

A Hybrid SOM-Altman Model for Bankruptcy Prediction

Egidijus Merkevičius, Gintautas Garšva, and Stasys Girdzijauskas

Department of Informatics, Kaunas Faculty of Humanities, Vilnius University
Muitinės st. 8, LT- 44280 Kaunas, Lithuania
{egidijus.merkevičius, gintautas.garsva,
stasys.girdzijauskas}@vukhf.lt

Abstract. This article analyzes the problems of business bankruptcy, and the methods for bankruptcy prediction. This study proposed to join two models, one is the multi-discriminate Z-Score created by Altman, and the other is the Self-organizing maps. We proposed to generate self-organizing maps based on the financial data of public companies that are included in the NASDAQ list. These maps were used for bankruptcy prediction as well as creating classification of financial risk for Lithuanian companies. Comparing the weak results of prediction we accelerated by changing of ratios weights of the Altman Z-Score model. In this way, it can fit to conditions of the Lithuanian conjuncture. Based on the original ratio weights in Altman's Z-Score the results predicting Lithuanian bankruptcy were weak. The weights of Altman's Z-Score model were changed to fit the Lithuanian economic circumstance.

Keywords: self-organizing maps, Z-Score, bankruptcy, prediction, bankruptcy class, multivariate discriminate model, Altman.

1 Introduction

The forecasting of bankruptcy has always been a relevant task in the finance markets. Available algorithms of statistical and artificial intelligence and the combination of these methods provide more accurate and predictable results [8], [1], [7], [5]. The early history of research attempts to classify and predict bankruptcy is well documented in [4]. The historical development of statistical bankruptcy models can be divided into 3 stages: 1) univariate analysis (by Beaver in 1966); 2) multivariate (or multiple discriminate [MDA]) analysis, and 3) Logit analysis (initiated by Ohlson).

The purpose of this paper is to propose a hybrid artificial-discriminate model to be used as a predictive measure of corporate financial health (0-healthy, 1-bankrupt), based in an unsupervised artificial neural network and a multivariate discriminate model by Altman. Altman's Z-Score model was created for companies that are best characterized in a perfect market economy as evidenced by his use of USA companies financial statements. In light of this mathematical basis a second purpose to this paper has been to present a methodology for adapting Altman's Z-Score model based in the economic reality of developing countries; specifically to propose changing the weight measures in Altman's Z-Score variables.

Therefore, the focus is to explore the capabilities of an unsupervised learning type of artificial neural network – self-organizing map (SOM) to solve such problems as bankruptcy and financial condition. Secondly it is to describe other related work through SOM in predicting bankruptcy and financial distress. In the third part a methodology is presented using a hybrid SOM-Altman model to bankruptcy prediction, and fourthly, the results of this study are demonstrated using the proposed model. In the last section the main conclusions are presented and discussed.

2 Related Work

Artificial neural networks (ANN) are divided into supervised and unsupervised learning [8]. When working in the realm of prediction supervised ANN are normally used. The aim of this investigation is to observe the capabilities of an unsupervised ANN to predict bankruptcy classes of specifically Lithuanian companies.

The Self-organizing map (SOM) is an unsupervised learning artificial neural network that is generated without defining output values. The outcome of this process is a two-dimensional cluster map that can visually demonstrate the financial units which are scattered according to similar characteristics. In this case the bankruptcy class of data is labeled on the map and the data distribution is analyzed. A detailed description of the SOM method is presented in [11].

During the past 15 years investigations in area of SOM applications to financial analysis have been done. Doebeck described and analyzed most cases in [6]. Martindel-Prio and Serrano-Cinca were one of the first to apply SOM in financial analysis. They generated SOM's of Spanish banks and subdivided those banks into two large groups, the configuration of banks allowed establishing root causes of the banking crisis [2].

Kiviluoto [10] made a map by means of including 1137 companies, out of which 304 companies were crashed. SOM's are said to give useful qualitative information for establishing similar input vectors. Based on Kiviluoto's study, through visual exploration one can see the distribution of important indicators (i.e. bankruptcy) on the map.

The previous authors work is based in an historical or current analysis of company and market conditions. Through this work they have been able to take past settings and predict forwards in time the outcomes of bankruptcy or crisis periods in market economy. It is proposed here that by generating SOM's to current information future segmentation of credit classes can be discerned for new or existing companies.

3 Methodology

In this section the SOM and Altman's Z-Score model are introduced as well as a hybrid SOM-Altman model. The hybrid model was created on the basis of the SOM and Altman's Z-Score with an applied new methodology to change the weights of Altman's Z-Score variables under a specific dataset test.

