



The 7th International Conference on Ambient Systems, Networks and Technologies
(ANT 2016)

Hybrid Discriminant Neural Networks for bankruptcy prediction and risk scoring

Fatima Zahra Azayite^a, Said Achchab^{b*}

^{a,b} National School for Computer Science and Systems analysis, Mohammed V University of Rabat, Morocco

Abstract

Determining the firm risk failure using financial statements has been one of the most interesting subjects for investors and decision makers. The discriminant variables that can be selected to predict firm health influence significantly the accuracy of the models especially if we have a missing data available. We developed a hybrid model of neural networks to study the risk of failure of Moroccan firms taking into account the data availability and reliability. Based on a three-step analysis, this methodology combines discriminant analysis, multilayer neural network and self-organizing-maps. This hybrid model considers the firms' behavior during three years to predict risk failure. It is qualified as a dynamic model because it adapts to financial environment and data availability. The model outperforms Discriminant analysis and gives a visual monitoring tool to supervise a firms' portfolio.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Conference Program Chairs

Keywords: Neural networks; Bankruptcy prediction; Self Organizing Maps; risk scoring

1. Introduction

Bankruptcy prediction is one of the most challenging subjects in economic and financial area. It is a crucial key indicator for making decision in many cases especially in investment and credit. Making a monitoring tool to evaluate failure risk of a firms' portfolio depending on their financial behavior can improve making decision. To predict financial distress, various models have been implemented based on mathematical, statistical or intelligent techniques.

* Corresponding author. Tel.: +212-053-777-8579; fax: +212-053-777-7230.

E-mail address: fatimazahra.azayite@um5s.net.ma, s.achchab@um5s.net.ma

Traditional methods were statistical techniques but in last years, new methods are employed such as machine learning analysis techniques (Decision trees, SVM or Neural Networks...). One of the most popular and performer tools are Artificial Neural Network model^{3,7} because they are able to learn nonlinear mappings between inputs and outputs. New studies focuses on soft computing techniques from the hybridization of techniques mentioned above to combine the advantages of individual models. A technique is called hybrid if several soft computing approaches were applied in analysis and only one predictor was used to make the final prediction¹². The accuracy of a hybrid model can be better than individual models used separately^{1,4}.

Bankruptcy prediction models discriminate between failing and healthy firms based on a set of variables. In one hand, there is no universally agreed ratios list to use in context and many researches are focused on listing the appropriate financial ratios. For instance, Altman⁵ used 5 ratios, Liang & Wu² used 7 ratios, Fedorova and Gilenko¹⁰ used 35 ratios. In this subject, F. Du jardin⁶ has shown that a neural-network-based model for predicting bankruptcy performs significantly better when designed with appropriate variable selection techniques than when designed with methods commonly used in the financial literature. As there is no theoretical method defining the best input variables of a neural network model, the discriminant analysis gives a statistical support in selecting the most relevant subset input vector for the designed neural network model¹. In the other hand, traditional failure models analyze the prediction field at solely 1-year horizon. They suppose that failure process is the same for all firms and the warning indicators occur in the same way. They do not pay attention to how different financial variables are connected in different phases of firm failure. However, in reality firms follow different strategies of declining¹³, some firms go bankrupt quickly even they appear healthy; others still survive even if their indicators are alarming.

In this study, the main goal is to increase the failure prediction accuracies of the existing approaches by analyzing companies' financial behavior through a period before failing using a dynamic layer for selecting appropriate financial variables depending on data availability in different failure phases. Because the majority of Moroccan firms have small size and they don't present complete balance sheet, it become very difficult to collect reliable data and calculate financial ratios over a long term consecutive years. This is the reason why we analyze the firms' behavior over 3 years by using the raw data collected from the balance sheet. We develop in this paper, a hybrid discriminant neural networks using Discriminant Analysis (DA), Back propagation Neural Networks (BPNN) and self-organizing maps (SOM) to evaluate the financial performance of Moroccan companies and to predict firms' failure especially small ones. The rest of paper is organized as follows: Section 2 presents the research methodology, describes the data pre-processing and describes the main characteristics of our hybrid model approach .Section 3 presents the empirical findings , Finally in Section 4 , we draw some conclusions and propose some further improvements.

