9. A Decision Support System to Predict Financial Distress. The Case of Romania

Liviu TUDOR[1]
Mădălina Ecaterina POPESCU[2]
Marin ANDREICA[3]

Abstract

Financial distress prediction has become a topic of great interest for most decision makers over the last decades, especially because of the valuable insights and effective early warnings of potential bankruptcy yielded by such prediction models. Therefore, discovering a suitable model for predicting financial distress is likely to be of great significance to global investors. Thus, this paper aims to offer a practical solution to predict financial distress in Romania by focusing on developing an integrated decision support system and on analysing the effectiveness of several prediction models based on decision trees, logit and hazard models, as well as neural networks.

Keywords: financial distress, decision support system, decision tree, logit and hazard model, neural networks

JEL Classification: G32, C40

I. Introduction

In the context of economic instability and uncertainty, when more and more companies struggle to keep in business and face serious financial difficulties, the need for a sound strategic planning and an efficient management system is quite obvious. Thus, financial distress prediction has become a topic of great interest for most decision makers over the last decades, especially because of the valuable insights and effective early warnings of potential bankruptcy yielded by such prediction models. Therefore, discovering a suitable model for predicting financial distress one year ahead is likely to be of great significance to global investors. Although the issue of bankruptcy prediction is of high relevance, previous studies mostly concentrated on building early warning models of financial distress based on financial indicators, due to the limited access to

---

[1] *Ph.D. Student at The Bucharest University of Economic Studies, E-mail: lt@idgrup.ro.*

[2] *The Bucharest University of Economic Studies and Scientific Researcher at the National Scientific Research Institute for Labour and Social Protection, E-mail: madalina.andreica@gmail.com*

[3] *The Bucharest University of Economic Studies, E-mail: marinandreica@gmail.com.*

financial data of bankrupt companies. However, since little attention has been paid to actually developing a decision support system for financial distress, this paper aims to address this gap by building several effective prediction models that will be integrated in a complex decision support system. The present research will be conducted for a set of Romanian listed companies, contributing, thus to the extension of the literature in the field for transition economies.

Having this in mind, the focus of this paper is on developing an integrated decision support system and on analysing the effectiveness of the prediction models of financial distress for the case of Romania. Hence, this paper aims to offer a practical solution for predicting financial distress in Romania one year in advance.

II. Literature Review on Financial Distress Prediction

Recently, there has been growing interest in the topic of corporate financial distress prediction or even bankruptcy prediction, especially after the economic crisis that caused economic instability and generated serious financial difficulties to a high number of companies, out of which thousands eventually turned into bankruptcy. The issue of financial distress first became a topic of great concern and a subject of thorough empirical research starting with Beaver (1966), who developed a dichotomous classification test based on a simple t-test in a univariate framework. His findings suggested that the financial indicator described by the ratio of cash flow to total debt is the best predictor of corporate bankruptcy.

Beaver’s study was then followed by Altman (1968) and many different approaches have been proposed ever since, in order to predict financial distress more efficiently. For instance, Eisenbeis (1977), Ohlson (1980) and Jones (1987) argued there were some inadequacies with the Multivariate Discriminant Analysis model used by Altman (1968) with respect to the assumptions of normality and group dispersion. Hence, the logit analysis (Ohlson, 1980) as well as the probit model (Zmijewski, 1984) were then introduced. But another issue soon aroused, since the logistic analysis only allows using single period data and works properly on the assumption that the failure process is fairly stable over a considerable period of time, which unfortunately does not hold in most cases (Hillegeist, 2004).

Shumway (2001) demonstrated that these problems could result in biased, inefficient, and inconsistent coefficient estimates and in order to overcome these problems introduced the hazard model. This is actually a multi-period logit model, since the likelihood functions of the two models are identical. The main particularities of the hazard model consist in the facts that firm specific covariates must be allowed to vary with time for the estimator to be more efficient and a baseline hazard function is also required, but which can be estimated directly with macroeconomic variables to reflect the radical changes in the environment. Such an example was proposed by Nam, Kim, Park and Lee (2008), who developed a duration model with time varying covariates and a baseline hazard function incorporating macroeconomic variables, such as exchange rate volatility and interest rate. Their findings suggest that the model built by allowing temporal and macroeconomic dependencies overcame both the traditional dichotomous static model and the logit model with time-varying covariates, but no baseline hazard function.

In recent years many heuristic algorithms, such as neural networks and decision trees have also been successfully applied to the bankruptcy prediction problem. For example, the studies made by Tam and Kiang (1992), Salchenberger *et al*. (1992) and Jain and Nag (1998) provided evidence to suggest that neural networks outperform conventional statistical models (such as discriminant analysis and logit models) in financial applications involving prediction issues.

