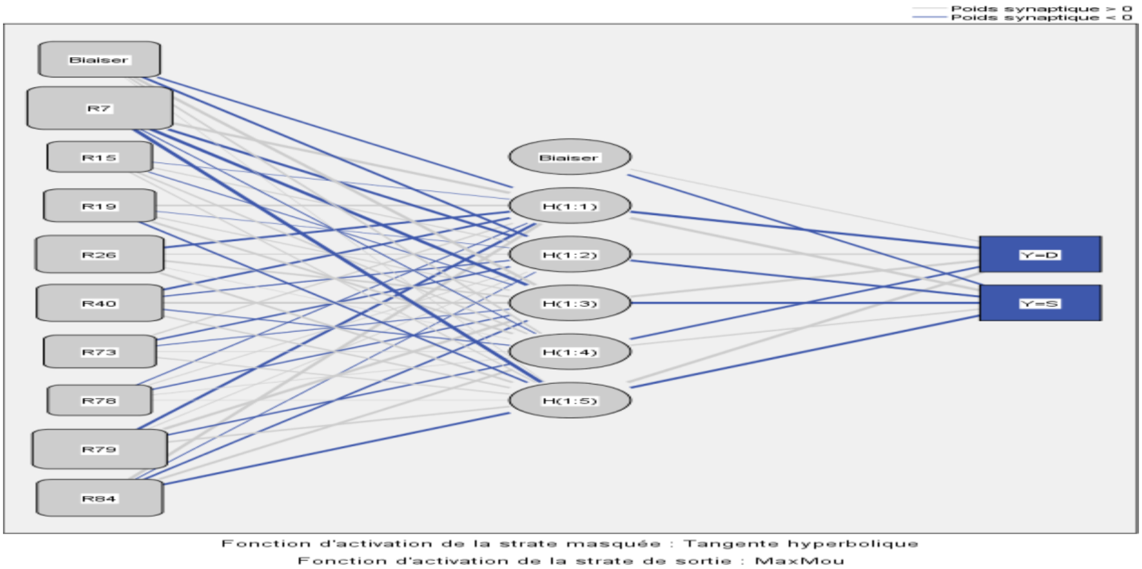

Fig 3 : Multilayer perceptron architecture

- the activation function used for the hidden layer is of the hyperbolic Tangent type whereas it is of the MaxMou type for the output layer (see appendix 5 table 35).

The second dimension is the model summary (see appendix 5 table 36) which displays information on the results of the learning of the final network and its application to the processed sample. Indeed,

- A cross-entropy error occurs because the output layer uses the MaxMou activation function. This is the error function that the network tries to minimize during learning.
- The percentage of incorrect forecasts comes from the league table and will be discussed later in this section.
- The stopping criterion is the indicator that should be imposed on the algorithm, i.e. the criterion which, once it satisfies the algorithm, stops and puts an end to all calculations. The stopping criterion can be either a number of variables or iterations, or the absence of a significant variation of an expected result after adding or removing a variable or still obtaining a satisfactory predictive capacity threshold. In our case, learning stopped when the error converged.

5.1. Estimation of the MLP model discriminatory power one year before distress:

The MLP method, applied to an original sample "a year before the distress", allows for correctly classifying 100% (152/152) of companies.

Table 37 : Estimates of initial sample one years before distress:

Classification table[b]

Observations			Predicted		
			Selected observations[a]		
			Y		Percentage correct
			0	1	
Stape 1	Y	0	76	0	100,0
		1	0	76	100,0
	global Percentage				100,0

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1

5.2. Validation of the discriminatory power of the method in time

5.2.1. for the same business two years before failure

In this section, we will keep the same companies, but we will use the data two years before distress.
The results show a slight reduction in accuracy of the method. Indeed, the correct classification percentage moved from 100% to 99.34% due to misclassification of one-distressed firms by type I error of about 1.32% (1/76). Type II error remained always zero.

Table 38: Estimates of initial sample two years before distress :

Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		correct Percentage	Y		Correct percentage
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	76	0	100,0
		1	0	76	100,0	1	75	98,68
	global Percentage				100,0			99,34

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

5.2.2. for the same business three years before failure

Three years from the date of the coming failure, the MLP method yielded the following results:

Table 39: Estimates of initial sample three years before distress:

Classification table[c]

Observations			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y		correct percentage	Y		
			0	1		0	1	
Stape 1	Y	0	76	0	100,0	73	3	96,05
		1	0	76	100,0	1	75	98,68
	global Percentage				100,0			97,37

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

Table 40: The results provided by the model over time

	1 year before distress	2 years before distress	3 years before distress
% of correct classification	100 %	99.34 %	97.37 %
% of classification error	0 %	0.66 %	2.63 %
% of error type I	0 %	1.32 %	1.32 %
% of error type II	0 %	0 %	3.95 %

When evaluating the predictive ability of the model, we found a correct classification percentage varying between 100% (152/152) and 97.37% (148/152), respectively for year one and three years before distress. Similarly, type I and II errors have increased from 0% (0/76) to 3.95% (3/76) during the same period. Despite its application to data located three years before the advent of the distress, the MLP method keeps a decent percentage of correct classification (97.37%), allowing it to keep almost all of its predictive capacity in time

5.3. Validation of the discriminatory power of the method in space

Since the MLP method was able to keep its predictive ability in time, we will now examine if it is able to keep its capacity in space. To find out, we will apply the method on data one, two and three years before distress for a new firm population, called the control sample. This test sample consists of 60 new firms 30 "healthy" and 30 "distressed”. The obtained results are as follows:

Table 41: Estimates of initial and control samples one year before distress:
Classification table[c]

			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
Observations			0	1	Correct percentage	0	1	Correct percentage
Stape 1	Y	0	76	0	100,0	30	0	100,0
		1	0	76	100,0	0	30	100,0
	global Percentage				100,0			100,0

Table 42: Estimates of control sample two years before distress:
Classification table[c]

			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
Observations			0	1	Percentage correct	0	1	Percentage correct
Stape 1	Y	0	76	0	100,0	30	30	100,0
		1	0	76	100,0	0	0	100,0
	global Percentage				100,0			100,0

Table 43: Estimates of control sample three years before distress:
Classification table[c]

			Predicted					
			Selected observations[a]			Excluded observations[b]		
			Y			Y		
Observations			0	1	Percentage correct	0	1	Percentage correct
Stape 1	Y	0	76	0	100,0	29	1	96,67
		1	0	76	100,0	0	30	100,0
	global Percentage				100,0			98,33

a. Selected observations Partition EQ 1
b. Excluded observations Partition NE 1
c. The cut value is ,500

Table 44: The results provided by the model over time and space

Sample	Initial			Control		
	1year	2 years	3 years	1year	2 years	3 years
% of correct classification	100 %	99.34 %	97.37 %	100 %	100 %	98,33 %
% of error classification	0 %	0.66 %	2.63 %	0 %	0 %	1,67 %
Error type I	0 %	1.32 %	1.32 %	0 %	0 %	0 %
Error type II	0 %	0 %	3.95 %	0 %	0 %	3,33 %

The results show a percentage of correct classification varying between 100 % (60/60) and 98,33% (59/60) for the coming three years before the failure. Type I error is null since the model to be correctly classified all failing companies. Type II error has reached 3.33 % (1/30) during the third year since the model has ranked one healthy business as failed.