In the self-organizing process the output data are configured in a visualization of the topologic original data. The unsupervised learning of the SOM is based on competitive learning ("winner takes all"). A detailed description of the SOM algorithm is presented in [6], [11], [5].

Altman's Z-Score predicts whether or not a company is likely to enter into bankruptcy within one or two years. Edward Altman developed the "Altman Z-Score" by examining 85 manufacturing companies in the year 1968 [3]. Later, additional "Z-Scores" were developed for private manufacturing companies (Z-Score - Model A) and another for general/service firms (Z-Score - Model B) [4].

According to E. Altman the Z-Score bankruptcy-predictor combines several of the most significant variables in a statistically derived combination. It was originally developed on a sampling of manufacturing firms. The algorithm has been consistently reported to have a 95 % accuracy of prediction of bankruptcy up to two years prior to failure on non-manufacturing firm. Z-Score for private firms is as follows [4]:

$$Z = 0.717(X1) + 0.847(X2) + 3.107(X3) + 0.420(X4) + 0.998(X5) \quad (1)$$

where

X1 = Working capital/Total assets (captures short-term liquidity risk),

X2 = Retained earnings/Total assets (captures accumulated profitability and relative age of a firm),

X3 = Earnings before interest and taxes/Total assets (measures current profitability),

X4 = Book value of Equity/Book value of total liabilities (a debt/equity ratio captures long-term solvency risk),

X5 = Net sales/Total assets (indicates the ability of a firm to use assets to generate sales),

and

Z = Overall index.

In the original model a healthy private company has a $Z > 3$; it is non-bankrupt if $2.7 < Z < 2.99$; it is in the watch-listed zone if $1.8 < Z < 2.69$; it is unhealthy (bankrupt) if it has a $Z < 1.79$. This paper has corrected the bankruptcy classes where a healthy private company has $Z > 1.8$ and bankrupt company has a Z score of < 1.8 , (e.g. we eliminated the "gray" zone).

In figure 1 is presented the algorithm of the proposed hybrid methodology for bankruptcy class prediction.

The main steps are as follows:

- 1 On the basis of the NASDAQ list companies financial statements, the Altman's Z-Score variables are calculated and converted to bankruptcy classes (0-healthy, 1-bankrupt).
- 2 Data preprocessing is executed. It consists of a normalized data set, a select map structure, a topology, a set of other options like data filter, and a set of delay etc.
- 3 The SOM is generated. The Inputs of SOM are the Altman's Z-Score variables and the labels are bankruptcy classes.
- 4 The SOM is labeled with the bankruptcy classes.
- 5 On the basis of the TEST list companies financial statements the Altman's Z-Score variables are calculated. Bankruptcy classes (0-healthy, 1-bankrupt) are assigned. Companies that are included in the TEST lists will be used in this study and we will be predicting their bankruptcy class and their weights Altman model weights. At this step we calculate Altman's Z-Score variables on the original weight basis.
- 6 The generated SOM is labeled with the bankruptcy classes of TEST companies.

- 7 Labeled units of the trained SOM are compared with the same units labeled with TEST bankruptcy classes.
- 8 Corresponded units are calculated.
- 9 The second part of the algorithm is created in order to increase the number of corresponding TRAIN and TEST checked labels which are located on the same SOM map unit number.

