Accounting and Management Information Systems
Vol. 10, No. 3, pp. 351–374, 2011

A HYBRID DEVICE OF SELF ORGANIZING MAPS (SOM) AND MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARS) FOR THE FORECASTING OF FIRMS' BANKRUPTCY

Javier DE ANDRÉS[1]
University of Oviedo, Department of Accounting, Spain

Fernando SÁNCHEZ-LASHERAS
University of Oviedo, Department of Construction and Manufacturing Engineering, Spain

Pedro LORCA
University of Oviedo, Department of Accounting, Spain

Francisco Javier DE COS JUEZ
University of Oviedo, Department of Exploitation and Exploration of Mines, Spain

ABSTRACT

This paper proposes a hybrid approach to the forecasting of firms' bankruptcy of Spanish enterprises from the construction sector. Our proposal starts splitting the group of healthy companies into two subgroups: borderline and non-borderline companies. Borderline companies are healthy companies with marked financial similarities with bankrupt ones. Then, each subgroup is divided in clusters according to their financial similarities and then each cluster is replaced by a director vector which represents the companies included in the cluster. In order to do this, we use Self Organizing Maps (SOM). Once the companies in clusters have been replaced by director vectors, we estimate a classification model through Multivariate Adaptive Regression Splines (MARS). Our results show that the proposed hybrid approach is much more accurate for the identification of the companies that go bankrupt than other approaches such as a multi-layer perceptron neural network and a simple MARS model.

Bankruptcy, Self Organized Maps (SOM), Multivariate Adaptive Regression Splines (MARS), Construction firms

[1] *Correspondence address:* Javier De Andrés, University of Oviedo, Department of Accounting, Avda. del Cristo s/n, Oviedo 33006, Spain; tel. +34985103287; fax +34985109599, E-mail: jdandres@uniovi.es

INTRODUCTION

During the last years the importance of bankruptcy forecasting models has been very high due to the current financial crisis, which demands an even more careful management of financial resources. Furthermore, under Basel II Accord recommendations (Bank for International Settlements, 2006), banks which choose to develop their own empirical model to quantify required capital for credit risk (Internal Rating-Based Approach) are required to maintain less capital than those using the Standardized Approach.

According to Sueyoshi and Goto (2009a), research on bankruptcy-based performance assessment can be classified into three broad categories. First, those studies centered on a particular model, which test how such model performs in comparison with others. Second, research focused on the selection of an appropriate set of variables to implement a particular model. The third category comprises papers which investigate the bankruptcy process.

Among these categories, the first is the one which has received most attention by researchers. The tested models are mainly statistical methodologies (for a review of the most outstanding studies see Keasey & Watson, 1991; Balcaen & Ooghe, 2006 among others) and Artificial Intelligence techniques (for a review see, e.g., Aziz & Dar, 2006; Ravi Kumar & Ravi, 2007).

Ravi Kumar and Ravi (2007) discuss the models which have been most frequently used in studies focused in insolvency prediction via intelligent systems. These models are Fuzzy Logic (FL), Neural Networks (NN), Genetic Algorithms (GA), Case-Based Reasoning Systems (CBR), Rough Sets (RS), Support Vector Machines (SVM), Decision trees (DT), Data Envelopment Analysis (DEA) and Hybrid Systems (HS). Among these, HS are the most promising. These combine two or more intelligent techniques in several forms to derive the advantages of all of them. HS have received considerable attention from researchers as they amplify the advantages of the intelligent techniques while simultaneously nullifying their disadvantages. Most HS require a considerable amount of data to reach to accurate estimations. This is not a problem nowadays, as publicly available databases containing financial information of listed and unlisted firms exist.

However, studies using HS for bankruptcy prediction suffer from a drawback which is that the majority of them estimate the model upon the basis of a sample in which non-failed companies are underrepresented. In most cases a matched-pairs design is used. The selection of non-failed firms is arbitrary, which makes the model to achieve a high in-sample percentage of correct classifications but it is likely to be inaccurate for failure prediction in new cases drawn from a real population.

Another strategy is to consider a “real” population as the sample. That is, to consider all the companies for which we have financial information available. However, as only a very small percentage of firms enter into financial distress in a normal economic situation, such samples are very unbalanced. This causes coefficient instability and leads to poor performance ability of the models.

As an alternative to both strategies we propose a HS model where, upon the basis of a real population of firms, data are preprocessed to summarize the information of healthy firms. So, the initial unbalanced sample is transformed into a balanced one which retains the main features of the healthy firms. Self Organized Maps (SOM) is used in this stage. Then a classification device is developed upon the transformed sample, for which we use the Multivariate Adaptive Regression Splines (MARS) approach. The results are compared with benchmarks which are popular in bankruptcy prediction literature. As an important application of the combined approach, this paper applies it to the solvency assessment of Spanish construction firms.

The remainder of the paper is structured as follows. Section 1 revises prior studies on bankruptcy prediction using HS. Section 2 is devoted to build the database. Section 3 describes the algorithm and the analytical procedures we used. Section 4 comments on the main results, including the benchmark techniques applied. Finally, section 5 is devoted to the summary and main conclusions, including also some further research avenues.

1. PRIOR BANKRUPTCY RESEARCH USING HYBRID SYSTEMS

Basically, there are four types of HS which have been applied to financial distress prediction:

- Hybrid Algorithms (HA).
- Ensemble Classifiers (EC).
- Feature Selectors (FS).
- Clustering and Classificatory devices (CC).

1.1. Hybrid Algorithms

In this kind of systems two or more intelligent algorithms are tightly integrated to form a new classification device (i.e., GA-trained NN, neuro-fuzzy systems). One of the first empirical research papers using the HA approach is that of Piramuthu (1999), which proposed a hybrid algorithm of neural networks and fuzzy sets. Although the learning results of the system could be more easily understood than those of NNs, no significant improvements were obtained with regard to prediction accuracy.

Later on, the models by Tseng and Lin (2005) and Wang *et al.* (2005) also integrated fuzzy sets into, respectively, logit models and SVMs. The results of both studies were not conclusive as for only some of the considered datasets the proposed models outperformed single approaches.

A successful research line which can be included into the HS approach is the use of GAs to estimate the parameters that drive a single model. In this regard, GAs successfully replaced back-propagation algorithm for the training of NNs (see, e.g., Sexton *et al.*, 2003; Pendharkar, 2005, among others). It is also remarkable the work by Wu *et al.* (2007), which used GAs to optimize the parameters of a SVM classification device. Furthermore, another related model was that proposed by Ahn and Kim (2009), which used GAs to reach to an optimal selection of the instances to be included into a CBR system. The prediction accuracy of this model was also higher than the best performance of different NNs.