The MLP method has retained its discriminatory ability both in time with a correct classification rate that remains above 97.37%, and in space with a good ranking ratio of about 98,33%.

Referring to the work done in the area, we find that our MLP model has better results than those of Min and Lee (2005), Hu and Tseng (2005), Boyacioglu and al (2009) and Serrano-canca and al (2013). However, our results remain below those reported by Wu et al (2007) (Table 45).

Table 45 : T**he results provided by the literature review**

Authors	Year	Method	Percentage of correct classification		
			One year	Two year	Three year
Boyacioglu and al	2009	MLP	95,5%		
Min and Lee	2005	MLP	85,25%		
Serrano-canca and al	2013	MLP	93,93%		
Wu and al	2007	MLP	100%	100%	100%
Hu and Tseng	2005	MLP	81,64%		

6. Estimation of the SVM model discriminatory power:

In this section, we will repeat the same work already done for MLPs using the nine explanatory variables and apply it for SVMs to test the ability of this method to maintain its predictive capacity in time and space.

6.1. Estimation of the SVM model discriminatory power one year before distress:

SVM method, applied to original sample "a year before the distress", allows to correctly classifying 99.34% (151/152) of companies (incorrect classification of one failed businesses) this corresponds to an error rate of type I of about 1.32% (1/76). The error rate of type II having remained zero.

Table 46: Initial one year before distress

Error rate	0,0066

Values prediction

Value	Recall	1-Precision
D	0,9868	0,0000
S	1,0000	0,0130

Confusion matrix

	D	S	Sum
D	75	1	76
S	0	76	76
Sum	75	77	152

6.2. Validation of the discriminatory power of the method in time

6.2.1. for the same business two years before failure

In this section, we will keep the same companies, but we will use the data relating to two years before the distress.
The results show a slight reduction in accuracy of the method. Indeed, the correct classification percentage passed from 100% to 95.39% due to misclassification of seven distressed firms either type I error of about 9.21% (7/76). The error of the second kind having remained always zero.

Table 47 : Initial two years before distress

Error rate			0,0461			
Values prediction			**Confusion matrix**			
Value	**Recall**	**1-Precision**		**D**	**S**	**Sum**
D	0,9079	0,0000	**D**	69	7	76
S	1,0000	0,0843	**S**	0	76	76
			Sum	69	83	152

6.2.2. for the same business three years before failure

Moving three years away from the date of the failure, the SVM method presents the following results:

Table 48 : Initial three years before distress

Error rate			0,0658			
Values prediction			**Confusion matrix**			
Value	**Recall**	**1-Precision**		**D**	**S**	**Sum**
D	0,8816	0,0147	**D**	67	9	76
S	0,9868	0,1071	**S**	1	75	76
			Sum	68	84	152

When evaluating the predictive ability of the model, we found a correct classification percentage varying between 99.34% (151/152) and 93.42% (142/152), respectively for years one and three years before distress. Similarly, the type I error has increased from 1.32% (1/76) to 11,84 % (9/76) for the same period, the error rate of type II has increased from 0 % (0/76) to 1,32 % (1/76). Despite its application to data located three years before the advent of the distress, the method of SVM keeps a respectable percentage of correct classification (93.42%), allowing it to keep almost all of its predictive capacity in time.

Table 49: The results provided by the model over time

Sample	**Initial**		
	1year	**2 years**	**3 years**
% of correct classification	99,34 %	95. 39 %	93.42 %
% of classification error	0,66 %	4. 61 %	6.58 %
Error type I	1,32 %	9.21 %	11.84 %
Error type II	0 %	0 %	1,32 %

6.3. Validation of the discriminatory power of the method in space

Since the method of SVM was able to keep its predictive ability in time, we will now see if it is able to keep its capacity in space. To find out, we will apply the method on data one, two and three years before distress for a new firm population, called control sample. This test sample consists of 60 new firms 30 "healthy" and 30 "distressed".

Table 50: The results provided by the model over time and space

Sample	Initial			Control		
	1year	2 years	3 years	1year	2 years	3 years
% of correct classification	99,34 %	95. 39 %	93.42 %	96,67 %	95 %	95 %
% of error classification	0,66 %	4. 61 %	6.58 %	3,34 %	5 %	5 %
Error type I	1,32 %	9.21 %	11.84 %	6,67%	10 %	6,67 %
Error type II	0 %	0 %	1,32 %	0 %	0 %	3,33 %

The results show a percentage of correct classification varying between 96.67 % (58/60) and 95% (57/60) for the coming three years before the failure. The error of the first kind has reached 10% (3/30) at the second year. The Type II error did not exceed 3.33% (1/30). (Table 51, 52 and 53).

Table 52 : Control one year before distress

Error rate	0,0333
Values prediction	**Confusion matrix**

Value	Recall	1-Precision
D	0,9333	0,0000
S	1,0000	0,0625

	D	S	Sum
D	28	2	30
S	0	30	30
Sum	28	32	60

Table 52 : Control two years before distress

Error rate	0,0500
Values prediction	**Confusion matrix**

Value	Recall	1-Precision
D	0,9000	0,0000
S	1,0000	0,0909

	D	S	Sum
D	27	3	30
S	0	30	30
Sum	27	33	60

Table 53 : Control three years before distress

Error rate	0,0500
Values prediction	**Confusion matrix**

Value	Recall	1-Precision
D	0,9333	0,0345
S	0,9667	0,0645

	D	S	Sum
D	28	2	30
S	1	29	30
Sum	29	31	60

The SVM method has retained its ability to discrimination both in time with a correct classification rate, which remains above 93.42%, and in space at a ratio of good ranking of about 95%.

Referring to the work done in the area, we find that our SVM model has better results than those achieved by Min and Lee (2005), Olson and al (2012), Serrano-canca and al (2013) and Yu and al (2014). As against our results remain below those reported by Wu and al (2007), Lee and al (2010) and Ahn and al (2011) (Table 54).

