CONNECTING CORPORATE GOVERNANCE TO COMPANIES' PERFORMANCE BY ARTIFICIAL NEURAL NETWORKS

Darie MOLDOVAN, PhD*
Mircea RUSU, PhD student**

Abstract

The objective of this paper is to demonstrate the utility of Artificial Neural Networks in connecting the corporate governance variables, among financial information, to companies' performance. We have considered two well known indicators for estimating companies' performance (Tobin's Q ratio and Altman Z-Score) and we used them as target variables for classification using the neural networks. The results proved to be robust after experimenting with three different datasets, containing information on 1400 companies from three stock indexes: S&P 500, STOXX Europe 600 and STOXX Eastern Europe 300.

Keywords: Artificial Neural Networks, classification, corporate governance, Altman Z-Score, Tobin's Q ratio

JEL Classification: C63, C90, G32, G33

1. Introduction

The prediction of companies' performance is a subject of great interest in the stock markets investment decision process. Different quantitative methods were adopted during the years in order to fundament the decisions of buying or selling a certain stock, including the fundamental analysis of the underlying company and technical analysis of the price evolution.

When analyzing financial data for a certain company, we obtain an image of the past, trying to extrapolate it in the future, to identify trends and flags. But how will the performance of the company be affected by a change in the corporate governance policies or practices? In the fast changing world of investments one cannot wait for the next financial reports in order to answer this question, but needs models to evaluate the impact of such changes.

---

* *Lecturer, Babeş-Bolyai University of Cluj-Napoca, Business Information Systems Department.*

** *Technical University of Cluj-Napoca, Computer Science Department.*

In our paper we develop a model using Artificial Neural Networks (ANN), trying to evaluate the relationships between the corporate governance variables, the financial data and some wide known indicators like the Tobin's Q (Tobin, 1969)(estimating the performance) and Altman Z-score (Altman, 1968)(estimating the bankruptcy risk).

We use a dataset containing 1400 listed companies, divided according to the three indexes to which they belong: S&P500, STOXX Europe 600 and STOXX Eastern Europe 300, covering in this way three different regions of the world. We aim to identify the most important corporate governance variables for each data subset and to assess the differences.

The remainder of the paper is structured as follows. Section 2 is dedicated to the literature review; section 3 presents the artificial neural network methodology for learning the relationship between the variables; in section 4 we present our findings following the experiments. We present our final conclusions in section 5.

2. Literature review

The use of Artificial Neural Networks was adopted in many fields to solve complex problems, being recognized for its prediction power and robustness. Compared to other data mining methods, such as decision trees and regressions, the ANN perform better in terms of accuracy, but due to their "closed-box" nature are much more difficult to explain.

The literature in this field shows us the interest in identifying different modelling methods in the finance and corporate governance fields. The related work can be categorized into two directions. One aims to enhance the fraud detection methods by using machine learning models, while the other is focused on assessing risk or performance. The general opinion in both categories is that by adding corporate governance variables to the models, their accuracy improves.

Fanning (1998) built a model for detecting fraud made by the management by altering the financial results and provides empirical evidence for a selection of flags related to frauds. Turnbull (2002) uses a model ground corporate governance in the field of cybernetics. Its main findings sustain the idea that a unitary board is not reliable to govern complex companies.

Kumar (2006) uses ANN vs linear models in order to predict credit risk, using the financial data available. It shows the results obtained by employing ANN (instead of linear prediction models) are superior in terms of both training and validation. They conclude that ANN are more suitable for large datasets.

Polsiri (2009) uses ANN and logit models to predict companies' distress on the Thai listed companies during the East Asian economic crisis. The study emphasises the importance of corporate governance variables in predicting companies' distress. The classification models global accuracy was high for both approaches, suggesting the results could be considered valid as signals for companies' distress.

The paper of Chiou (2010) proposes the approach of finding the relation between the corporate governance and the pricing of the initial public offerings by employing the ANN as a learning method. Their results show that using the corporate governance variables among the financial variables, the prediction accuracy increases, concluding that corporate governance is closely related to the price of the initial public offering.