Another stream of research which has obtained good results consists in the hybridization of outranking methods (such as for example ELECTRE) and a single classification device. The works by Li and Sun (2009, 2011) are fair examples of this research line.

Finally, it is noticeable that more complex models have also been published. An interesting one is that of Chuang and Lin (2009). These authors designed a hybrid system of NNs and MARS and added a reassigning stage in which rejected good credit applicants were re-evaluated using a CBR model. The proposed model outperformed a variety of single classification devices. Another successful model which integrates more than two systems is that of Yeh *et al.* (2010). This system mainly consists of a hybrid model of RS and SVM, but efficiency estimates obtained through DEA are also considered as features for financial failure prediction.

1.2. Ensemble Classifiers

The second type of HS which have been applied to financial distress prediction are EC, which consist of multiple single classifiers whose decision is combined to form that of the combined system, usually by applying a voting scheme.

Among the EC researches some of them proceed to ensemble NN either using evolutionary computation techniques (Kim & Cho, 2008) or using the voting strategy (Tsai & Wu, 2008), others build consecutive classifiers on modified versions of one training set which are generated according to the error rate of the previous classifier, while focusing on the hardest examples of the training set (Alfaro *et al.*, 2008). Yu *et al.* (2008) propose a multistage NN ensemble learning model where the NN ensemble aggregates the decision values from the different neural ensemble members, instead of their classification results directly. Hung and Chen (2009) developed a selective ensemble of three classifiers: DT, back-propagation NN and SVM. More recently, Yu *et al.* (2010) propose a four-stage SVM based multiagent ensemble learning approach and Sun *et al.* (2011) constructed an ensemble using Single Attribute Tests (SAT) and DT, among other techniques.

Other papers increased the number of techniques in the ensemble. Karthik Chandra *et al.* (2009) developed a hybrid intelligent system through ensembling a Multilayer Perceptron (MLP), Random Forest (RF), Logistic Regression (LR), SVM, and Classification and Regression Trees (CART). In the same vein, Nanni and Lumini (2009) tested four different methods for creating an ensemble of classifiers (Bagging, Random Subspace, Class Switching, Rotation Forests), and they tested four other classifiers (Levenberg–Marquardt neural net with five hidden units, MLP with five hidden units, Radial Basis function SVM, and 5-nearest neighbor). Ensemble methods proved to be superior. Finally, Wang *et al.* (2011) conducted a comparative assessment of the performance of three popular ensemble methods (Bagging, Boosting, and Stacking) based on four base learners (LR, DT, Artificial Neural Network and SVM). Ensemble methods also outperformed base learners.

Conclusions are clear: the majority of the HS applied improve the results of the single classifiers. The only relevant work that does not evidence the superiority of ensembles is that by Kim and Cho (2008), which concluded that multiple neural network classifiers do not outperform a single best neural network classifier in many cases.

1.3. Feature Selectors

In these systems, an algorithm is used for the selection of the predictors of failure among a list of feasible variables and another model is used to predict the bankruptcy status using the selected indicators. Considering that many indicators can be computed upon the financial statements of a company, the consideration of a preprocessing stage where some of the indicators are selected for the estimation of a further model is a possibility that is worth exploring.

In this regard, a first and somewhat basic approach is to use statistical methods for the selection of the ratios which will be subsequently used for the estimation of a classification model. The most popular procedure is the analysis of the t-statistic. The works by Tsai (2009) and Ravisankar and Ravi (2010), which used NNs for the classification stage, are examples of successful applications. A most refined approach is to use multivariate statistical models (see, e.g., Yang *et al.*, 2011, which considered partial least squares to select the financial ratios to be entered into a SVM).

Furthermore, GAs has also been frequently used for feature selection. A first attempt was the work by Back *et al.* (1995), which trained a conventional NN using the ratios previously selected by a GA. This model outperformed back-propagation trained NN and other traditional classification techniques. Later on, other authors developed on this idea. We can highlight the work by Huang *et al.* (2007), which used GAs for feature selection and a hybrid SVM-GA system for classification. Their model obtained a good classification performance. Another interesting effort research is that by Li *et al.* (2010), which considered GA and statistical methods for feature selection, and CBR for classification. This approach improved the results of single CBR models.

In addition, some authors used other Artificial Intelligence-based models for the selection of indicators. We can mention the paper by Chaudhuri and De (2010), which selected the most relevant ratios using fuzzy clustering and classified through SVM, and Cho *et al.* (2010), which used DTs in the selection stage and CBR for classification.

Finally, we must highlight the paper by Chen *et al.* (2009). In this research work MARS is used for the selection of indicators and SVM for classifying firms. This model is of special interest as it outperforms not only some individual approaches (CART, SVM and MARS) but also another hybrid system which combines SVM and CART.

1.4. Clustering and Classificatory devices

These HSs preprocess the financial information on the failed and non-failed firms and identify groups based on similarities. The grouping information is used in the subsequent estimation of a classification model.

One of the first empirical research papers is that of Alam *et al.* (2000), which presented experimental results of fuzzy clustering and two SOM used as classification tools for identifying potentially failing banks. The estimated model provides an ordinal rating of the data set in terms of failing likelihood possibility.

Later, Hsieh (2005) proposed clustering algorithms for identifying unrepresentative subsamples and constructs NN using the remainder of the sample. Defu *et al.* (2008) extended the proposal of Hsieh (2005) using also DT. They tested the models in two datasets and concluded that they are efficient in comparison with benchmark methods. Boyacioglu *et al.* (2009) evaluated four different NN models (MLP, competitive learning, SOM and learning vector quantization), SVM and three multivariate statistical methods (multivariate discriminant analysis, k-means cluster analysis and logistic regression). Results showed that MLP and learning vector quantization can be considered as the most successful models in predicting the financial failure.

Finally, De Andrés *et al.* (2011) proposed a fuzzy clustering and then a MARS model was estimated on the clusterized data. They used a wide sample of 59,336 healthy and 138 failed firms. Results revealed that the proposed model outperforms single classification devices (NN, multivariate discriminant analysis, MARS).

1.5. Assessment of previous studies: strengths and weaknesses

It must be pointed out that if the bankruptcy prediction models are eventually to be used in a predictive context, the estimation samples of failing and non-failing firms should be representative of the whole population of firms (Ooghe & Joos, 1990). Nevertheless, in the great majority of the hybrid prediction models revised, the samples are not representative of the whole population. Most studies oversample failing companies because of the low frequency rate of failing firms in the economy. A common strategy is the use of matched pairs samples (on the basis of size, sector, and/or age). This can lead to biased parameter estimates especially if the sample is made up of failed firms and very healthy companies. In that case the model will achieve a high percentage of correct classifications but it is likely to be inaccurate for failure prediction in new cases drawn from a real population.