Soon after that, hybrid Artificial Neural Network methods were used in some financial distress prediction studies and were found to outperform other models, concluding that there could be very useful in early warning systems for firm failure prediction (Yim and Mitchell, 2005)

However, Zheng and Yanhui (2007) as well as Koyuncugil and Ozgulbas (2007) highlighted several disadvantages of neural network models, consisting mainly in the difficulty of building up and interpreting the model, as well as the time required to accomplish iterative process. On the other hand, they presented the advantages of using CHAID decision trees in comparison to a neural network model, which is complicated to build and to interpret or to a statistic model such as multivariate discriminate regression and logistic regression, where the patterns need to be linearly separable and samples are assumed to follow a multivariate normal distribution.

As noticing from the literature review presented above, the bankruptcy and distress prediction issues still remain an opened challenge, especially in times of economic instability when each company’s surviving skills become crucial. In this context, early warning signals could be of great help in preventing financial distress or even bankruptcy.

III. The Architecture of the Decision Support System

This study was designed to develop a decision support system for financial distress, based on the particularities of a sample of 102 Romanian listed companies. The architecture of the proposed decision support system is further on described, based on the following four components:

- The database
- The model-base
- The knowledge management system
- The user system interface

The database consists of 14 financial ratios reflecting the company’s profitability, solvency, asset utilization, growth ability and size for a set of 102 Romanian listed companies on the Bucharest Stock Exchange over the period 2011-2013. Out of the total sample, 50 firms were facing financial difficulties, while the rest of 52 firms were considered healthy companies, as they had not registered any losses or debts during the last three financial years starting with 2011.

Since no standard definition stands for "distressed" companies, we decided to follow the same criteria used in other similar studies (Psillaki *et al*., 2008; Zheng and Yanhui, 2007) and define "distressed" if a company had losses and outstanding payments for at least 2 consecutive years. The selection of the main set of financial ratios for each company (see Table 1) was conditioned by the variables used in most empirical work and restricted by data availability.

Table 1

Financial Ratios

CATEGORY	CODE	FINANCIAL RATIOS	DEFINITION
Profitability	**I1**	Profit Margin	Net Profit or Loss / Turnover *100
	I2	Return on Assets	Net Profit or Loss / Total Assets *100
	I3	Return on Equity	Net Profit or Loss / Equity *100
	I4	Profit per employee	Net Profit or Loss / number of employees
	I5	Operating Revenue per employee	Ln(Operating revenue / number of employees)
Solvency	**I6**	Current ratio	Current assets / Current liabilities
	I7	Debts on Equity	Total Debts / Equity *100
	I8	Debts on Total Assets	Total Debts / Total Assets *100
Asset utilization	**I9**	Working capital per employee	Working capital / number of employees
	I10	Total Assets per employee	Ln(Total Assets / number employees)
Growth ability	**I11**	Growth rate on net profit	(Net P/L_1 - Net P/L_0) / Net P/L_0
	I12	Growth rate on total assets	Total $Assets_1$ – Total $Assets_0$) / Total $Assets_0$
	I13	Turnover growth	($Turnover_1$- $Turnover_0$) / $Turnover_0$
Size	**I14**	Company size	ln (Total Assets)

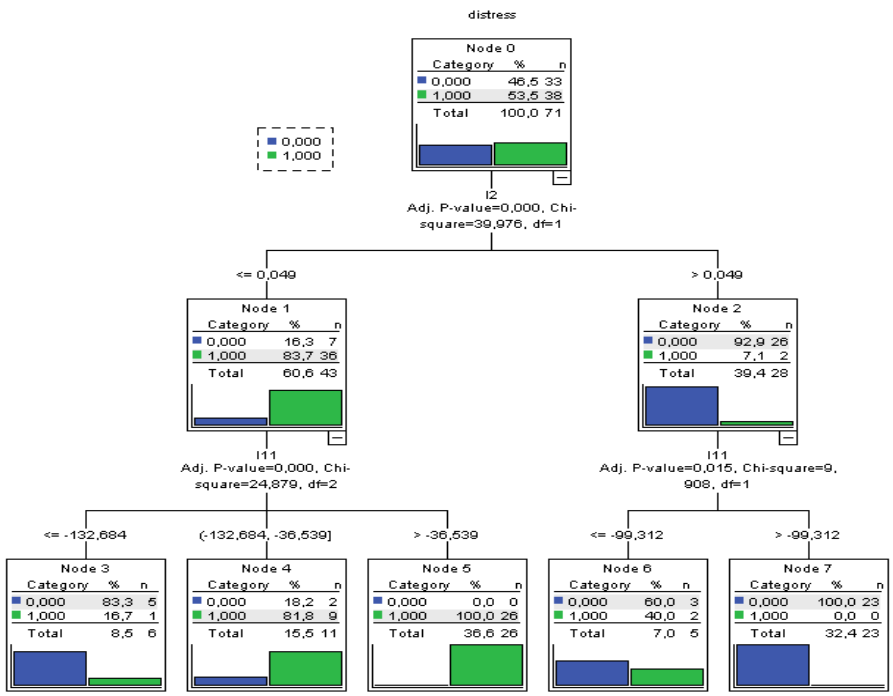
Regarding **the model-base of the decision support system,** four types of prediction models were tested: CHAID decision tree models, logit and hazard models, as well as neural network models with the purpose to predict financial distress. In all four cases the initial sample of 102 companies was divided into a 70% training sample and a 30% test sample. Then the out-of-sample performances were calculated in order to measure the models' efficiency. The prediction models are further on presented.