Nor (2011) proposes a stock market trading strategy based on corporate governance variables. It investigates whether this approach can return economically significant results. ANN are used for learning the model and the results show superior returns than the classic benchmark buy and hold strategy. The study uses data from the Malaysian stock market and also refers to the information efficiency on this market.

The relation between ownership structure of a company and the dividend yield is learned by Soni (2011) using ANN. The papers shows evidence from the Indian stock market for a period of five years. The results show that individual investors presence in a large proportion in the shareholder structure is negatively impacting the dividend yield, while companies with banks and financial institutions as shareholders relate to a higher dividend. The study also shows that foreign investors in the shareholders do not have a significant influence on the dividend policy, suggesting their neutrality in this matter.

In a more recent study by Chen (2014) several data mining methods (decision trees, neural networks, random forest) are used in order to detect fraudulent financial statements, considering also corporate governance variables. The results show that introducing the corporate governance variables in the model the classification results improve. The study was concerned companies from Taiwan. Financial crises are also subject to data mining prediction. Li (2015) is using corporate governance variable as predictors along with financial data available. Their results are similar with those previously mentioned in terms of usefulness for the corporate governance variables, which provide better prediction results if used together with other independent variables.

Zhang (2015) proposes a methodology for creating sustainability reports for organizations, considering economic, environmental, social, and governance indicators. The paper presents ways for using the neural networks as classification method for the financial position of a company.

The relation between the corporate governance practices and voluntary financial information disclosure for the most important French companies is studied by Botti (2014).

3. Classification with Artificial Neural Networks

The ANNs are widely used to solve complex, nonlinear hypothesis. They can model shapes that are otherwise difficult to build using the other classification methods. Methods like linear regressions have limited learning abilities when dealing with many features, especially when we want to model the interactions between these features. If a regression would imply quadratic or cubic terms, the number of resulting feature would grow exponentially. The regression would become soon impractical because of increasing needs in terms of computation. Because of the complex form of the ANNs it is difficult to explain the models built, which are regarded as “black-boxes” by many. This is a result of the hidden layers, and the lack of transparency of the learning algorithms which computes the weights of the variables in the layers.

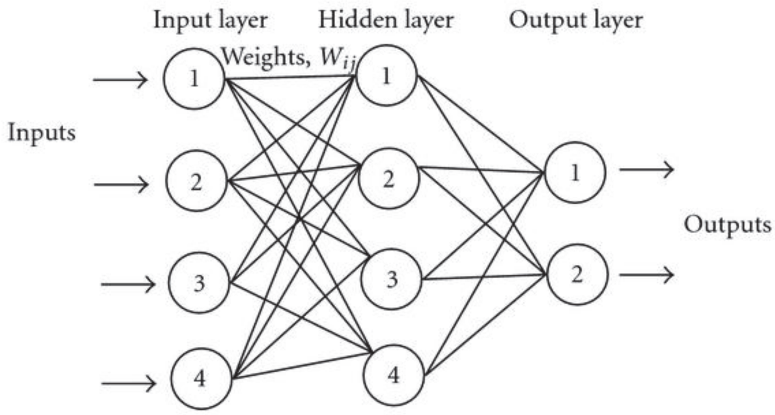
In Figure 1 we show an example of ANN with one hidden layer and two classes as outputs. In the input layer we have the initial values of the variables. Their influence is transmitted to the hidden layer by weighting each variable in order to obtain the best possible classification for the output layer.

The steps for training a neural network are the following:

1. Weights initialization (randomly);
2. Forward propagation;
3. Cost function definition;
4. Back propagation;
5. Optimization function for minimizing the cost function.

Figure 3

An ANN example



4. Experiments and results

In order to apply the neural networks methodology we had to prepare the datasets for learning. In the knowledge discovery process (Fayyad, 1996) this step represents the data pre processing and transformation phases.