Table 54: **The results provided by the literature review**

Authors	Year	Method	Percentage of correct classification
			One year
Ahn and al	2011	SVM	100%
Boyacioglu and al	2009	SVM	90,90%
Min and Lee	2005	SVM	88,01%
Min and al	2006	SVM	82,45%
Kim and Kang	2012	SVM	72,45%
Olson and al	2012	SVM	66,1%
Serrano-canca and al	2013	SVM	90,71%
Tserng and al	2011	SVM	80,31%
Wang and al	2014	SVM	79,99%
Wu and al	2007	SVM	100%
Yu and al	2014	SVM	93,2%
Lee and al	2010	SVM	100%

7. Comparison of methods

7.1. Results from the models applied to initial samples

Table 55: Comparison of the two models applied to initial samples

	MDA (16 ratios)	LR (12 ratios)	MLP (9 ratios)	SVM (9 ratios)
a) one year before distress				
- % of correct classification	100 %	100 %	100 %	99,34 %
- % of error classification	0 %	0 %	0 %	0,66 %
- Error du type I	0 %	0 %	0 %	1,32 %
- Error du type II	0 %	0 %	0 %	0 %
b) two years before distress				
- % of correct classification	96,71 %	99,34 %	99,34 %	95,39 %
- % of error classification	3,29 %	0,66 %	0,66 %	4,61 %
- Error du type I	6,58 %	1,32 %	1,32 %	9,21 %
- Error du type II	0 %	0 %	0 %	0 %
c) three years before distress				
- % of correct classification	**95,4 %**	**96,71 %**	**97,37 %**	**93,42 %**
- % of error classification	4,6 %	3,29 %	2,63 %	6,58 %
- Error du type I	7,89 %	6,58 %	1,32 %	11,84 %
- Error du type II	1,32 %	0 %	3,95 %	1,32 %

The results obtained using the initial samples (validation in time) show a superiority of the MLP compared to other methods. Indeed, the MLP has a correct classification percentage that remains above 97.37% compared respectively to 96.71% for the LR, 95.4% for the MDA and 93.42% for the SVM. Similarly, to the extent that the cost of a Type I error is much higher than that of a Type II error, we found that the maximum rate of error for MLP (1,32%) is largely lower than that committed by SVM (7.89%), LR (6,58%) and SVM (11,84).

7.2. Results from the models applied to control samples

Table 56: Comparison of the two models applied to control samples

	MDA (16 ratios)	LR (12 ratios)	MLP (9 ratios)	SVM (9 ratios)
a) one year before distress				
- % of correct classification	93,34 %	**95 %**	100 %	**96,67 %**
- % of error classification	6,66 %	5 %	0 %	3,33 %
- Error du type I	10 %	10 %	0 %	6,67 %
- Error du type II	3,33 %	0 %	0 %	0 %
b) two years before distress				
- % of correct classification	95 %	98,33 %	100 %	95 %
- % of error classification	5 %	1,67 %	0 %	5 %
- Error du type I	6,67 %	3,33 %	0 %	10 %

- Error du type II	3,33 %	0 %	0 %	0 %
c) three years before distress				
- % of correct classification	**91,67 %**	96,67 %	**98,33 %**	**95 %**
- % of error classification	8,33 %	3,33 %	1,67 %	5 %
- Error du type I	6,67 %	6,67 %	0 %	6,67 %
- Error du type II	10 %	0 %	3,33 %	3,33 %

The above comparative table shows a clear superiority of the multilayer perceptron method, both in time and in space, compared to the other methods. Indeed, even over three years of the control sample, the correct classification rate has always remained greater than or equal to 98,33%, well above 91.67% for the MDA, 95% for the SVM and 96,67 for the LR. The Type I error is zero and the Type II error is 1.33%, i.e. out of the sixty companies in the test sample, the model (MLP) made only one error by classifying a 'healthy' company as a 'defaulter'.
To check this superiority, we applied MLPs and SVMs to estimate the 16 variables retained by the multivariate discriminant analysis as well as the 12 variables retained by the logistic regression. The results obtained confirm the supremacy of the MLPs. Indeed, the results obtained by the MLPs on the 9, 12 or 16 ratios are much better than those of the other used methods (see appendix 6 table 57 and 58)
Based on the results obtained on the initial sample, we can classify the used forecast models as follows:

- The multilayer perceptron will rank first followed by the logistic regression while the multivariate discriminant analysis and the supports vectors machines will rank third and fourth respectively.
- The results on the test sample confirm this ranking with only one change in rank between the multivariate discriminant analysis and the supports vectors machines.

In the literature, the superiority of the multilayer perceptron model (MLP) is confirmed by Udo (1993), Kumar et al (1997), Wu (1999), Brabnazon and Keenan (2004), Yi-Chung Hu et al (2005), olson and al (2012), Sangjae Lee et al (2013) and Serrano-Cinca and al (2013). However, for Wu et al (1997), Min and Lee (2005), Min et al (2006) and Zhou et al. (2012), the SVM is best suited to discriminate between healthy companies and distressed companies.

Table 59: Comparison between MDA, LR, MLP and SVM

Authors	Year	Conclusion
Serrano-canca and al	2013	MLP > MDA MDA > SVM LR > SVM
Min and Lee	2005	SVM > MDA SVM > LR SVM > MLP
Olson and al	2012	SVM < MLP SVM < LR
Wu, Wu and al	1999 and 2007	MLP > MDA SVM > LR MLP = SVM
Min and al	2006	SVM > LR SVM > MLP
Zhou and al	2012	SVM > LR SVM > MDA

8. Conclusion:

Either for the initial sample or the control sample, the results provided by the selected methods perform well either from the point of view of the percentage of correct classification or from the point of view of the stability of the discriminatory power in time and space.

The supremacy of the multilayer perceptron model is confirmed both experimentally and methodologically.

Experimentally, the MLP has two advantages over conventional statistical methods (logistic regression (LR) and multivariate discriminant analysis (MDA)).

- a relative ease of manipulation, since the model only retains nine independent variables in its prediction equation compared to twelve for the logistic regression and sixteen for the multivariate discriminant analysis;
- the first-order error is always lower for the multilayer perceptron model than for the logistic regression and the multivariate discriminant analysis, which means a lower cost for creditors.