An alternate sampling strategy is to consider a real population. As Foglia *et al.* (2001) point out, this procedure increases the variance of the estimates of coefficients due to the data imbalance between healthy and bankrupted firms. An additional drawback is that, having into account that in a normal economy most companies are non-bankrupt, classifying all the firms as "not-bankrupt" would let the model reach a high percentage of correct classifications. To avoid this, the algorithm can be designed to consider the different misclassification costs (the costs of classifying as insolvent a company which is solvent are much lower than those of the opposite error). Such a model will pay more attention to accurately classifying the failing companies at the expense of more misclassifications of non-failing firms.

However, the estimation of the different misclassification costs is not straightforward as it depends on the financial decision to be taken. Furthermore, such estimation is a subjective task as it also depends on the risk profile of the agent who makes the decision.

As an alternative to both approaches, we propose a method which enables the formation of a sample which is representative of the main features of the population but retains the balanced design and the stability of the coefficients.

Our proposal is a hybrid method in which healthy companies are divided in clusters according to their financial similarities and then each cluster is replaced by a director vector which summarizes all of them. The clustering process is made by means of a SOM procedure. The most relevant reasons for choosing SOM among the different methods for clustering are the following two: first, this technique was specifically designed for multidimensional datasets, and is able to take advantage of their complexity and second, unlike other methods for data-reduction and clustering, this family of algorithms is characterized by a learning process that is constantly updated as it takes more information from the input data, improving the output dynamically over the training stage and therefore producing more reliable results.

Prior to the calculation of clusters, healthy companies are divided into two groups:

1. Companies which are actually healthy but whose financial features have a certain degree of similarity with those of failed ones. These are called "borderline" companies.
2. Companies which are healthy and whose financial features are clearly different from those of bankrupt companies.

The clustering process is carried out separately for each group of firms. Although the idea of considering a "grey zone" or group of doubtful firms has been previously introduced by other researchers (see, i.e., Alam, *et al.*, 2000; Tseng & Lin, 2005), we made the discrimination between healthy and doubtful firms on a multivariate basis by using a non-euclidean distance measure (the Mahalanobis distance).

Once the companies in clusters have been replaced by director vectors, we estimate a classification model through MARS. The reason for choosing MARS as the second part of the hybrid system lies in the fact that this technique is a flexible procedure, which models relationships that are nearly additive or involve interactions with fewer variables (Hastie & Tibshirani, 1990). MARS builds flexible models by fitting piecewise linear regressions; that is, the nonlinearity of a model is approximated through the use of separate regression slopes in a limited number of intervals of the variable space. This is made by using a procedure which is inspired by the recursive partitioning technique governing Classification And Regression Trees (CART) algorithm (Breiman *et al.*, 1984). Such features make it especially suitable for the bankruptcy prediction problem, as the variety of indicators that can be computed upon the financial statements of a firm can be considered as manifestations of a small number of financial features (i.e. profitability, solvency, etc.). So, a small number of indicators can represent most of the information contained in the annual accounts (Yli-Olli & Virtanen, 1989). Consequently, some studies (see, i.e., Lee *et al.*, 2006; Chen *et al.*, 2006) found evidence that MARS performs better than other approaches when applied to financial classification purposes.

As benchmarks for our hybrid system we estimated a simple MARS model (without the SOM-preprocessing stage) and a multilayer BP-trained NN.

2. THE DATABASE

In the present research we consider failing and nonfailing firms from the construction sector in Spain. The recent credit crisis and economic downturn have had some serious implications for the Spanish construction sector. As the economic situation changed, along with the increase in unemployment and the rise of the interest rates, the expectations of house prices' evolution that sustained demand and encouraged new developments disappeared. Consequently, firms in the real estate and construction sectors are facing difficulties and challenges which affect their future viability.

2.1. Enterprises in the sample

In Spain, bankruptcy is regulated by the Bankruptcy Act 22/2003, of 9th July. This Act contemplates a unique proceeding, which is called "bankruptcy" (span. *concurso de acreedores*). This procedure can conclude either with the approval of the settlement of creditors or with the liquidation of the company. Filing for bankruptcy does not necessarily means that the firm is insolvent. However, the recovery rate (understood as cents on the euro recouped by creditors through the regulated procedures) in Spain is lower than in many developed countries, i.e. Belgium, Denmark, Finland, Iceland, Ireland, Norway, Netherlands, Sweden, United Kingdom, Canada, United States, Hong Kong, Japan, Korea, Singapore, Taiwan, New Zealand, or Australia (IFC, 2010). So, in practice bankruptcy procedure can be understood as insolvency.

Many papers on bankruptcy prediction have focused on the manufacturing sector (i.e. Altman, 1968; Begley *et al.*, 1996; Becchetti & Sierra, 2003). Nevertheless, there are several papers examining the bankruptcy in sectors other than manufacturing. For example, telecommunications industry (Foreman, 2003); restaurant industry (Gu, 2002; Kim & Gu, 2006; Young & Gu, 2010); air carriers (Davalos *et al.*, 1999); nursing facility industry (Knox *et al.*, 2009); oil companies (Sena & Williams, 1998); retail sector (Bhargava *et al.*, 1998); construction industry (Sueyoshi & Goto, 2009b).

Therefore, a database with Spanish construction firms was drawn up. As bankrupt companies we considered those whose judicial declaration took place in 2008. In accordance with Spanish legislation, limited liability companies are required to deposit their annual accounts in the *Registro Mercantil*. This information is gathered and provided by *Bureau van Dijk* and *Informa* for Spanish firms in the SABI database, one of Europe´s leading publishers of electronic business information. We deleted from the sample companies that did not provide full information about all the variables from the year prior to bankruptcy. To avoid the distortions caused by defects in the preparation of financial information of small enterprises, whose annual accounts are generally unaudited, we also deleted from the database those firms whose total assets were below 100K €. Once these filters were applied, we obtained a final data set that was made up of 63.107 firms. Of these, a total of 256 companies went bankrupt in 2008.

2.2. The financial ratios for predicting bankruptcy

In this paper we used the five variables proposed by E.I. Altman in his seminal paper on the usefulness of linear discriminant analysis (Altman, 1968). The reasons for this choice were the following: i) these are variables that are readily available for any company. It must be borne in mind that increasing the number of variables has the undesirable effect of reducing the number of companies in the dataset, since not all companies provide equal levels of information; ii) several papers used this same set of variables to test the effectiveness of statistical techniques and/or other models for bankruptcy prediction (i.e., Odom & Sharda, 1993, for neural networks and Lizarraga

Dallo, 1998 for the logit model); iii) it should be noted that some authors (i.e., Begley *et al.*, 1996; Lizárraga Dallo, 1997, and Grice & Ingram, 2001) have studied the validity of the Altman function when applied in other geographical settings and time spans. They concluded that with a proper reassessment of the coefficients, the model proposed by Altman in 1968 remains as a valid approximation to the issue of predicting insolvency.