Figure 4

KDD process (Fayyad, 1996)

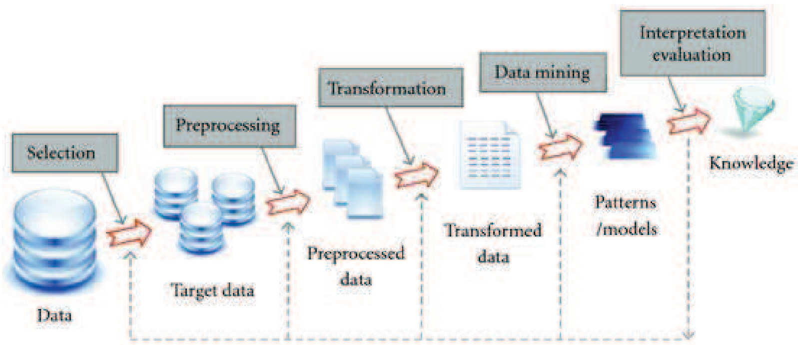


Figure 2 depicts the entire process, showing the steps needed in order to obtain the knowledge, but also the cyclic character of this type of research. After obtaining the results, one can return to a previous step in order to adjust some parameters, until obtaining meaningful and robust information.

Our datasets were extracted according to the composition of the three stock indexes chosen: S&P500 (SPX), STOXX Europe 600 (SXXP) and STOXX Eastern Europe 300 (EEBP). The total number of instances was 1400, and the initial number of variables was 52.

The target variables for our experiments are Tobin's Q and Altman Z-Score. Tobin's Q is used as a measure for the companies' performance, while Altman Z-Score is an indicator for the financial distress. The following two formulas (1 and 2) show the way to calculate each of the two indicators.

$$\text{Tobin's Q} = (\text{Market Cap} + \text{Total Liabilities} + \text{Preferred Equity} + \text{Minority Interest}) \,/\, \text{Total Assets} \tag{1}$$

$$\text{Altman's Z-Score} = 1.2 * (\text{Working Capital} \,/\, \text{Tangible Assets}) + 1.4 * (\text{Retained Earnings} \,/\, \text{Tangible Assets}) + 3.3 * (\text{EBIT} \,/\, \text{Tangible Assets}) + 0.6 * (\text{Market Value of Equity} \,/\, \text{Total Liabilities}) + (\text{Sales} \,/\, \text{Tangible Assets}) \tag{1}$$

As independent variables we have chosen 50 financial and corporate governance variables. A list of selected variables can be found on Annex 1. Our objective is to build a model that can incorporate the information regarding both financial and corporate governance in relation with the two dependent variables. In this way, one could easily verify if certain changes in the corporate governance policies or practices could affect the performance of the company and the default risk.

After conducting a descriptive analysis of the data we defined the necessary steps to take in order to clean the data and prepare it for the learning algorithm.

The datasets contained outliers and missing data that we had to treat. We eliminated the instances containing outliers and replaced the missing data with the average value of the specific variable. Also, in order to facilitate the computation, we normalized the values for each variable, using the feature scaling shown in the following formula:

$$X' = \frac{X - \mu}{X_{mx} - X_{min}} \tag{1}$$

In this way all the variables will be brought into a certain range. This is useful in order to speed up the learning, without affecting the accuracy of the classification.

The neural network contained one hidden layer and the maximum number of epochs was set to 500. The validation threshold was set to 20. Due to the limited number of observations, we preferred the 10-fold cross validation method for training and validation.

We can note the imbalance between the classes of target variables, which can lead to poor classification for the class with less observation. In order to tackle this aspect, we employed a method for oversampling: Synthetic Minority Oversample Technique (SMOTE), proposed by Chawla (2002).

The results after training the neural networks for the three datasets having the Altman Z-Score as target variable are presented in Table 1.

Table 2

Altman Z-score as class

Dataset	Correctly classified instances	Coverage of cases (0.95 level)	Precision Class 0	Precision Class 1	Precision Class 2	ROC Area
SPX	84.3938 %	90.9964 %	0.875	0.817	0.836	0.936
SXXP	76.456 %	84.8823 %	0.831	0.68	0.767	0.906
EEBP	77.2523 %	86.4865 %	0.876	0.706	0.721	0.882

We can note the overall accuracy of the classification is robust, the SPX index dataset showing the best overall performance. The EEBP dataset contained more missing data than the others, this aspect being a major drawback for the classification accuracy.