From a methodological point of view, the multilayer perceptron model has several advantages over the discriminant analysis and the logistic regression:

- the absence of hypotheses on the distribution of variables makes allows for overcoming the restrictions on the multivariate discriminant analysis like the multi-normality of the variables or the equality of the variance-covariance matrices. The multilayer perceptron model perfectly suits the treatment of non-linearly separable problems, which coincides with the needs of failure prediction where the separation between healthy and failing companies is non-linear;
- In contrast to the logistic regression, which requires that the a priori distribution of the classes of healthy and failing firms in the population be fair, i.e. that the sample should contain exactly the same proportion of healthy and failing firms, the multilayer perceptron model helps to overcome this constraint and complies with the reality that the percentage of firms in difficulty is much lower than that of healthy firms;
- The multilayer perceptron model has a specific processing capacity. Indeed, they are able to determine relationships and links between the used variables, which allows for moving from the simple predictive feature of statistical tools to an explanatory approach based on the interactions between the selected variables.
- The multilayer perceptron model is able to learn despite the presence of missing information.

In the four forecasting methods, namely the multivariate discriminant analysis, the logistic regression, the multilayer perceptron model, and the vector machine supports, we note the presence of the following independent variables:

- R_7 = Current assets / Total Assets
- R_{15} = Equity /Total Assets
- R_{19} = Short-term debt / Total Liabilities
- R_{26} = Amortization of Capital Assets / Gross Fixed Assets
- R_{40} = Current assets (excluding stocks) / Total Assets
- R_{79} = Total Liabilities / Total Assets

The predominance of structural ratios can be seen. Indeed, the debt ratio R19 expresses the degree of maturity of the company's debts. As a solvency ratio, the R79 ratio is assessed when the company goes insolvent. It is a kind of security measure for creditors since it allows them to determine the capacity of the real assets, in the event of a sale, to cover all the company's debts. Similarly, this ratio serves as an indication of the asset funding mix. The presence of the R15 financial autonomy ratio is expected since most Tunisian companies suffer from a structural problem linked to low equity capital. The two liquidity ratios R7 and R40 measure the degree of liquidity of current assets, or even the short-term solvency of the company, which is of a primary interest to creditors. The aging ratio of tangible fixed assets R26 indirectly informs us about the firm's ability to renew its equipment.

References :

[1] Ahn, J., and al. (2011). "Trade finance and the great trade collapse." The American Economic Review **101**(3): 298-302.

[2] Altman, E. I. (1968). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy." The journal of finance **23**(4): 589-609.

[3] Altman, E. I. (1984). "A further empirical investigation of the bankruptcy cost question." The journal of finance 39(4): 1067-1089.

[4] Altman, E. I., and al. (1977). " Zeta Analyses : A new model to identify bankruptcy risk of corporations." journal of banking & finance **1**(1): 29-54.

[5] Altman, E. I., and al. (1994). "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)." journal of banking & finance**18**(3): 505-529.

[6] Altman, E. I., and al. (2005). "The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications*." The Journal of Business **78**(6): 2203-2228.

[7] Altman, E., (1985). The Success of Business Failure Prediction Models (An International Survey), Journal of Banking and Finance, 8, pp. 171-184.

[8] Anderson, J. A. (1972). "Separate sample logistic discrimination." Biometrika **59**(1): 19-35.

[9] Anderson, J. and V. Blair (1982). "Penalized maximum likelihood estimation in logistic regression and discrimination." Biometrika **69**(1): 123-136.

[10] Aziz, A., and al. (1988). "Bankruptcy prediction-an investigation of cash flow based models [1]." Journal of Management Studies **25**(5): 419-437.

[11] Back, B., and al. (1996). "Neural networks and genetic algorithms for bankruptcy predictions." Expert Systems with Applications **11**(4): 407-413.

[12] Bardos M. (1984) « Le risque de défaillance d’entreprise » Cahiers économiques et monétaires de la banque de France, p. 35, n° 19, 190 p

[13] Bardos M. (1998a). « Detecting the risk of company failure at the Banque de France”, Journal of Banking and Finance, vol. 22, pp. 1405-1419.

[14] Bardos, M. (1989). "Trois méthodes d’analyse discriminante." Cahiers économiques et monétaires **33**: 151-190.

[15] Bardos, M. and W. Zhu (1997). "Comparaison de l'analyse discriminante linéaire et des réseaux de neurones: Application à la détection de défaillance d'entreprises." Revue de statistique appliquée **45**(4): 65-92.

[16] Barniv, R. and J. B. McDonald (1992). "Identifying financial distress in the insurance industry: A synthesis of methodological and empirical issues." Journal of Risk and Insurance **59**: 543-543.

[17] Beaver, W. H. (1966). "Financial ratios as predictors of failure." Journal of accounting research: 71-111.

[18] Blum, M. (1974), “Failing Company Discriminant Analysis, Journal of Accounting Research”, vol. 12, n° 1, pp. 1-25.

[19] Boser, B. E., and al. (1992). A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory, ACM.

[20] Boyacioglu, M. A., and al. (2009). "Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey." Expert Systems with Applications **36**(2): 3355-3366.

[21] Brabazon, A. and P. B. Keenan (2004). "A hybrid genetic model for the prediction of corporate failure." Computational Management Science **1**(3-4): 293-310.

[22] Burgstahler, D., and al. (1989). "Changes in the probability of bankruptcy and equity value." Journal of Accounting and Economics**11**(2): 207-224.

[23] Charalambous, C., and al. (2000). "Comparative analysis of artificial neural network models: Application in bankruptcy prediction." Annals of operations research **99**(1-4): 403-425.

[24] Charitou, A., and al. (2004). "Predicting corporate failure: empirical evidence for the UK." European Accounting Review**13**(3): 465-497.

[25] Chen, J., and al. (2006). "Financial distress prediction in China." Review of Pacific Basin Financial Markets and Policies **9**(02): 317-336.

[26] Conan, J. and M. Holder (1979). "Variables explicatives de performances et contrôle de gestion dans les PMI." These d'Etat, CERG, Université Paris Dauphine.

[27] Cox, D. (1970). "Analysis of binary data, 1970." Methuen, London: 103-108.

[28] Dambolena, I. G. and S. J. Khoury (1980). "Ratio stability and corporate failure." The journal of finance 35(4): 1017-1026.

[29] Day, N. E. and D. F. Kerridge (1967). "A general maximum likelihood discriminant." Biometrics **23**(2): 313-323.

[30] Deakin, E. B. (1972). "A discriminant analysis of predictors of business failure." Journal of accounting research: 167-179.

[31] Diamond, J. M., and al. (1976). "Island biogeography and conservation: strategy and limitations." CR Seances Soc. Biol. Paris 160(3): l966.

[32] Dimitras, A., and al. (1999). "Business failure prediction using rough sets." European Journal of Operational Research **114**(2): 263-280.

[33] Edmister, R. O. (1972). "An empirical test of financial ratio analysis for small business failure prediction." Journal of Financial and Quantitative Analysis**7**(02): 1477-1493.