The CHAID Decision Tree Model for Financial Distress

According to Andreica (2008; 2009; 2013) and Popescu (2015) a decision tree is a predictive model build in the process of learning from instances, which can be viewed as a tree. Each branch of the tree is a classification question and the leaves of the tree are partitions of the dataset with their classification. Out of the main types of decision tree algorithms, Chi-square Automatic Interaction Detector called CHAID was used, as it has the advantage of generating non-binary trees. CHAID model finds the pair of values that is least significantly different with respect to the target attribute and the significant difference is measured by the Pearson chi-square test p-value. For each selected pair, CHAID checks if the obtained p-value is greater than a certain merge threshold and it merges the values in case so. It then searches for an additional potential. The two alpha levels: α_{merge} and α_{split} values were set at a 5% level.

Figure 1

The Testing Decision Tree for Financial Distress Prediction



Source: Own calculation using SPSS.

The CHAID model has three layers and two splits, indicating that the two variables that are relevant to classify the initial sample into "healthy" and "distressed" companies are *Return on Assets (ROA)* ***(I2)*** and *Growth rate on net profit* ***(I11).*** As noticing, the results indicated a profitability financial ratio and also a growth ability ratio to be the best predictors on financial distress. The selection of a financial indicator as being among the best predictors of distressed firms is also supported by Zheng and Yanhui's work (2007). When computing the prediction ability of the model based on both in-sample and out-of-sample data-sets, we notice that the decision tree has a prediction accuracy of 93% in the learning phase and smoothly drops to 90.3% in the testing phase. The results are shown in Table 2.

Table 2

Prediction Accuracy of the CHAID Models

	In-sample			**Out-of-sample**		
	healthy	**distress**	**TOTAL**	**healthy**	**distress**	**TOTAL**
Total	33	38	71	19	12	31
incorrect	2	3	5	0	3	3
correct	31	35	66	19	9	28
% incorrect	6.1	7.9	7.0	0.0	25.0	9.7
% correct	93.9	92.1	**93.0**	100.0	75.0	**90.3**

Source: Own calculations.

The LOGIT and the HAZARD Models

As already presented in the literature review regarding distress prediction, there is quite a large number of studies focusing on the logistic and hazard models in order to predict the probabilities of a company to become distressed in the following years. Therefore, both econometric models were applied in this study in order to predict financial distress one year ahead. Based on Shumway's (2001) theory, the main difference between the logistic and the hazard model is that the classical dichotomous static model uses only one year financial data for each company, while a hazard model is actually a multi-period logit model that considers each annual financial ratio of a company to be distinct observations. Thus, the logistic model was built using financial ratios of the year 2013, while for the hazard model all financial ratios available for the period 2011-2013 were taken into consideration. Given the short time frame for which financial data was available in this study we considered to be more appropriate building a hazard model with time invariant baseline hazard function.

The following steps were taken to find the best logistic and hazard model for distress prediction:

- First, a backward looking procedure was applied, by estimating a logistic model with all financial ratios included as explanatory variables, followed by a step by step procedure of exclusion any statistically insignificant variable.
- Then, for each resulting model, each coefficient sign was checked if it corresponds to the economic theory and in case of contradiction, the corresponding variable would be dropped.
- Lastly, the remaining models (in case of more than just one) were compared based on the following criteria: out-of-sample performance, McFadden value, LR value, AIC value, the goodness of fit Test (H-L Statistics) and total gain in comparison to the simple constant model and the best model was then selected.