2. Research methodology

2.1. Selecting a Data pre-processing

The database used in this research contains a three years period before failure of annual financial statements data for a sample of Moroccan firms. Taking into account that changes in macroeconomic environment can influence to the model estimated from one period to another, different samples are collected over the period from 2012 and to 2014 (Fig. 1). Three samples are collected and each one contains firms that their status (Failed vs. Healthy) is identified in one year and their financial declarations are available for the period of three years before. Furthermore, to reduce the influence of the industry, only three sectors are selected in each sample: retail, services and manufacturing. The firms are balanced 50% healthy and 50% failed by sample and by sector. Data collected contains 933 companies (Table 1). A binary variable is created with two values (0 if it is failed and 1 if it is healthy). Before using Data as input for Neural Network model, a normalizing function² (1) was applied to bound data values to -1 and +1 with X is the input matrix, Y is the normalized matrix, x_{min} and x_{max} are respectively the maximal and the minimum values of a variable.

$$Y = \frac{0.9-0.1}{x_{max} - x_{min}} X + \left(0,9 - \frac{0.9-0.1}{x_{max}-x_{min}} x_{max}\right) \tag{1}$$

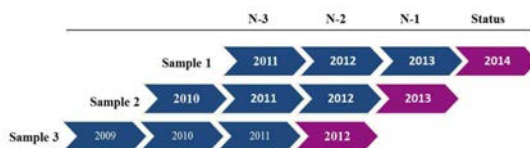


Fig. 1: Data collecting process
The purple represents the year (N) firms' failure status is identified and the blue represents the 3 years period (N-1, N-2 and N-3) financial data is collected.

Table 1 : Number of firms selected in the data collecting process.

	Failed	Healthy	All
Sample 1	195	195	390
Sample 2	155	151	306
Sample 3	120	117	237
All	470	463	933

2.2. Hybrid Discriminant neural network approach (HDNN)

In this study, our main objective is to predict failure based on a sample of firms that their status is already known. Furthermore, we want to score the risk of failure depending to the financial history behavior. To meet these objectives, we develop a Hybrid Discriminant Neural Networks model (Fig.2) combining DA, BPNN and SOM This model is qualified as a dynamic model for predicting failure and risk scoring, because it adapts to financial environment and data availability.

1) Financial ratios selection model

The first step in HDNN is to select the appropriate financial variables that predict the firms' bankruptcy. This model is necessary especially when we have many variables as input. To meet this objective we design a model based on DA with the stepwise discriminant approach. This method is used in statistics, pattern recognition and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events and, more commonly, for dimensionality reduction before later classification.

Each year before the failure has its own relevant variables that have the best abilities to predict the risk failure at a time horizon before bankruptcy (1, 2 or 3 years before). Therefore, this model selects three subsets of appropriate variables for each year of the period study as shown in Fig.3.

2) Predict failure model

The second step has the goal to predict the probability of failure for period of three years starting from a database of failed and healthy firms. The methodology chosen is a parallel run of three BPNN with one hidden layer (Fig.3).

As a reminder, Artificial Neural Network (ANN) is a collection of interconnected neurons that incrementally learn from their environment (data) to capture essential linear and nonlinear trends in a complex data to provide reliable predictions for new situations⁷. When it is trained to assign the correct target classes to a set of input patterns, it can be used to classify new patterns. The topology of ANN plays a fundamental role in its functionality and performance. In literature, ANN with one hidden layer is the best structure to use for classification problems¹¹. The commonly used type of neurons connection is feed-forward. It means that information moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes with no cycles or loops in the network. In literature⁹, feed-forward network with a one hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of R^n . The back-propagation algorithm is one of the most applied methods to feed-forward networks.

This model contains three BPNN (Fig. 3) modeling failure predictions respectively for 1, 2 and 3 years before bankruptcy. Each network has its own architecture: The model input is respectively the preselected subsets in the first step corresponding to 1, 2 and 3 years horizon. The number of hidden neurons for each BPNN is chosen by experience. Tests are made to choose the number of neurons that gives the minimum of Mean Squared Error (MSE). The output is one binary variable that separates default and healthy firms (0 for default and 1 for healthy firms). The activation function is a sigmoid function and the algorithm of training is a gradient descendant. The performance of this model is compared with DA as a statistic model. This model, implemented with adapted variables, is also compared with the same model implemented with common ratios to evaluate the importance of the first layer.

3) Failure risk monitoring

Based on unsupervised learning method, this layer creates a visual representation of firms clustering depending on their risk failure behavior during three years. One of the most popular neural network models is SOM. With its ability to project multidimensional data onto a less dimensional data, is an interesting model to define an intermediate risk classes between failed and healthy targets. Based only on input vectors with a competitive learning, without specifying the desired output in similarity with BPNN, the SOM can determine clusters in one or two-dimensional relationships in data without using an externally provided target output⁷. The SOM has two layers of neurons: Input layer that represents the input variables and Output layer that is composed on neurons arranged in a single line (one-dimensional) or two-dimensional. Each output neuron is connected to all the source nodes in the input layer.