[34] Fitzpatrick, P. J. (1932), "A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Firms", The Certified Public Accountant October 1932, 598-605

[35] Flagg, J. C., and al. (1991). "Predicting corporate bankruptcy using failing firms." Review of financial Economics**1**(1): 67-78.

[36] Frydman, H., and al. (1985). "Introducing recursive partitioning for financial classification: the case of financial distress." The journal of finance **40**(1): 269-291.

[37] Gunn, S. R. (1998). "Support vector machines for classification and regression." ISIS technical report **14**.

[38] Holder, M., and al. (1984). "Le score de l'entreprise." Nouvelles Editions Fiduciaires, Paris.

[39] Houghton, K. A. and R. Sengupta (1984). "The effect of prior probability disclosure and information set construction on bankers' ability to predict failure." Journal of accounting research: 768-775.

[40] Izan, H. (1984). "Corporate distress in Australia." journal of banking & finance **8**(2): 303-320.

[41] Jones S. and Hensher, D. A. (2007a). "Forecasting corporate bankruptcy: optimizing the performance of the mixed logit model." Abacus **43**(3): 241-264.

[42] Jones, S. and D. A. Hensher (2004). "Predicting firm financial distress: a mixed logit model." The Accounting Review **79**(4): 1011-1038.

[43] Jones, S. and D. A. Hensher (2007). "Modelling corporate failure: A multinomial nested logit analysis for unordered outcomes." The British Accounting Review **39**(1): 89-107.

[44] Kim, M.-J. and D.-K. Kang (2012). "Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction." Expert Systems with Applications **39**(10): 9308-9314.

[45] Kira D.S, Doreen D., Ngnyen D. (1997) « An application of artificial neutral networks and statistical methods in qualitative evaluation of small business loans » ASMD Summer, April.

[46] Kumar, N., and al. (1997). "Financial decision support with hybrid genetic and neural based modeling tools." European Journal of Operational Research 103(2): 339-349.

[47] Lachenbruch, P. A., and al. (1973). "Robustness of the linear and quadratic discriminant function to certain types of non-normality." Communications in Statistics **1**(1): 39-56.

[48] Laitinen, E. K. and T. Laitinen (1998). "Cash management behavior and failure prediction." Journal of Business Finance & Accounting **25**(7-8): 893-919.

[49] Laitinen, E. K. and T. Laitinen (2001). "Bankruptcy prediction: application of the Taylor's expansion in logistic regression." International Review of Financial Analysis **9**(4): 327-349.

[50] Lau, A. H.-L. (1987). "A five-state financial distress prediction model." Journal of accounting research: 127-138.

[51] Lee, M.-C. and C. To (2010). "Comparison of multilayer perceptronand back propagation neural network in evaluating the enterprise financial distress." arXiv preprint arXiv:1007.5133.

[52] Lee, S. and W. S. Choi (2013). "A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis." Expert Systems with Applications **40**(8): 2941-2946.

[53] Li, H. and J. Sun (2011). "Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction." Expert Systems with Applications **38**(5): 6244-6253.

[54] Marks, S. and O. J. Dunn (1974). "Discriminant functions when covariance matrices are unequal." Journal of the American Statistical Association **69**(346): 555-559.

[55] Martin, D. (1977). "Early warning of bank failure: A logit regression approach." journal of banking & finance**1**(3): 249-276.

[56] Merwin, C. L. (1942), "Financing Small Corporations in Five Manufacturing Industries", 1926-1936, National Bureau of Economic Research, Financial Research Program III, Studies in Business Financing, 172 p. December, pp. 598-605, 656-662, 727-731.

[57] Min, J. H. and Y.-C. Lee (2005). "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters." Expert Systems with Applications **28**(4): 603-614.

[58] Min, S.-H., and al. (2006). "Hybrid genetic algorithms and support vector machines for bankruptcy prediction." Expert Systems with Applications **31**(3): 652-660.

[59] Mossman, C. E., and al. (1998). "An empirical comparison of bankruptcy models." Financial Review **33**(2): 35-54.

[60] Moyer, R. C. (1977). Forecasting Financial Failure: A Re-examination. Financial Management, 6(1), 11-17.

[61] Nam, J. H. and T. Jinn (2000). "Bankruptcy prediction: Evidence from Korean listed companies during the IMF crisis." Journal of International Financial Management & Accounting **11**(3): 178-197.

[62] Ohlson, J. A. (1980). "Financial ratios and the probabilistic prediction of bankruptcy." Journal of accounting research: 109-131.

[63] Olson, D. L., and al. (2012). "Comparative analysis of data mining methods for bankruptcy prediction." Decision Support Systems**52**(2): 464-473.

[64] Platt, H. D. and M. B. Platt (1991). "A note on the use of industry-relative ratios in bankruptcy prediction." journal of banking & finance**15**(6): 1183-1194.

[65] Ramser, J., Foster, L. (1931), “A Demonstration of Ratio Analysis”, Bulletin n° 40, University of Illinois, Bureau of Business Research, 52 p.

[66] Rose P. S.andGiroux G.A. (1984). “Predicting Corporate Bankruptcy: An Analytical and Empirical Evaluation”, Review of Business and Economic Research, vol. XIX, n° 2, pp. 1-12.

[67] Salchenberger, L. M., and al. (1992). "Neural Networks: A New Tool for Predicting Thrift Failures*." Decision Sciences **23**(4): 899-916.

[68] Serrano-Cinca, C. and B. Gutiérrez-Nieto (2013). "Partial least square discriminant analysis for bankruptcy prediction." Decision Support Systems**54**(3): 1245-1255.

[69] Taffler, R. J. (1982). "Forecasting company failure in the UK using discriminant analysis and financial ratio data." Journal of the Royal Statistical Society. Series A (General): 342-358.

[70] Taffler, R. J. (1983), “The Assessment of Company Solvency and Performance Using a Statistical Model, Accounting and Business Research”, vol. 13, n° 52, pp. 295-307.

[71] Taffler, R. J. (1983). "Forecasting company failure in the UK using discriminant analysis and financial ratio data." Journall of the Royal Statistical Society. Series A (General): 342-358.

[72] Tam, K. Y. and M. Y. Kiang (1992). "Managerial applications of neural networks: the case of bank failure predictions." Management Science **38**(7): 926-947.

[73] Tserng, H. P., and al. (2011). "An enforced multilayer perceptronmodel for construction contractor default prediction." Automation in Construction**20**(8): 1242-1249.

[74] Udo, G. (1993). "Neural network performance on the bankruptcy classification problem." Computers & industrial engineering **25**(1): 377-380

[75] Vapnik, V. N. and A. Y. Chervonenkis (1968). On the uniform convergence of relative frequencies of events to their probabilities. Soviet Math. Dokl.