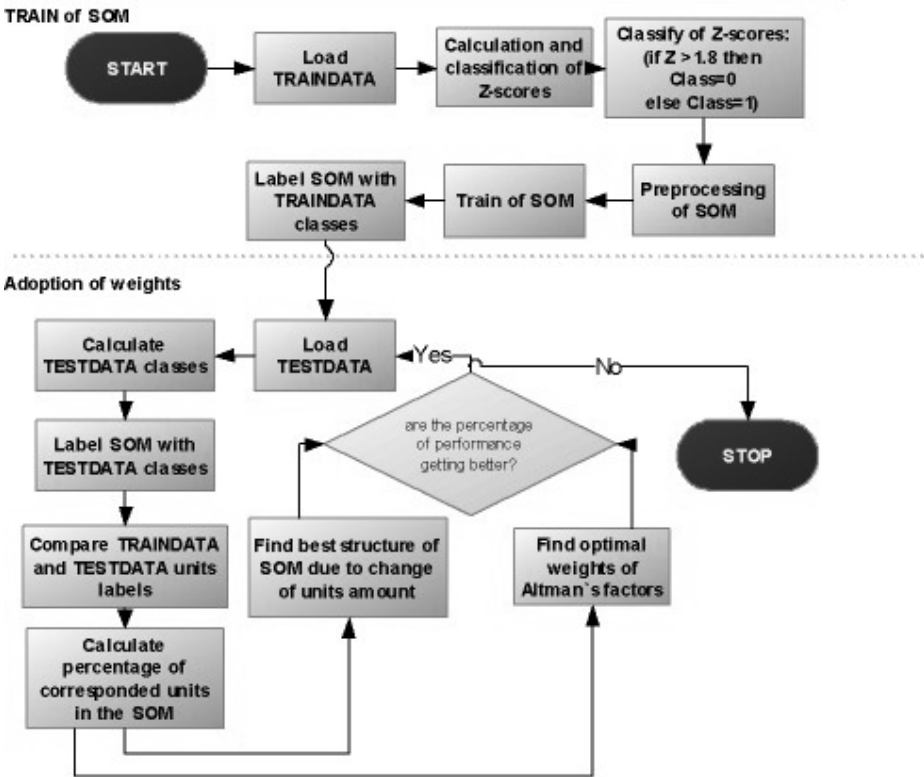


Fig. 1. Algorithm of proposed hybrid model and methodology of weights adoption

- 10 The attempt is made to create such a map structure within which the amount of unit numbers has the biggest corresponding label number.
- 11 The assessment of the influence of each of Altman's model variables to the number of corresponding labels.
- 12 When the performance of the prediction doesn't change the algorithm is stopped.

The results of this algorithm can be presented as follows:

1. A new SOM with a concrete prediction percentage;
2. A new multivariate discriminator model that is based on Altman's Z score but with corrected weight variables.

4 Results of Experiments

In this paper the possibilities of the use of SOM's have been studied using two real financial datasets: companies from NASDAQ list, (or TRAINDATA) loaded from EDGAR PRO Online database, and a dataset of Lithuanian company financial statements (TESTDATA) presented by one of the Lithuanian banks.

The basis for generating the SOM is TRAINDATA. The calculated bankruptcy ratios are used as inputs and the Z-Scores from Altman's Z-Score model are used as labels for the identification of units in the SOM.

Table 1. Characteristics of financial datasets

Dataset	TRAINDATA	TESTDATA
Taken from	EDGAR PRO Online Database (free trial)	Database of Lithuanian bank.
Period of financial data	2004Y	2004Y
Count of records	1108	742
Number of inputs (attributes)	5	
Risk class of bankruptcy	0-1 (>1.8 is healthy - <=1.8 is bankrupt)	

The SOM was trained using the SOM Toolbox for Matlab package [9]. From the U-matrix in Figure 2 the SOM formed four clusters. By looking at the labels, it can be seen that the two clusters in the left corresponds to the bankrupt class. The two other clusters in the right part of U-matrix correspond to the healthy class.

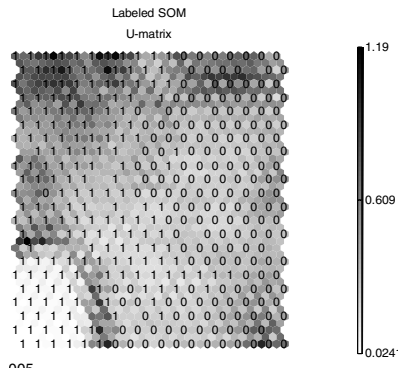


Fig. 2. U-matrix of SOM with TRAINDATA labels

The next step is to label the map with the TESTDATA labels and calculate the corresponding units between the TRAINDATA and the TESTDATA labels. It is important to note that only the units that were not empty in both TRAIN and TEST cases are being compared. The ratio between corresponding units and the number of

all TESTDATA labels in the SOM map defines accurateness of bankruptcy prediction which is equal to 72.678%. The U-matrix with TESTDATA labels presents at Figure 3(a).

In order to increase the accuracy of bankruptcy prediction the cycle of SOM structure change is created. By changing SOM size, the following accuracy of bankruptcy prediction results is acquired:

Table 2. Performance of bankruptcy prediction via change of SOM size

Bankruptcy prediction (x100%)	Map size		
0,69767	15	x	13
0,69811	20	x	15
0,72678	22	x	18
0,68681	25	x	20
0,68557	27	x	22
0,68182	30	x	23
0,68817	32	x	25
0,65761	33	x	27
0,66667	36	x	28
0,66486	38	x	29

From the results we can see that the bigger the size of the map, the lower the accuracy of bankruptcy prediction. The best results are acquired with the map size of 22x18.