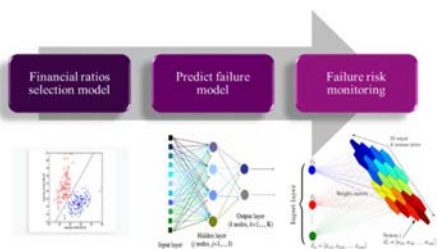


Fig.2 Hybrid Discriminant neural network approach

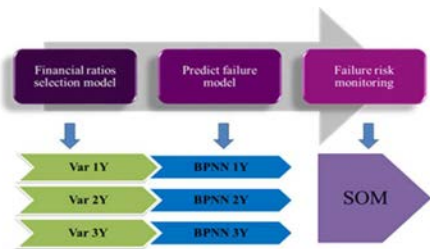


Fig.3 Hybrid Discriminant neural network process

The first layer of the network has three inputs representing the firms’ probability of failure during three years. The output layer is a one dimensional network 1x5 map to define intermediate risk failure levels (1: Healthy, 2: Probably healthy, 3: Moderate failure risk, 4: High risk failure, 5: Very high risk failure).

3. Empirical study

1) Financial ratios selection model

Taking into account that there is no universally agreed ratios list used for financial prediction models and that neural-network-based model for predicting bankruptcy performs better when designed with adapted variable selection techniques⁶, this model is designed to define appropriate variables depending on data availability.

Because the majority of firms don’t present complete balance sheet, ratios are not available for the entire database. For this constraint, we use as input raw data collected from financial declarations based on Moroccan declaration model. Rather than designing one variable selection model, we designed 3 models according to 3 time horizons study (1, 2 and 3 years before failure).

In this step, DA is used to define the appropriate variables that predict failure depending on the time horizon. It is also used in this article to compare its classification capability as a traditional classification tool with the HDNN proposed. For each time horizon, a subset of variables is selected using the stepwise discriminant approach. 17, 14 and 18 significant predictor variables are selected respectively for 1, 2 and 3 years before failure models. The results, presented in Table 2, revealed that the average of correct classification rate through 3 years is 70.3% and that decrease as the time horizon increases. The model output is the subsets selected as appropriate variables to predict firms’ failure in the three time horizons.

2) Predict failure model

As explained before, the predict failure model is composed by three BPNN associated to predicting failure 1, 2 and 3 years before and the vector of behavior during this three years will be evaluated to score risk failure.

To fix the architecture of the three BPNN, several test experiences were done, to choose the number of hidden nodes and the learning rate that minimize the Mean Squared Error and give the best accuracy rate. Each year, 30 BPNN structures were tested having a number of hidden nodes from 11 to 40 and trained with several learning rates values (0.002, 0.004 and 0.006). The convergence criteria used for training are a mean squared error (MSE) less or equal to 0.00001 or a maximum iterations equal to 3000. The BPNN topology with minimum MSE is considered as the optimal one. As results, the topologies chosen are 11, 13 and 12 hidden nodes for respectively BPNN 1Y, BPNN 2Y, BPNN 3Y with 0.004 learning rate. After that, a 4-fold cross validation technique is used to train and test the model to avoid over fitting. The data set is randomly split into 4 equal size subsets. In each time, data is trained on 3 folds and tested on the remaining fold. The accuracy of the model is the average of the 4 individual accuracy measures. As shown in Table 3 this model gives a higher accuracy than DA.

In order to test the hypothesis that a model trained with appropriate variables gives better results than with commonly used variables specially with missing data, we do the same process to build predict failure model with a ratios commonly used in literature. We choose 5 financial ratios (Working capital/Total assets, Retained Earnings/Total assets, Earnings before interest and taxes/Total assets, Market value equity/Book value of total debt, Sales/Total assets). To define the 3 topologies of BPNN, the same test experiences were done and the results summarized in Table 3 (c) confirm the hypothesis.

3) Failure risk monitoring

After checking the validity of the Predict failure model, we collect the probability of failure computed before to analyze firms’ behavior before failure and simulate risk scoring.

Combining the output of the BPNN models, a one dimensional 1x5 map is created (Fig. 4). We analyze the five behavior groups created (G1, G2, G3, G4 and G5) based on the distribution of failed and healthy firms (fig. 5) in each group. We confirm that the map created gives a risk failure scale from 1 to 5 as presented in Fig. 6 legend. Group 1, with 275 healthy firms, gives 95% as healthy accuracy in the same group and it can be labeled as a healthy group.