Therefore, the five variables used in this paper are the following:

X_1 = working capital/total assets
X_2 = retained earnings/total assets
X_3 = earnings before interest and taxes (EBIT)/total assets
X_4 = market value of equity/book value of total debt
X_5 = sales/total assets

Regarding the fourth of the variables, it should be noted that its calculation is difficult in environments where only a small percentage of companies are quoted. Therefore, in subsequent sectoral applications of this model to predict insolvency, the author replaced the market value of equity by the book value of equity (Altman, 1993). In this research we considered such a definition.

Tables 1 and *2* show some descriptive statistics for the variables.

Table 1. **Descriptive statistics (bankrupt companies)**

Var.	Q1	Median	Q3	Mean	StDev	Asym.	Kurt.
X_1	-0.138	0.006	0.157	-0.024	0.450	-1.692	7.848
X_2	-0.123	0.015	0.069	-0.122	0.412	-2.838	10.725
X_3	-0.170	0.013	0.052	-0.109	0.310	-2.277	5.742
X_4	-0.092	0.031	0.103	0.008	0.304	4.028	41.531
X_5	0.786	1.407	2.229	1.602	1.101	0.905	0.682

Table 2. **Descriptive statistics (healthy companies)**

Var.	Q1	Median	Q3	Mean	StDev	Asym.	Kurt.
X_1	-0.016	0.136	0.367	0.160	0.352	-2.560	69.241
X_2	0.025	0.126	0.310	0.163	0.321	-8.391	456.771
X_3	0.019	0.051	0.104	0.060	0.173	-5.728	234.588
X_4	0.062	0.217	0.607	1.237	51.701	239.9	59307.8
X_5	0.802	1.400	2.130	1.596	1.212	3.077	48.180

From a first examination of the information contained in *Tables 1* and *2* it is clear that the statistical distribution of the considered variables is asymmetric and extremely leptokurtic. This corroborates previous results on the statistical distribution of the financial indicators (Lau *et al.*, 1995; Martikainen *et al.*, 1995, among others) and advises against the use of parametrical models.

3. ALGORITHM AND ANALYTICAL PROCEDURE

3.1. The proposed hybrid model

The model proposed in the present research combines the use of MARS models with a clustering technique which is SOM mapping in order to obtain a MARS model which uses as training information only those companies considered as representative of each cluster. The steps of the algorithm are the following (a more detailed explanation of each one of the steps is provided in subsequent sections):

Step 1: Study of the similarities of the bankrupt companies by means of Mahalanobis' distances. The Mahalanobis distance of all the bankrupt companies was calculated.

Step 2: Those bankrupted companies that were more dissimilar to the rest of the sample were signaled as outliers and removed from the data set to be employed for step 3 although they were taken into account for the training and validation of the model. The determination of the bankrupted companies considered as outliers was done by means of the robust estimation of the parameters in the Mahalanobis distance (Rousseeuw & Van Zomeren, 1990) and the comparison with a critical value of the Chi-square distribution (in our case the 95% quantile).

Step 3: The Mahalanobis distance of each one of the non-bankrupt companies versus the set of all the bankrupted companies not considered as outliers was calculated.

Step 4: A new category of companies was created, which was called "borderline". The companies that were not considered as outliers when compared with the sample of bankrupt companies are supposed to be more likely to go bankrupt than the rest of non-bankrupted companies. Therefore they were included in this new category.

Step 5: Companies belonging to non-bankrupted and borderline populations were classified in clusters using the SOM algorithm proposed by Kohonen (1995). Two clusters of similar dimensions to the number of bankrupted companies were defined and trained with the non-bankrupted and borderline sets. This step is performed in order to obtain a more balanced set of data for the training of the models in the next steps.

Step 6: An algorithm based on the MARS model (Friedman, 1991) was fed with the reduced sets of borderline and non-bankrupt companies and the original set of bankrupt companies. The performance of this model was evaluated by means of their specificity and sensibility (more details on this point are provided below).

3.2. The Mahalanobis distance

The Mahalanobis distance (Mahalanobis, 1936) is a non-euclidean distance measure based on correlations between variables by means of which different patterns can be identified and analyzed. It is a useful way of determining the similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations in the data set and is scale-invariant, i.e. not dependent on the scale of measurements.

Given the vectors that represents the set of variables of two companies $x_1 \in \Re^n$ and $x_2 \in \Re^n$, their Mahalanobis distance can be calculated as follows:

$$d_A(x_1, x_2) = \sqrt{(x_1 - x_2)^T \cdot A \cdot (x_1 - x_2)} \tag{1}$$

Where $A \in \Re^{nxn}$ is positively semi-definite and represents the inverse of the covariance matrix of class $\{I\}$. The Mahalanobis distance is therefore a weighted Euclidean distance where the weighting is determined by the range of variability of the sample point; expressed by the covariance matrix (Avishek & Maiti, 2010). Using the eigenvalue decomposition, *A* can be decomposed into $A = W \cdot W^T$. Thus, it is also feasible to learn the matrix *W*. Then, we have

$$d_A(x_1, x_2) = \sqrt{(x_1 - x_2)^T \cdot (W \cdot W^T) \cdot (x_1 - x_2)} \tag{2}$$

3.3. Self-organized Maps neural networks

SOM is a class of neural-network algorithms which belong to the unsupervised-learning category. SOM is an algorithm used to visualize and interpret large high-dimensional data sets.

The SOM map (Jeong *et al.*, 2010) consists of a regular grid of processing units, "neurons". A model of some multidimensional observation, eventually a vector consisting of features, is associated with each unit. The map attempts to represent all the available observations with optimal accuracy using a restricted set of models. At the same time the models become ordered on the grid so that similar models are close to each other and dissimilar models far from each other.

Let *N* be the dimension of the *n* sample vectors $X(t) \in \Re^n$, $t = 1,2,..n$, where each sample vector is identified by a label. The two-dimensional output layer contains a rectangular mesh of $k = 1,...,x_{\dim} \times y_{\dim}$ nodes, each serving as codebook vector W_k of dimension *N*. The training of the weight (codebook) vectors of the map's nodes is realized by the following algorithm (Kohonen, 1995).