[76] Vapnik, V. N. and S. Kotz (1982). Estimation of dependences based on empirical data, Springer-Verlag New York.

[77] Vapnik, Vladimir N.(1995) ; The Nature of Statistical Learning Theory, Springer-Verlag. ISBN 0-387-98780-0

[78] Varetto F. (1998). “Genetic Algorithms Applications in the Analysis of Insolvency Risk”, Journal of Banking and Finance, vol. 22, n° 10-11, October, pp. 1421-39.

[79] Wang, Jian Ma, Shanlin Yang (2014)“An improved boosting based on feature selection for corporate bankruptcy prediction”. Expert Systems with Applications, Volume 41, Issue 5, April, Pages 2353-2361

[80] Wilcox, J. W. (1973). "A prediction of business failure using accounting data." Journal of accounting research: 163-179.

[81] Winakor, A. H., Smith, R. F. (1935), “Changes in the Financial Structure of Unsuccessful Industrial Corporations”, University of Illinois, Bureau of Business Research, Bulletin n° 51, 44 p.

[82] Wu, C.-H., and al. (2007). "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy." Expert Systems with Applications**32**(2): 397-408.

[83] Yu, Q., and al. (2014). "Bankruptcy prediction using Extreme Learning Machine and financial expertise." Neurocomputing **128**: 296-302.

[84] 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.

[85] Zeitun, R., & al. (2007). "Default probability for the Jordanian companies: A test of cash flow theory." International Research Journal of Finance and Economics, 8, 147-162

[86] Zhou, L., and al. (2012). "Empirical models based on features ranking techniques for corporate financial distress prediction." Computers & mathematics with applications **64**(8): 2484-2496.

[87] Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research, 22, 59-82.

[88] Zopounidis, C. (1995). Evaluation du risque de défaillance de l'entreprise: Méthodes et cas d'application, Economica

The appendix

Appendix 1: The battery of 87 ratios initially used:

R1= Financial expenses / Operating income
R2= Cash-flow / Turnoverexcluding taxes
R3= Cash-flow / Total debt
R4= Cash-flow / Equity
R5 = Cash and cash equivalents/ Current liabilities
R6= Permanent capital/ Total Balance Sheet
R7= Current assets / Total Assets
R8= Financial expenses / Turnover
R9= Personnel costs / Added value
R10= Operating income / Added value
R11= Total debt / Equity
R12= Working Capital /Turnover
R13= Added value / Fixed assets
R14= Financial expenses/ Added value
R15= Equity /Total Assets
R16= Working Capital / Cash-flow
R17= Cash and cash equivalents/ Short-term debt
R18= Stocks / Total Assets
R19= Short-term debt / Total Liabilities

R20= Turnovers / Equity
R21= Total Debts/ Total Liabilities
R22= Equity / Permanent equity
R23= Permanent equity / Net fixed assets
R24= Equity / Net fixed assets
R25= Current assets / Current liabilities
R26= Amortization of Capital Assets / Gross Fixed Assets
R27= Added value / Actifs non courants
R28= Working Capital / Total Assets
R29= Added value / Total Assets
R30= Turnover / Total Assets
R31= Cash-Flow / Short-term debt
R32= Short-term debt / Equity
R33= Current assets (excluding stocks)/ Current liabilities
R34= Added value / Turnovers
R35 = Staff costs / Trade accounts payable
R36 = Current assets t – Current assets t-1 / Current assets t-1
R37 = Non-current assetst – Non-current assetst-1 / Non-current assetst-1
R38 = Current assets (excluding stocks) / Turnover
R39 = Current assets (excluding stocks) / Current bank accounts
R40 = Current assets (excluding stocks) / Total Assets
R41 = Current assets (excluding stocks) / Current assets
R42 = Current assets / Turnover
R43 = EBIT(Earnings Before Interest and Taxes) (/ Total Assets
R44 = EBIT / Turnover
R45 = EBIT / Financial expenses
R46 = Net operating result / Equity
R47 = Net operating result / Turnover
R48 = Net operating result / Total Assets
R49 = Working capital requirements / Working capital
R50 = Cash Flow / Total Liabilities
R51 = Cash-Flow / Turnoverexcluding taxes
R52 = Cash-Flow / Non-current liabilities
R53 = Cash Flow / Total Assets
R54 = Staff costs / Gross operating incomes
R55 = Turnover t – Turnover t-1 / Turnover t-1
R56 = Turnover t-1 / Total Assets t-1
R57 = Purchase cost of materials consumed (or purchase cost of production sold) / Average stock material or production
R58 = Receivables/ Total Assets
R59 = Receivables + Stocks / Suppliers
R60 = Non-current liabilities/ Equity
R61 = Medium and long-term debt / Cash flow
R62 = Customer credits Duration
R63 = Credits suppliersDuration

R64 = Gross operating incomes/ Turnover
R65 = Gross operating incomes/ Total Assets
R66 = Gross operating incomes/ Added value
R67 = Working Capital/ Added value
R68 = Non-current liabilities / Non-current assets
R69 = Reserves / Total Assets
R70 = Pre-tax income/ Current liabilities
R71 = Gross operating incomes / Total Assets
R72 = Net Income / Equity
R73 = Net Income / Turnover
R74 = Net Income / Total Liabilities
R75 = Inventory turnover
R76 = Working capital requirements turnover
R77 = Stocks / Total Assets
R78 = Size[Ln (total assets)]
R79 = Total Liabilities / Total Assets
R80 = Growth rate of real assets = (Total Assets t – Total Assets t-1) / Total Assets t-1
R81 = Growth rate of Equity – Growth rate of assets
R82 = Added value t – Added value t-1 / Added value t-1
R83 = Added value / Total Liabilities
R84 = Net fixed assets / Total Assets
R85 = Working Capital/ Cash-flow
R86 = 1 if net income is negative for the past two years, zero otherwise
R87 = 1 if total liabilities exceed total assets, zero otherwise

Appendix 2:

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
R1	,991	1,348	1	150	,247
R2	,850	26,417	1	150	,000
R3	1,000	,000	1	150	,989
R4	,926	12,027	1	150	,001
R5	,928	11,667	1	150	,001
R6	,864	23,515	1	150	,000
R7	,883	19,885	1	150	,000
R8	,887	19,065	1	150	,000
R9	,990	1,540	1	150	,216
R10	,998	,234	1	150	,629
R11	,993	1,093	1	150	,298
R12	,849	26,615	1	150	,000
R13	,998	,358	1	150	,551
R14	,976	3,721	1	150	,056
R15	,828	31,080	1	150	,000
R16	,995	,780	1	150	,379
R17	,994	,878	1	150	,350
R18	,943	9,010	1	150	,003
R19	,759	47,732	1	150	,000
R20	1,000	,028	1	150	,868
R21	,981	2,836	1	150	,094
R22	,978	3,432	1	150	,066
R23	,982	2,808	1	150	,096
R24	,979	3,140	1	150	,078