Table 2 shows the results obtained after training the neural networks for classifying the datasets having Tobin's Q ratio as target variable. The results are showing a classification slightly less robust than the previous case, where Altman Z-Sore was the target variable. The precision is consistent for the post classes, showing about the same values for each of them. Surprisingly the highest ROC area was obtained for the EEBP dataset, although the best overall classification was obtained for the SXXP dataset.

Table 3

Tobin's Q Ratio as class

Dataset	Correctly classified instances	Coverage of cases (0.95 level)	Precision Class 0	Precision Class 1	ROC Area
SPX	64.1129 %	77.8226 %	0.64	0.642	0.71
SXXP	76.6387 %	86.3866 %	0.784	0.751	0.848
EEBP	70.0337 %	81.8182 %	0.811	0.805	0.893

5. Conclusion

In this paper we demonstrated the use of artificial neural networks in classifying three datasets containing information regarding financial information and corporate governance practices from 1400 companies located in Eurozone, Eastern Europe and United States. The results show the neural networks can be successfully employed in classifying the companies using financial and corporate governance information in order to estimate their bankruptcy risk or their general performance according to two wide known indicators: Altman Z-Score and Tobin's Q Ratio.

For future development we consider using data mining methods for multi target prediction in order to classify the date according to both target variables at the same time and observe the relationship between them.

Acknowledgments

This work was co financed from the European Social Fund through Sectoral Operational Program Human Resources Development 2007-2013, project number POSDRU/159/1.5/S/134197 „Performance and excellence in doctoral and postdoctoral research in Romanian economics science domain".

References

1. Altman, E. I. (1968). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy". The journal of finance, 23(4), 589-609.
2. Botti, L., Boubaker, S., Hamrouni, A., & Solonandrasana, B. (2014). "Corporate governance efficiency and internet financial reporting quality". Review of Accounting and Finance, 13(1), 43-64.
3. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). "SMOTE: synthetic minority over-sampling technique". Journal of artificial intelligence research, 16(1), 321-357.
4. Chen, F. H., Chi, D. J., & Zhu, J. Y. (2014). "Application of Random Forest, Rough Set Theory, Decision Tree and Neural Network to Detect Financial Statement Fraud–Taking Corporate Governance into Consideration". InIntelligent Computing Theory (pp. 221-234). Springer International Publishing.
5. Chiou, C. C., & Sen-Wei, W. (2010). "A Study on the Relationship between Corporate Governance and Pricing for Initial Public Offerings: The Application of Artificial Neural Networks. In Advances in Intelligent Decision Technologies" (pp. 153-162). Springer Berlin Heidelberg.
6. Fanning, K. M., & Cogger, K. O. (1998). "Neural network detection of management fraud using published financial data. International Journal of Intelligent Systems in Accounting", Finance & Management, 7(1), 21-41.
7. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). "From data mining to knowledge discovery in databases". AI magazine, 17(3)
8. Kumar, K., & Bhattacharya, S. (2006). "Artificial neural network vs linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances". Review of Accounting and Finance, 5(3), 216-227.
9. Li, C. K., Liang, D., Lin, F., & Chen, K. L. (2015). "The Application of Corporate Governance Indicators With XBRL Technology to Financial Crisis Prediction.Emerging" Markets Finance and Trade, (ahead-of-print), 1-15.
10. Nor, S. M., & Islam, S. (2011). "Return Predictability of a Corporate Governance Neural Networks Trading System". Available at SSRN 1916643.
11. Polsiri, P., & Sookhanaphibarn, K. (2009). "Corporate distress prediction models using governance and financial variables:

evidence from Thai listed firms during the East Asian economic crisis". Journal of Economics and Management, 5(2), 273-304.

12. Soni, T. K. (2011). "Impact of Ownership Structure on Dividend Yields: A Neural Network Approach". Available at SSRN 1841629.

13. Tobin, J. (1969). "A general equilibrium approach to monetary theory". Journal of money, credit and banking, 1(1), 15-29.

14. Turnbull, S. (2002). "The science of corporate governance. Corporate Governance: An International Review", 10(4), 261-277.

15. Zhang, N., Liu, J., & Wang, X. (2015). "Analysis on the Effect and Transmission Mechanism of Corporate Governance on Security of Banking Industry". In LISS 2013 (pp. 723-728). Springer Berlin Heidelberg.