For a given number of iterations do:

1. Pick up randomly one sample vector $X(t)$
2. Find the nearest weight vector W_c : $\|X - W_c\| = \min_j \|X - W_j\|$
3. Update the weights W_i according to the rule:

$$W_i(t+1) = W_i(t) + h_{ci}(t)[X(t) - W_i(t)] \tag{3}$$

Where $h_{ci}(t)$ is the neighbor function that is usually of the Gaussian type:
$h_{ci}(t) = \alpha(t)\exp(-\|W_c - W_i\| / 2\sigma^2(t))$ or of a local "bubble" type (Kohonen, 1995).

Weights of neurons laying in the neighborhood $h_{ci}(t)$ of the winning neuron are moved closer to $X(t)$. The learning rate $\alpha(t) \in [0,1]$ decreases monotonically with time, $\sigma(t)$ determining that the radius of the neighborhood also decreases monotonically. After many iterations and a slow reduction of $\alpha(t)$ and $\sigma(t)$, the neighborhood covers only a single node and the map is formed: neurons with weights that are close in the parameter space *W* are also close on the mesh and can be labeled with names (classes) of input clusters. A graphical interpretation of the Mahalanobis distance can be found in the work of Maesschalck *et al.* (2000).

3.4. Multivariate Adaptive Regression Splines (MARS) model

As stated earlier, MARS is a multivariate nonparametric regression technique developed by Friedman (1991). Its main purpose is to predict the values of a continuous dependent variable, $\vec{y}(n \times 1)$, from a set of independent explanatory variables, $\vec{X}(n \times p)$. The MARS model can be represented as:

$$\vec{y} = f(\vec{X}) + \vec{e} \tag{4}$$

where $\vec{e}$ is an error vector of dimension $(n \times 1)$.

MARS can be considered as a generalization of classification and regression trees (CART) (Hastie *et al.*, 2003), and is able to overcome some of its limitations. MARS does not require any a priori assumptions about the underlying functional relationship between dependent and independent variables. Instead, this relation is covered from a set of coefficients and piecewise polynomials of degree q (basis functions) that are entirely driven from the regression data $(\vec{X}, \vec{y})$. The MARS regression model is constructed by fitting basis functions to distinct intervals of the independent variables. Generally, piecewise polynomials, also called splines, have pieces smoothly connected together. In MARS terminology, the joining points of the polynomials are called knots, nodes or breakdown points. These will be denoted by the small letter *t*. For a spline of degree *q* each segment is a polynomial function. MARS uses two-sided truncated power functions as spline basis functions. These are described by the following equations (Sekulic & Kowalski, 1992):

$$[-(x-t)]_+^q = \begin{cases} (t-x)^q & if \ x < t \\ 0 & otherwise \end{cases} \tag{5}$$

$$[+(x-t)]_+^q = \begin{cases} (t-x)^q & if \ x \geq t \\ 0 & otherwise \end{cases} \tag{6}$$

where $q(\geq 0)$ is the power to which the splines are raised and which determines the degree of smoothness of the resultant function estimate.

The MARS model of a dependent variable $\vec{y}$ with M basis functions (terms) can be written as follows (Friedman & Roosen, 1995):

$$\hat{\vec{y}} = \hat{f}_M(\vec{x}) = c_0 + \sum_{m=1}^{M} c_m B_m(\vec{x}) \tag{7}$$

where $\hat{\vec{y}}$ is the dependent variable predicted by the MARS model, c_0 is a constant, $B_m(\vec{x})$ is the m-th basis function, which may be a single spline basis function, and c_m is the coefficient of the m-th basis function.

Both the variables to be introduced into the model and the knot positions for each individual variable have to be optimized. For a data set $\vec{X}$ containing *n* objects and *P* explanatory variables, there are $N = n \times p$ pairs of spline basis functions, given by equations (5) and (6), with knot locations x_{ij} ($i = 1,2,...,n;\ j = 1,2,...,p$).

A two-step procedure is followed to construct the final model. First, in order to select the consecutive pairs of basis functions of the model, a two-at-a-time forward stepwise procedure is implemented (Friedman & Roosen, 1995). This forward stepwise selection process leads to a very complex and overfitted model. Such a model, although adequately fitting the estimation data, has poor predictive abilities for new objects. To improve the prediction, the redundant basis functions are removed one at a time using a backward stepwise procedure. To determine which basis functions should be included in the model, MARS utilizes the generalized cross-validation (GVC) criterion (Sekulic & Kowalski, 1992). GVC is the mean squared residual error divided by a penalty which is dependent on model complexity. Then, GVC is defined in the following way:

$$GVC(M) = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}_M(\vec{x}_i)\right)^2}{(1 - C(M)/n)^2} \tag{8}$$

where $C(M)$ is a complexity penalty that increases with the number of basis functions in the model and which is defined as:

$$C(M) = (M + 1) + d\,M \tag{9}$$

where M is the number of basis functions in equation 7, and the parameter d is a penalty for each basis function included into the model. d can be also regarded as a smoothing parameter. In the present research, d equals 2. This value can be chosen by model user but it must be remarked that a smaller d generates a larger model with more basis functions; a larger d creates a smaller model with less basis functions (Kriner, 2007). Further details about the selection of the d parameter can be seen in Friedman (1991).

The main steps of the MARS algorithm as applied in this research can be summarized as follows (Sekulic and Kowalski, 1992):

1. Select the maximum allowed complexity for the model and define the d parameter.
2. Forward stepwise selection:
   a. Start with the simplest model, i.e. with the constant coefficient only.
   b. Explore the space of the basis functions for each explanatory variable.
   c. Determine the number of basis functions (M) that minimizes the prediction error and include them into the model.
   d. Go to step 2.b until a model with a predetermined complexity is derived.
3. Backward stepwise selection:
   a. Search the entire set of basis functions (excluding the constant) and delete from the model the one that contributes least to the overall goodness of fit using the GCV criterion.
   b. Repeat 3.b until GCV reaches its maximum.

The predetermined complexity of MARS model in step 3 should be considerably larger than the optimal (minimal GCV) model size M^*, so choosing $2M^*$ as the minimum predetermined complexity for the model is enough in general (Friedman & Roosen, 1995).

The predictive ability of the MARS model can be evaluated in terms of the root mean squared error of cross-validation (RMSECV) and the squared leave-one-out correlation coefficient (q^2). To compute RMSECV, one object is left out from the data set and the model is constructed for the remaining $n-1$ objects. Then the model is used to predict the value for the object which is left out. When all objects have been left out once, RMSECV is given by the following expression (Friedman & Roosen, 1995):

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_{-i}\right)^2}{n}} \tag{10}$$

where y_i is the value of the dependent variable of the *i-th* object and $\hat{y}_{-i}$ is the predicted value of the dependent variable of the *i-th* object with the model built without the *i-th* object.

The value of q^2 is given as:

$$q^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_{-i})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (11)$$

where $\bar{y}$ is the mean value of the dependent variable for all n objects.