R25	,848	26,807	1	150	,000
R26	,652	79,976	1	150	,000
R27	,998	,352	1	150	,554
R28	,859	24,701	1	150	,000
R29	,987	1,919	1	150	,168
R30	,997	,427	1	150	,514
R31	,890	18,517	1	150	,000
R32	,999	,110	1	150	,740
R33	,883	19,909	1	150	,000
R34	,968	4,950	1	150	,028
R35	,993	1,073	1	150	,302
R36	,995	,730	1	150	,394
R37	,986	2,159	1	150	,144
R38	,959	6,447	1	150	,012
R39	,993	1,030	1	150	,312
R40	,981	2,921	1	150	,090
R41	,977	3,575	1	150	,061
R42	,970	4,677	1	150	,032
R43	,865	23,501	1	150	,000
R44	,857	24,960	1	150	,000
R45	,979	3,290	1	150	,072
R46	,978	3,435	1	150	,066
R47	,813	34,409	1	150	,000
R48	,834	29,925	1	150	,000
R49	,999	,193	1	150	,661
R50	,832	30,369	1	150	,000
R51	,858	24,904	1	150	,000
R52	,957	6,773	1	150	,010
R53	,916	13,843	1	150	,000
R54	,999	,106	1	150	,746
R55	,984	2,372	1	150	,126
R56	,977	3,552	1	150	,061
R57	,999	,225	1	150	,636
R58	,990	1,559	1	150	,214
R59	,991	1,396	1	150	,239
R60	,999	,200	1	150	,655
R61	,970	4,629	1	150	,033
R62	,923	12,465	1	150	,001
R63	,985	2,244	1	150	,136
R64	,933	10,785	1	150	,001
R65	,918	13,351	1	150	,000
R66	,990	1,540	1	150	,216
R67	,992	1,232	1	150	,269
R68	,980	3,008	1	150	,085
R69	,996	,541	1	150	,463
R70	,944	8,910	1	150	,003
R71	,833	29,967	1	150	,000
R72	,985	2,292	1	150	,132
R73	,817	33,588	1	150	,000
R74	,829	30,994	1	150	,000
R75	,998	,268	1	150	,605
R76	,995	,738	1	150	,392
R77	,943	9,010	1	150	,003
R78	,958	6,633	1	150	,011
R79	,785	41,038	1	150	,000
R80	,963	5,803	1	150	,017
R81	,992	1,141	1	150	,287
R82	1,000	,042	1	150	,838
R83	,891	18,401	1	150	,000
R84	1,000	,045	1	150	,832

R85	,988	1,898	1	150	,170
R86	,799	37,684	1	150	,000
R87	,765	45,996	1	150	,000

Appendix 3:

Table 15: The results of literature review

Authors	**Year**	**Method**	**Percentage of correct classification**		
			One year	**Two years**	**Three years**
Altman	1968	MDA	95%	72%	48%
Altman et al	1977	MDA	92,8%	89%	83,5%
Altman et al	1994	MDA	93,2%	88,2%	91,1%
Altman et al	1985	MDA	100%	75%	50%
BACK et al	1996	MDA	85,14%	78,38%	72,97%
Blum	1974	MDA	87%	79%	72%
Boyacioglu et al	2009	MDA	68,18%		
Brabazon et KEENAN	2004	MDA	80,67%	72%	
Brabazon et Keenan	2004	MDA	76% c	69,33% c	64,67% c
CALIA et GANUCI	1997	MDA	60,9%		
Charitou et al	2004	MDA	82,5%	62,5%	68%
Dambolena et Khoury	1980	MDA	91,2%	84,8%	82,6%
DEAKIN	1972	MDA	87% (c)	82% (c)	
DEAKIN	1972	MDA	91,2%	84,8%	
Dimitras et al	1999	MDA	90%	81,3%	77,5%
Gombola et al	1987	MDA	89%	70%	78%
Izan	1984	MDA	100%	70%	40%
Jae H. Min, Young-Chan Lee	2005	MDA	78,81%		
KIRA et al	1997	MDA	93,3%		
Levitan et al	1985	MDA	95%	91%	83%
Moyer	1977	MDA	84,1%	76,6%	68,2%
Myoung-Jong Kim, Dae-Ki Kang	2012	MDA	71,02%		
Rafiei et al	2011	MDA	80,6%		
Serrano-canca et al	2013	MDA	91,79%		
Sharma et Mahajan	1980	MDA	91,7%	78,3%	73,9%
Weinrich	1978	MDA	89%	84,3%	78,1%
WILSON et SHARDA	1994	MDA	88,65%		
Wu et al	2007	MDA	87,5%	85,22%	75%
Yi-Chung Hu et Fang-MeiTseng	2005	MDA	77,94%		
Yu et al	2014	MDA	86,5%		

Table 16: The results of literature review

Authors	Year	Method	Percentage of correct classification					
			Distressed			healthy		
			1year	2 years	3 years	1year	2 years	3 years
AGARWAL	1999	MDA	93,70%			99,00%		
ALTMAN	1968	MDA	93,39	71,2%	48,3%	97%	93,9%	
ALTMAN	1983	MDA	94,2%			92,4%		
ALTMAN et al	1994	MDA	92,8%	90,3%		96,5%	86,4%	
BACK et al	1996	MDA	86,49%	75,68%	83,78%	83,78%	81,08%	62,16%
Brabazon et Keenan	2004	MDA	82,7%	74,7%	65,3%	78,7%	69,3%	66,7%
Cadden	1991	MDA	80%	60%	60%	90%	80%	70%
Dambolena et Khoury	1980	MDA	83%	83%	78%	100%	87,%	87%
Deakin	1972	MDA	77%	96%	94%	82%	92%	82%
Diamond J.R	1976	MDA	97,3%	87,8%	80%	90,7%	85,3%	80%
Dimitras et al	1999	MDA	87,5%	75%	67,5%	92,5%	87,5%	87,5%
Dwyer	1992	MDA	76%	70%	43%	57%	54%	57%
Gloubos et Grammatikos	1988	MDA	66,7%	60,9%	64,3%	66,7%	82,6%	85,7%
Laitinen	1991	MDA	90%	72,5%	57,5%	87,5%	65%	52,5%
Moyer	1977	MDA	95%	80%	70%	82%	86%	73%
ROSE et GIROUX	1984	MDA	84,6% (1c)	87,5% (2c)		97,1% (1c)	96,2% (c)	
TAFFLER	1982	MDA	87,9% (1c)	48% (c)		100% (1c)		