Table 2: Comparative Classification results between Predict failure model (b), DA (a) and specific ratios model (c)

Predicted class	Actual Class																	
	Classification results using Discriminant analysis (a)						Classification results using Predict failure model (b)						Classification results using specific ratios (c)					
	N-3		N-2		N-1		N-3		N-2		N-1		N-3		N-2		N-1	
Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	Failed	Healthy	
Failed	442(94,0%)	243(52,5%)	444(94,5%)	250(54,0%)	442(94,0%)	234(50,5%)	351(74,4%)	47(10,1%)	385(82,0%)	72(15,6%)	395(83,9%)	57(12,4%)	375(80,1%)	95(20,4%)	392(80,8%)	78(17,4%)	386(82,1%)	84(18,1%)
Healthy	28(6%)	220(47,5%)	26(5,5%)	218(46,0%)	28(6,0%)	229(49,5%)	119(25,3%)	416(89,9%)	85(18,0%)	391(84,4%)	75(16,1%)	406(87,6%)	93(19,9%)	370(79,6%)	93(19,2%)	370(82,6%)	84(17,9%)	379(81,9%)
AVG correct classification	70,8%		70,2%		71,8%		81,4%		82,3%		84,5%		79,8%		81,7%		82,0%	

Group 5, with 360 failed firms, represents 90% of failure prediction accuracy in the same group and 76.5% in the total failed firms and it can be designed as a very high-risk failure group. Intermediate scales are groups 2, 3 and 4. The first two groups represent respectively 86%, 64% of healthy firms whereas the last group displays 63% of failed ones. Based on these groups definition, we plot the average behavior of firms in Fig. 8. This Chart presents a prototype of firms' behavior during three years to predict the risk failure and it confirm the hypothesis that some firms go bankrupt quickly even they appear healthy (G4); others still survive even if their indicators are alarming (G2 and G3).



Fig. 4: SOM result

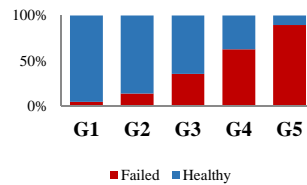


Fig. 5 Distribution of groups based on firms status (Failed vs. Healthy)

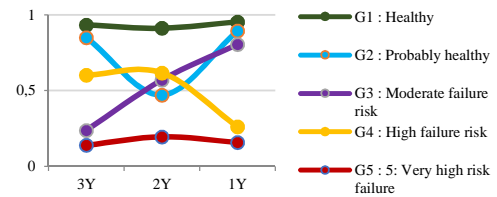


Fig. 6 Firms' failure behavior process.

The X-axis represents the years before failure and Y-axis represents the degree of failure (0 failure, 1 healthy)

4. Conclusion

In this study, we have designed a hybrid discriminant neural network based on Discriminant analysis, backpropagation neural networks and self-organizing maps to predict risk failure. This model takes into account the way firms move in failure space through a period of three years and the constraints of missing data to define their risk failure. Based on results, the HDNN gives a good accuracy in comparison with DA. Moreover, it confirms the hypothesis that a model trained with appropriate variables gives better results than with commonly used variables specially with missing data. This hybrid model can be a useful tool for investors and stakeholders to define the risk profile of a firms' portfolio. Nevertheless, it is necessary to perform the model by introducing a dynamic layer to define the number of hidden neurons in the BPNN models.

References

1. T.-S. Lee, C.-C. Chiu, C.-J. Lu et I.-F. Chen, «Credit scoring using the hybrid neural discriminant technique,» *Expert Systems with Applications*, p. 245–254, 2002.
2. L. Liang et D. Wu, «An application of pattern recognition on scoring chinese corporations financial conditions based on backpropagation neural network,» *Computers & Operations Research*, pp. 1115-1129, 2005.
3. S. X. Zongyuan Zhao, «Investigation and improvement of multi-layer perceptron neural,» *Expert Systems with Applications*, 2015.
4. F. J. López-Iturriaga, «Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks,» *Expert Systems with Applications*, 2015.
5. E. I. Altman, «Financial ratios, discriminant analysis and the prediction of corporate bankruptcy,» *Journal of Finance*, 1968.
6. P. d. Jardin, «Predicting bankruptcy using neural networks and other classification methods: the influence of variable selection techniques on model accuracy,» *Neurocomputing*, 2010.
7. S. Samarasinghe, *neural networks for applied sciences and engineering*, Auebach Publications, 2006.
8. C.-F. Tsai et J.-W. Wu, «Using neural network ensembles for bankruptcy prediction and credit scoring,» *Expert Systems with Applications*, pp. 2639-2649, 2008.
9. G. Cybenko, «Approximation by superpositions of a sigmoidal function,» *Mathematics of Control, Signals and Systems*, 1989.
10. E. Fedorova, E. Gilenko et S. Dovzhenko, «Bankruptcy prediction fo Russian companies : Application of combined classifiers,» *Expert Systems with Applications*, pp. 7285-7293, 2013.
11. J. H. Min et Y.-C. Lee, «Bankruptcy prediction using support vector machine,» *Expert Systems with Applications*, p. 603–614, 2005.
12. J. Sun et al., «Prediction financial distress and corporate failure : A review from the state-of-art definitions, modeling, sampling, and featuring approaches,» *Knowledge-Base Systems*, pp. 41-56, 2014.
13. E. K. Laitinen, O. Lukason et A. Suvas, «Are firm failure processes different? Evidence from seven countries,» *Investment Management and Financial Innovations*, 2014.