The next investigation was concerned with the influence of each of Altman’s model variables to the performance of bankruptcy prediction. In the process of the cycle the weight of each variable is being changed and the change in accuracy in bankruptcy prediction is being monitored. It was noticed that during the change of label weights the accuracy of prediction achieved its highest score and it gradually dropped afterwards. The results show that the most important influence in the performance of bankruptcy prediction is the Net Sales/Total assets variable, and secondly is the EBIT/Total assets variable. The performance result after the correction of the variable weights increased from 72.678% up to 92.352% as follows:

Table 3. Comparison of performance results via change variable weights before and after

Name	Variable	Weight before	Weight after
Earnings before interest and taxes/Total assets	X3	3,107	2,800
Net sales/Total assets	X5	0,998	0,400
Book value of Equity/Book value of total liabilities	X4	0,420	0,440
Working capital/Total assets	X1	0,717	0,717
Retained earnings/Total assets	X2	0,847	0,843
Performance of bankruptcy prediction (%)		72,678%	92,352%

Other variables have less influence between American and Lithuanian companies. The variables of profitability and turnover are the most important factors which have the largest weight to predict bankruptcy.

U-matrix with TESTDATA labels presented as follows (Figure 3 (b)) and can be visually compared with U-matrix before corrections of variable weights (Figure 3(a)). It can be seen that bankrupt and healthy class labels distributed comparatively to the labels in TRAINDATA (Figure 2).

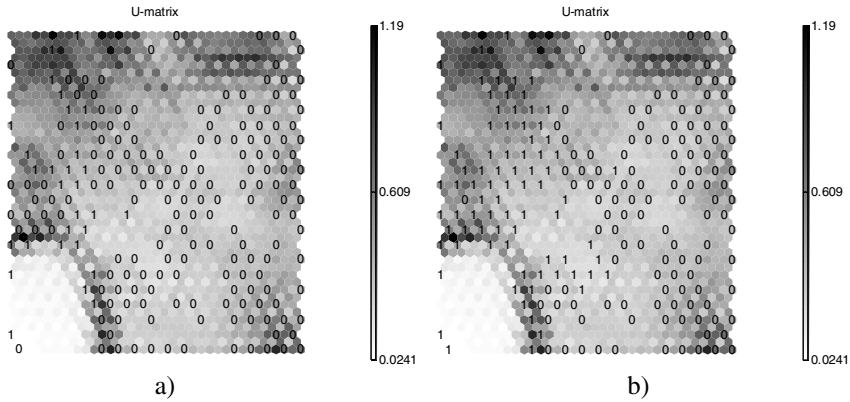


Fig. 3. U-matrix of trained SOM with TESTDATA labels before and after correction of weights

5 Conclusions

Our experiments and the results present several conclusions:

- The presented methodology works well with real world data, hybrid SOM-Altman’s bankruptcy model with presented datasets predicted with 92.352% performance.
- Methodology of presented hybrid bankruptcy model is flexible to adopt every datasets because rules and steps of methodology algorithm are universal.
- In this paper it was shown, that hybrid SOM-Altman’s adopted bankruptcy model can present differences between financial statements conjuncture of two types of economic countries: west market economy and development economy.

References

1. A. Atiya. Bankruptcy prediction for credit risk using neural networks: a survey and new results. *IEEE Transactions on Neural Networks*, Vol. 12, No. 4, pp. 929-935, July 2001.
2. B. Martin-del-Prio, K. Serrano-Cinca, Self-Organizing Neural Network: The Financial State of Spanish Companies. In *Neural Networks in Finance and Investing. Using Artificial Intelligence to Improve Real-World Performance*. R.Trippi, E.Turban, Eds. Probus Publishing, 1993., 341-357

3. E. Altman. Financial Ratios, Discrimination Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23, September, 1968.
4. E. Altman. Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models. (working paper at <http://pages.stern.nyu.edu/~ealtman/Zscores.pdf>) 2000
5. E. Merkevičius, R. Simutis, G.Garšva. Forecasting of credit classes with the Self-organizing maps. *Information technology and control*. 2004, 4(33), ISSN 1392-124X. 61-66 pp.
6. G. Deboeck. Financial Applications of Self-Organizing Maps. *American Heuristics Electronic Newsletter*, Jan, 1998.
7. G. Deboeck. Self-Organizing Maps Facilitate Knowledge Discovery In Finance. *Financial Engineering News*, December 1998.
8. J. Galindo, P. Tamayo. Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications. *Computational Economics*. April 2000, Volume 15.
9. J. Vesanto, J. Himberg, E. Alhoniemi, J. Parhankangas, SOM toolbox for Matlab 5, Technical report A57 (2000), Helsinki University of Technology, Finland.
10. K. Kiviluoto. Predicting bankruptcies with the self-organizing map. *Neurocomputing*, 21:191–201, 1998.
11. T. Kohonen. The Self-Organizing Map. *Proceedings of the IEEE*, 78:1464-1480.