The output estimations of the resulted logit and hazard models are presented in Table 3. Surprisingly, both binary models identified the same two financial ratios as the best predictors of financial distress, namely *Profit margin* (I1) and *Debts on Total Assets* (I8).

However, when testing the out-of-sample prediction accuracy of the logit model, it resulted that it predicted one year ahead financial distress with quite a high probability of 87.3% in the learning phase, but dropped to only 77.4% in the testing phase. On the other hand, when checking the prediction accuracy of the hazard model with time invariant baseline hazard function, we notice that the model outperforms the logit model in the testing phase, by over 5.8 percentage points. The high probability of 82.8% of correct prediction of the hazard model suggests that econometric models yield better prediction results when the variables' time invariant restriction is dropped.

Table 3

Output Estimations of the Logit and the Hazard Model

THE LOGIT MODEL		THE HAZARD MODEL	
Coefficients		*Coefficients*	
I1	-0.06406 (0.0212)***	I1	-0.01803 (0.0052)***
I8	0.01937 (0.0113)*	I8	0.02309 (0.0052)***
Const.	-2.0459 (0.6917)***	Const.	-1.5776 (0.3067)***
Pseudo R^2	52.14%	*Pseudo R^2*	31.70%
LR chi2	51.14***	*LR chi2*	93.28***
Goodness of fit Test	91.99%	*Goodness of fit Test*	33,8%
Total Gain in comparison to the simple constant	87.32%	*Total Gain in comparison to the simple constant*	77,9%
In-sample prediction accuracy	87.3%	*In-sample prediction accuracy*	78.4%
Out-of-sample prediction accuracy	77.4%	*Out-of-sample prediction accuracy*	82.8%

*Notes: Between brackets are the standard errors, and the ***,**,* stand for 1%, 5% and 10% significance level, respectively.*
Source: Own calculations.

Artificial Neural Network Model

The Artificial Neural Network (ANN) was built using all 14 financial ratios corresponding to the year 2013 in order to obtain distress prediction models. Before feeding the data into the neural network, some variable transformations were required. For instance, all positive values of the financial indicators were rescaled to the [0,1] range, while all negative values were rescaled within the interval [-1,0]. The ANN was built based on the following structure: one input layer, one hidden layer (with only one neuron) and one output layer and was trained on the same data sets as the previous methods. The same division of the initial sample was kept, meaning 70% of observations for the learning phase and 30% for the testing phase. Prediction based on the neural network had accuracy of 83.9% in the testing phase and of 91.5% respectively in the learning phase. The results are presented in table 4, were the total number of correct and the incorrect prediction cases was also computed.

Table 4

Prediction Accuracy of the ANN Model

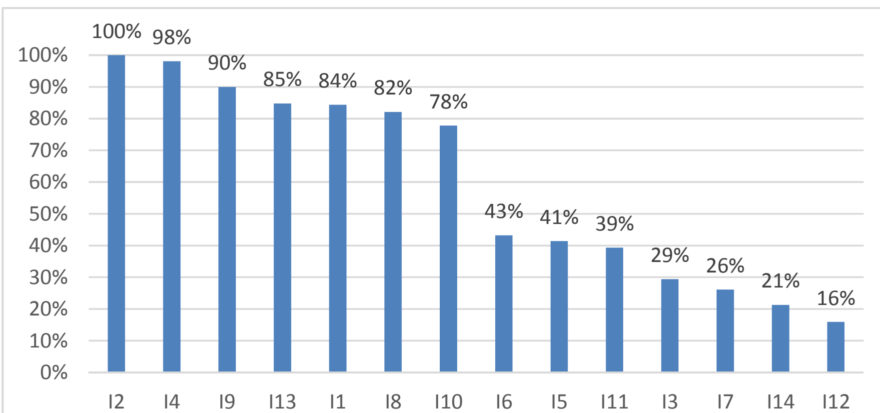
	In-sample			**Out-of-sample**		
	healthy	**distress**	**TOTAL**	**healthy**	**distress**	**TOTAL**
Total	33	38	71	19	12	31
incorrect	1	5	6	2	3	5
correct	32	33	65	17	9	26
% incorrect	3.0	13.2	8.5	10.5	25.0	16.1
% correct	97.0	86.8	**91.5**	89.5	75.0	**83.9**

Source: Own calculations.

If we were to consider the weights of the 14 financial indicators in the neural network, our findings indicate that the most relevant predictors of financial distress using ANN have proven to be the following: *Return on Assets* (ROA) ***(I2), Profit per employee (I4), Working capital per employee (I9), Turnover growth (I13), Profit margin (I1)*** and *Debt on total assets* ***(I8),*** suggesting once again just how relevant profitability indicators are, followed by indicators of asset utilization, solvency and growth ability.