Finally, we must comment on the procedure used to assess the performance of the model. The first measure is accuracy, which is the global percentage of correct classifications. We also computed the sensitivity, which is the percentage of bankrupt companies which were correctly classified. The last measure is specificity, which is the proportion of healthy companies correctly identified.

4. RESULTS

In this section we detail the results of the algorithm, as well as those of the benchmark techniques.

4.1. The algorithm

First, Table 3 details the number of clusters for each set (non-bankrupt and borderline companies). All companies belonging to non-bankrupt and borderline populations were classified in clusters using SOM. The clusters were obtained as the output of step 5 of the algorithm. As can be observed, the number of clusters used for the models is 256. This means that the original SOM was of (16x16) neurons. Please note that each cluster is represented by a director vector. A director vector (Perner, 2008) can be described as the expected value for each one of the independent variables for all the companies that belong to a certain cluster. Models with less neurons were tested but not included in the present research due to their lower performance. As it was already mentioned before, this step was performed in order to obtain a more balanced set of data for the training of the models in the following steps, in which each cluster was represented by a director vector that aims to summarize the information of all the individuals contained in each subset.

Table 3. **Number of clusters used for the model**

Number of director vectors (clusters)	
Non-bankrupt companies	**Borderline companies**
256	256

An algorithm based on MARS models (step 6) was then used for the implementation of a predictive model. In order to reach this aim, this model was trained using a set which comprises (a) all the bankrupted companies, (b) the director vectors corresponding to non-bankrupt non-borderline companies and (c) the director vectors corresponding to non-bankrupt borderline companies. The validation was made by calculating the confusion matrix using the information of the original database. Table 4 shows the average percentage of correctly classified companies of the mentioned model. The last column of the mentioned table represents the total percentage of companies of the database that were correctly classified by the model. This is the most important parameter as it gives us an outlook of the global performance of the model.

Table 4. **Average percentage of companies that are correctly classified in their corresponding category**

% of companies correctly classified				
Bankrupt	**Non-bankrupt**	**Borderline**	**Non-bankrupt + Borderline**	**Total**
88.70	60.40	91.60	84.63	84.29

In order to validate the predictive model we repeated the estimation process five times. For each run we randomly divided the original sample into a training subsample, which contained 80% of the non-bankrupt firms and 80% of the bankrupt firms, and a validation subsample (remaining 20% of the bankrupt and non-bankrupt companies).

Table 5 contains a confusion matrix in which the mean values obtained in the validation of the results of the five different runs are shown.

Table 5. **Confusion matrix: average values of the validation results of 5 different runs**

		Real category	
		Non-bankrupt	**Bankrupt**
Predicted category	**Non-bankrupt**	11,405	8
	Bankrupt	1,447	44

In addition, according to the information contained in *Table 5* it must be remarked that the specificity of the model is 88.74%, that is, it is able to detect 88.74% of the companies that did not go bankrupt. It also detects 84.61% of all those companies that went bankrupt (sensitivity). Finally, we must also underline that the global accuracy of the model is 88.72%.

4.2. Benchmark techniques

As indicated above, the benchmark techniques used to compare with the results obtained by the algorithm proposed in the present paper were two: back propagation NN and MARS. The model has 5 neurons in the input layer and 7 in the intermediate. The MARS model obtained was of degree 2 although no maximum degree condition was imposed.

For the estimation of the accuracy of NN and MARS, we followed a procedure similar to that proposed to test the accuracy of the proposed algorithm. NN and the MARS model were applied to five random selected training data bases (80% of the data chosen at random) and tested over their corresponding validation subsets (the remaining 20% of the database).

For the case of the NN model, the results obtained in the five runs yielded an average specificity of 99.95 %, an average sensitivity of 21.00 % and an average global accuracy of 99.01%.Although the specificity the NN-based device is higher than that of our proposal, it is inefficient for the detection of bankrupt companies, due to its low sensitivity. This makes this model useless for decision-aid purposes because the costs of the error consisting in considering a bankrupt company as non-bankrupt are very much higher than that of the opposite error.

The results obtained for the MARS model were as follows: average specificity of 99.79 %, average sensitivity of 3.85 % and average global accuracy of 99.78%. These results are even worse than those of NN, so it can be concluded that the MARS model is also useless for practical purposes.

CONCLUSIONS

This paper proposes a new approach to the forecasting of firms' bankruptcy. Our proposal is a hybrid method in which healthy companies are divided in clusters according to their financial similarities and then each cluster is replaced by a director vector which summarizes all of them. In order to do this, we used SOM mapping. Once the companies in clusters have been replaced by director vectors, we estimated a classification model through MARS.

For the test of the model we considered a real setting of Spanish enterprises from the construction sector because of the importance of this branch of activity in the Spanish economy. It is also remarkable that in our dataset the proportion of distressed firms is very close to that which is derived from Economic statistics. We also used two benchmark techniques to compare with the results obtained by the algorithm proposed in the present paper: a back propagation neural network and a MARS model.

Our results show that the proposed hybrid approach is much more accurate than the benchmark techniques for the identification of the companies that go bankrupt. As future research efforts we can mention the application of the procedure proposed in the present research to other related tasks in the field of financial statements analysis (i.e. prediction of takeovers, analysis of bond ratings, etc.). It could be also of interest the use of other models apart from MARS in the classification stage of the algorithm (e.g. SVM).

REFERENCES

Ahn, H. & Kim, K. (2009) "Bankruptcy prediction modelling with hybrid case-based reasoning and genetic algorithms approach", *Applied Soft Computing*, vol. 9, no. 2: 599-607

Alam, P., Booth, D., Lee, K. & Thordarson T. (2000) "The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: an experimental study", *Expert Systems with Applications,* vol. 18: 185–199

Alfaro, E., García, N., Gámez, M. & Elizondo, D. (2008) "Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks", *Decision Support Systems*, vol. 45: 110-122

Altman, E.I. (1968) "Financial ratios, discriminant analysis and the prediction of the corporate bankruptcy", *Journal of Finance,* vol. 23, no. 4: 589-609

Altman, E.I. (1993) *Corporate Financial Distress and Bankruptcy,* New York: John Wiley and Sons

Avishek, P. & Maiti J. (2010) "Development of a hybrid methodology for dimensionality reduction in Mahalanobis-Taguchi system using Mahalanobis distance and binary particle swarm optimization", *Expert Systems with Applications*, vol. 37, no. 2: 1286-1293

Aziz, M.A. & Dar, H.A. (2006) "Predicting corporate bankruptcy: where we stand?", *Corporate Governance*, vol. 6, no. 1: 18–33