Appendix 4:

Table 33: The results of literature review

Authors	Year	Method	Percentage of correct classification		
			One year	**Two years**	**Three years**
Ahn & al	2011	LR	89,47%		
Aziz & al	1988	LR	91,8%	84,7%	78,6%
Back & al	1996	LR	96,49%	71,6%	74,3%
Barniv & Hershbarger	1990	LR	91,1%	85,7%	
Barniv& MCdonald	1992	LR	83,7%	80%	71,9%
Boyacioglu & al	2009	LR	81,81%		
Charalambous & al	2000	LR	82,3%	74,5%	69,8%
Charitou & al	2004	LR	80,95%	73,81%	72,92%
Dimitras & al	1999	LR	90%	82,5%	78,75%
Min & Lee	2005	LR	79,31%		
Kira & al	1997	LR	95,5%		
Laitinen & Laitinen	1998	LR	86,6%	68,3%	
Laitinen & Laitinen	2001	LR	74,7%	65,3%	
Lau	1987	LR	80%	79%	85%
Min & al	2006	LR	78,13%		
Nam & Jinn	2000	LR	84,4%	76,1%	76,1%
Ohlson	1980	LR	82,84%	86%	
Olson & al	2012	LR	79,8%		
Serrano-canca & al	2013	LR	95,36%		
Tserng & al	2011	LR	73,61%		
Wang & al	2014	LR	73,9%		
Wilcox	1973	LR	94%	90%	88%
Wu & al	2007	LR	92,05%	89,78%	80,68%
Zavgren	1985	LR	96%	96%	96%
Chen & al	2006	LR	84,68%		

Appendix 5:

Table 34 : The Ratios Retained by the M.L.P. Method

Ratios	Formulas
R_7	Current assets / Total assets
R_{15}	Equity / Total assets
R_{19}	Short-Term Debt / Total Liabilities
R_{26}	Amortization of Capital Assets / Gross Fixed Assets
R_{40}	current assets (excluding stock) / Total assets
R_{73}	Net income / Turnover
R_{78}	Size Ln (Total assets)
R_{79}	Total Liabilities / Total Assets
R_{84}	Total Fixed Asset / Total assets

Table 35 : Network information

Entrance stratum	Covariables	1	R7
		2	R15
		3	R19
		4	R26
		5	R40

		6	R73
		7	R78
		8	R79
		9	R84
	Number of units[a]		9
	Rescaling method for covariates		standardized
Hidden stratum (s)	Number of hidden layers		1
	Number of units in the hidden stratum 1[a]		6
	Activation function		Hyperbolic tangent
Output stratum	Dependent variables	1	Y
	Number of units		2
	Activation function		MaxMou
	Error function		Cross entropy

a. Exclusion of the biased unit

Table 36 : Summary of models

Learning	Cross entropy error	,087
	Incorrect percentage forecasts	,0%
	Stopping the rule used	1 consecutive step (s) without decrease in error
	Duration of training	00:00:00,016

Dependent variable: Y
a. Error calculations are based on the test sample.

Appendix 6:

Table 57: Comparison of the two models applied to initial samples

	MDA (16 ratios)	MLP (16 ratios)	SVM (16 ratios)	LR (12 ratios)	MLP (12 ratios)	SVM (12 ratios)	MLP (9 ratios)	SVM (9 ratios)
a) one year before distress								
- % of correct classification	100 %	100 %	100 %	100 %	100 %	98,68 %	100 %	99,34 %
- % of error classification	0 %	0 %	0 %	0 %	0 %	1,32 %	0 %	0,66 %
- Error du type I	0 %	0 %	0 %	0 %	0 %	2,63 %	0 %	1,32 %
- Error du type II	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
b) two years before distress								
- % of correct classification	96,71 %	98,68 %	94,74 %	99,34 %	98,03 %	93,42 %	99,34 %	95,39 %
- % of error classification	3,29 %	1,32 %	5,26 %	0,66 %	1,97 %	6,58 %	0,66 %	4,61 %
- Error du type I	6,58 %	2,63 %	10,53 %	1,32 %	2,61 %	13,16 %	1,32 %	9,21 %
- Error du type II	0 %	0 %	0 %	0 %	1,32 %	0 %	0 %	0 %
c) three years before distress								
- % of correct classification	**95,4 %**	**96,05** %	**93,42 %**	**96,71 %**	**96,71 %**	**91,45 %**	**97,37 %**	**93,42 %**
- % of error classification	4,6 %	3,95 %	6,58 %	3,29 %	3,29 %	8,55 %	2,63 %	6,58 %
- Error du type I	7,89 %	3,95 %	13,16 %	6,58 %	3,95 %	17,11 %	1,32 %	11,84 %
- Error du type II	1,32 %	3,95 %	0 %	0 %	2,61 %	0 %	3,95 %	1,32 %

Table 58: Comparison of the two models applied to control samples

	MDA (16 ratios)	MLP (16 ratios)	SVM (16 ratios)	LR (12 ratios)	MLP (12 ratios)	SVM (12 ratios)	MLP (9 ratios)	SVM (9 ratios)
a) one year before distress								
- % of correct classification	93,34 %	98,33 %	96,67 %	**95 %**	98,33 %	**95 %**	100 %	**96,67 %**
- % of error classification	6,66 %	1,67 %	3,33 %	5 %	1,67 %	5 %	0 %	3,33 %
- Error du type I	10 %	3,33 %	6,67 %	10 %	0 %	6,67 %	0 %	6,67 %
- Error du type II	3,33 %	0 %	0 %	0 %	3,33 %	3,33 %	0 %	0 %
b) two years before distress								
- % of correct classification	95 %	100 %	96,67 %	98,33 %	100 %	98,33 %	100 %	95 %
- % of error classification	5 %	0 %	3,33 %	1,67 %	0 %	1,67 %	0 %	5 %
- Error du type I	6,67 %	0 %	6,67 %	3,33 %	0 %	3,33 %	0 %	10 %
- Error du type II	3,33 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %

c) three years before distress								
- % of correct classification	**91,67 %**	**98,33 %**	**93,33 %**	96,67 %	**96,67 %**	**95 %**	**98,33 %**	**95 %**
- % of error classification	8,33 %	1,67 %	6,67 %	3,33 %	3,33 %	5 %	1,67 %	5 %
- Error du type I	6,67 %	0 %	10 %	6,67 %	3,33 %	6,67 %	0 %	6,67 %
- Error du type II	10 %	3,33 %	3,33 %	0 %	3,33 %	3,33 %	3,33 %	3,33 %