Back, B., Laitinen, T., Sere, K. & Van Wezel, M. (1995) "Choosing bankruptcy predictors using discriminant analysis, logit analysis and genetic algorithms", in Sierra Molina, G. and Bonsón Ponte, E. (Eds.): *Artificial Intelligence in Accounting, Finance and Tax*, pp. 337-356

Balcaen, S. & Ooghe, H. (2006) "35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems", *The British Accounting Review*, vol. 38: 63–93

Bank for International Settlements (BIS) (2006) *International Convergence of Capital Measurement and Capital Standards. A Revised Framework*, Basel: BIS

Becchetti, L., & Sierra, J. (2003) "Bankruptcy risk and productive efficiency in manufacturing firms", *Journal of Banking and Finance,* vol.27: 2099-2120

Begley, J., Ming, J. & Watts, S. (1996) "Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's models", *Review of Accounting Studies*, vol. 1, no. 4: 267-284

Bhargava, M., Dubelaar, Ch. & Scott, T. (1998) "Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance", *Journal of Retailing and Consumer Services*, vol. 5, no. 2: 105-117

Boyacioglu, A., Kara, Y. & Baykan, O.M. (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*, vol. 36: 3355–3366

Breiman, L., Friedman, J.H., Olshen, R. & Stone, C.J. (1984) *Classification and Regression Trees*, Belmont: Wadsworth International Group

Chaudhuri, A. & De, K. (2011) "Fuzzy Support Vector Machine for bankruptcy prediction", *Applied Soft Computing*, vol. 11, no. 2: 2472-2486

Chen, W., Ma, C. & Ma, L. (2009) "Mining the customer credit using hybrid support vector machine technique", *Expert Systems with Applications*, vol. 36, no. 4: 7611-7616

Chen, L.H., Song, J.S. & Ji, F. (2006) "MARS-based research of personal credit scoring: verification of Chinese data", *Proceedings of the International Conference on Management Science and Engineering*, 5-7 October: 1498-1501

Cho, S., Hong, H. & Ha, B.Ch. (2010) "A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction", *Expert Systems with Applications*, vol. 37, no. 4: 3482-3488

Chuang, Ch.L. & Lin, R.H. (2009) "Constructing a reassigning credit scoring model", *Expert Systems with Applications*, vol. 36, no. 2: 1685-1694

Davalos, S., Gritta, R.D. & Chow, G. (1999) "The application of a neural network approach to predicting bankruptcy risks facing the major US air carriers: 1979–1996", *Journal of Air Transport Management,* vol.5, no. 2: 81-86

De Andrés, J., Lorca, P., de Cos Juez, F.J. & Sánchez-Lasheras, F. (2011) "Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS) ", *Expert Systems with Applications*, vol. 38, no. 3: 1866-1875

Defu, Z., Leung, S.C.H. & Zhimei, Y. (2008) "A decision tree scoring model based on genetic algorithm and k-means algorithm", *Third International Conference on Convergence and Hybrid Information Technology ICCIT '08*, vol. 1, 11-13 Nov. 2008: 1043-1047

Foreman, R.D. (2003) "A logistic analysis of bankruptcy within the US local telecommunications industry", *Journal of Economics and Business*, vol. 5, no. 2: 135-166

Friedman, J.H. (1991). "Multivariate adaptive regression splines", *Annals of Statistics*, vol. 19: 1-141

Friedman, J.H. & Roosen, C.B. (1995) "An introduction to multivariate adaptive regression splines", *Statistical Methods in Medical Research*, vol. 4: 197-217

Grice, J.S. & Ingram, R.W. (2001) "Test of generalizability of Altman's bankruptcy prediction model", *Journal of Business Research,* vol. 54, no. 1: 53-61

Gu, Z. (2002) "Analyzing bankruptcy in the restaurant industry: A multiple discriminant model", *International Journal of Hospitality Management*, vol. 21, no. 1: 25-42

Hastie, T. & Tibshirani, R. (1990) *Generalized Additive Models*, London: Chapman and Hall

Hastie, T., Tibshirani, R. & Friedman, J.H. (2003) *The Elements of Statistical Learning*, New York: Springer-Verlag

Huang, C.L., Chen, M.C. & Wang, C.J. (2007) "Credit scoring with a data mining approach based on Support Vector Machines", *Expert Systems with Applications*, vol. 33: 847-856

Hung, Ch. & Chen, J.H. (2009) "A selective ensemble based on expected probabilities for bankruptcy prediction", *Expert Systems with Applications*, vol. 36, no. 3: 5297-5303

Hsieh, N.Ch. (2005) "Hybrid mining approach in the design of credit scoring models", *Expert Systems with Applications*, vol. 28: 655-665

International Finance Corporation (IFC) (2010) *Doing Business 2011. Making a Difference for Entrepreneurs*, Washington: Palgrave Macmillan, World Bank/International Finance Corporation

Jeong K., Hong D., Byeon M., Jeong J., Kim H., Kim D. & Joo G. (2010) "Stream modification patterns in a river basin: Field survey and self-organizing map (SOM) application", *Ecological Informatics*, vol. 5, no. 4: 293-303

Karthik Chandra, D., Ravi, V. & Bose, I. (2009) "Failure prediction of dotcom companies using hybrid intelligent techniques", *Expert Systems with Applications*, vol. 36, no. 3: 4830-4837

Keasey, K. & Watson, R. (1991) "Financial distress prediction models: A review of their usefulness", *British Journal of Management*, vol. 2, no. 2: 89–102

Kim, H. & Gu, Z. (2006) "Predicting restaurant bankruptcy: A logit model in comparison with a discriminant model", *Journal of Hospitality and Tourism Research,* vol.30, no. 4: 474-493

Kim, K.J. & Cho, S.B. (2008) "Evolutionary ensemble of diverse artificial neural networks using speciation", *Neurocomputing,* vol. 71: 1604-1618

Knox, K.J., Blankmeyer, E.C., Trinidad, J.A. & Stutzman, J.R. (2009) "Predicting bankruptcy in the Texas nursing facility industry", *The Quarterly Review of Economics and Finance,* vol. 49, no. 3: 1047-1064

Kriner, M. (2007) “Survival Analysis with Multivariate Adaptive Regression Splines”, PhD Thesis, Fakultät für Mathematik, Informatik und Statistik, Ludwig-Maximilians-Universität München

Kohonen T. (1995) *Self-Organizing Maps*, (1st ed.) Berlin: Springer-Verlag

Lau, H.S., Hing-Ling, A. & Gribbin, D.W. (1995) “On modelling cross sectional distributions of financial ratios”, *Journal of Business Finance and Accounting*, vol. 22, no. 4: 521-549

Lee, T.S., Chiu, Ch.Ch., Chou, Y.Ch. & Lu, Ch.J. (2006) “Mining the customer credit using classification and regression tree and multivariate adaptive regression splines”, *Computational Statistics and Data Analysis,* vol. 50: 1113-1130

Li, H., Huang, H.B., Sun, J. & Lin, Ch. (2010) “On sensitivity of case-based reasoning to optimal feature subsets in business failure prediction”, *Expert Systems with Applications*, vol. 37, no. 7: 4811-4821

Li, H. & Sun, J. (2009) “Hybridizing principles of the Electre method with case-based reasoning for data mining: Electre-CBR-I and Electre-CBR-II”, *European Journal of Operational Research*, vol. 197, no. 1: 214-224

Li, H. & Sun, J. (2011) “Predicting business failure using forward ranking-order case-based reasoning”, *Expert Systems with Applications*, vol. 38, no. 4: 3075-3084

Lizarraga Dallo, F. (1998) “Modelos de previsión del fracaso empresarial: ¿funciona entre nuestras empresas el modelo de Altman de 1968?”, *Revista de Contabilidad*, vol. 1, no. 1: 137-166

Mahalanobis, P.C. (1936) “On the generalised distance in statistics”, *Proceedings of the National Institute of Science of India*, vol. 12: 49–55

Martikainen, T., Perttunen, J., Yli-Olli, P. & Gunasekaran, A. (1995) “Financial ratio distribution irregularities: implications for ratio classification”, *European Journal of Operational Research, vol.* 80: 34-44

Maesschalck, R.D., Jouan-Rimbaud, D. & Massart, D.L. (2000) “The Mahalanobis distance”, *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 1: 1-18

Nanni, L. & Lumini, A. (2009) “An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring”, *Expert Systems with Applications*, vol. 36: 3028-3033

Odom, M.D. & Sharda, R. (1993) “A neural network model for bankruptcy prediction”, in R. R. Trippi and E. Turban (Eds.) *Neural networks in Finance and Investing. Using Artificial Intelligence to Improve Real-World Performance*: 163-168, Chicago: Probus Publishing

Ooghe, H. & Joos, P. (1990) *Failure prediction, explanation of misclassifications and incorporation of other relevant variables: result of empirical research in Belgium*, Working paper, Department of Corporate Finance, Ghent University (Belgium)

Pendharkar, P.C. (2005) “A threshold-varying artificial neural network approach for classification and its application to bankruptcy prediction problem”, *Computers & Operations Research*, vol. 32, no. 10: 2561-2582

Perner P. (2008) *Advances in Data Mining - Medical Applications, E-commerce, Marketing, and Theoretical Aspects*, (1st ed.). Berlin: Springer-Verlag

Piramuthu, S. (1999) "Financial credit-risk evaluation with neural and neurofuzzy systems", *European Journal of Operational Research*, vol. 112: 310-321

Ravi Kumar, P. & Ravi, V. (2007) "Bankruptcy prediction in banks and firms via statistical and intelligent techniques-A review", *European Journal of Operational Research*, vol. 180, no. 1: 1-28

Ravisankar, P., & Ravi, V. (2010) "Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP", *Knowledge-Based Systems*, vol. 23, no. 8: 823-831

Rousseeuw P.J. & Van Zomeren B.C. (1990) "Unmasking multivariate outliers and leverage points", *Journal of the American Statistical Association*, vol. 85, no. 411: 633-651

Sekulic, S. & Kowalski, B.R. (1992) "MARS: a tutorial", *Journal of Chemometrics*, vol. 6: 199-216

Sena, J. & Williams, D. (1998) "Using the Altman bankruptcy model to analyze the performance of oil companies", *Petroleum Accounting and Financial Management Journal,* April 1.

Sexton, R.S., Sriram, R.S. & Etheridge, H. (2003) "Improving Decision Effectiveness of Artificial Neural Networks: A Modified Genetic Algorithm Approach", *Decision Sciences*, vol. 34, no. 3: 421-442

Sueyoshi, T. & Goto, M. (2009a) "Methodological comparison between DEA (data envelopment analysis) and DEA–DA (discriminant analysis) from the perspective of bankruptcy assessment", *European Journal of Operational Research*, vol. 199, no. 2: 561-575

Sueyoshi, T. & Goto, M. (2009b) "DEA-DA for bankruptcy-based performance assessment: Misclassification analysis of Japanese construction industry", *European Journal of Operational Research*, vol. 199, no. 2: 576-594

Sun, J., Jia, M. & Li, H, (2011) "AdaBoost ensemble for financial distress prediction: An empirical comparison with data from Chinese listed companies", *Expert Systems with Applications*, vol. 38, no. 8: 9305-9312

Tsai, Ch.F. (2009) "Feature selection in bankruptcy prediction", *Knowledge-Based Systems*, vol. 22, no. 2: 120-127

Tsai, Ch.F. & Wu, J.W. (2008) "Using neural network ensembles for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, vol. 34: 2639-2649

Tseng, F.M. & Lin, L. (2005) "A quadratic interval logit model for forecasting bankruptcy", *Omega*, vol. 33, no. 1: 85-91

Wang, Y.Q., Wang, S.Y. & Lai, K.K. (2005) "A new fuzzy support vector machine to evaluate credit risk", *IEEE Transactions on Fuzzy Systems*, vol. 13: 820-831

Wu, Ch.H., Tzeng, G.H., Goo, Y.J. & Fang, W.Ch. (2007) "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy", *Expert Systems with Applications*, vol. 32, no. 2: 397-408

Yang, Z, You, W. & Ji, G. (2011) "Using partial least squares and support vector machines for bankruptcy prediction", *Expert Systems with Applications*, vol. 38, no. 7: 8336-8342

Yeh, Ch.Ch., Chi, D.J. & Hsu, M.F. (2010) "A hybrid approach of DEA, rough set and support vector machines for business failure prediction", *Expert Systems with Applications*, vol. 37, no. 2: 1535-1541

Yli-Olly, P. & Virtanen, I. (1989) "On the long term stability and cross-country invariance of financial ratio patterns", *European Journal of Operational Research,* vol. 39: 40-53

Youn, H. & Gu, Z. (2010) "Predict US restaurant firm failures: The artificial neural network model versus logistic regression model", *Tourism and Hospitality Research*, vol. 10, no. 3: 171-187

Yu, L., Wang, S. & Lai, K.K. (2008) "Credit risk assessment with a multistage neural network ensemble learning approach", *Expert Systems with Applications*, vol. 34: 1434-1444

Yu, L., Yue, W., Wang. S. & Lai, K.K. (2010) "Support vector machine based multiagent ensemble learning for credit risk evaluation", *Expert Systems with Applications*, vol. 37, no. 2: 1351-1360