Figure 2

Weights of the Financial Ratios in the ANN



Source: Own calculations.

Regarding **the knowledge management system component**, the decision tree plays an important role not only by defining the variables that can be used in the measurement of financial distress, but also by determining consistent classification rules, mainly because of the tree structure and its ability to easily generate rules for segmentation of the original database. Since a decision tree generates a rule for each of its leaves, in the prediction model there are three classification rules, based on the values of the I2 and I11. More precisely, the decision tree classifies a company as being healthy if I2 is higher than 0.049. In the other case, the company is considered distressed only if the I11 is higher than -132.7%. However, it is obvious that these rules are very sensitive to the initial data set.

IV. Conclusions

In this paper, we showed how a decision support system provides early warning signals of financial distress to a company one year ahead. The aim of this paper consisted in offering a practical solution for predicting financial distress in Romania by focusing on developing an integrated decision support system and on analysing the effectiveness of several prediction models based on decision trees, logit and hazard models, as well as neural networks. Out of the four prediction models tested, best out-of-sample results were obtained by the CHAID decision tree model. The prediction accuracy of the classification tree was quite high, reaching over 90% in the testing phase, as compared to the neural networks (83.9%), the Hazard model with time invariant function (82.8%) or the single-period logit model (77.3%).

In addition, regarding the top best predictors of financial distress, profitability ratios turned out to perform best. The results are consistent to those obtained in other similar studies (Zheng and Yanhui, 2007; Koyuncugil and Ozgulbas, 2007). Relevant conclusions can also be drawn from these findings, regarding the most important financial indicators recommended to effectively predict financial distress in Romania. More precisely, in case of using financial ratios of the year 2013 in order to predict financial distress one year ahead, the decision tree model identified one profitability ratio, *Return on Assets (ROA)* and one growth ability ratio, namely *Growth rate on net profit* as best predictors. The two binary econometric models identified the same two financial ratios as the best predictors of financial distress, namely *Profit margin* and *Debts on Total Assets*, while for the case of the Artificial Neural Network all ratios were included in the model, but the following were proven to have the highest weights: *Return on Assets* (ROA)*, Profit per employee, Working capital per employee, Turnover growth, Profit margin* and *Debt on total assets,* suggesting once again just how relevant profitability indicators are, followed by indicators of asset utilization, solvency and growth ability of a company.

Thus, we can conclude that our results are consistent with the economic theory and the literature review and the high prediction accuracy of the model of over 90% suggest that the proposed decision support system can become a practical solution for any decision maker.

References

Andreica, M., Nicolae, D., Andreica, M.E., Todor, L. (2009) „Working out forecasts on the Basis of Statistical Extrapolation Principle”, *Economic Computation and Economic Cybernetics Studies and Research*, (2) 43: 85-108

Andreica, M.E. (2013), “Early warning models of financial distress. Case study of the Romanian firms listed on RASDAQ”, *Theoretical and Applied Economics*, Vol. XX, No. 5(582): 4-12

Andreica, M.I., Tapus N. (2008). Reliability Analysis of Tree Networks Applied to Balanced Content Replication, *Proceedings of the IEEE Int. Conf. on Automation, Quality and Testing*, Robotics: 79-84

Altman, E.I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance* 23: 589-609.

Beaver, W. (1966), "Financial ratios as predictors of failure", *Journal of Accounting Research (Supplement)* 4: 71-102.

Eisenbeis, R. (1977), “Pitfalls in the application of discriminant analysis in business, finance and economics”, *Journal of Finance* 32: 875-900

Hillegeist, S.A., Keating, E.K., Cram, D.P., Lundstedt, K.G. (2004), “Assessing the probability of Bankruptcy”, *Review of Accounting Studies* 9: 5-34.

Jain, B.A., Nag, B.N. (1998), „A neural network model to predict long-run operating performance of new ventures”, *Annals of operations research* 78: 83 – 110

Jones, F.L. (1987), "Current techniques in bankruptcy prediction", *Journal of Accounting Literature* 6: 131-164.

Koyuncugil, A.S., Ozgulbas, N. (2007), “Detecting financial early warning signs in Istanbul Stock Exchange by data mining”, *International Journal of Business Research*.

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

Nam, C.W., Kim, T.S., Park N.J., Lee, H.K. (2008), “Bankruptcy prediction using a Discrete-Time Duration Model Incorporating Temporal and Macroeconomic dependencies”, *Journal of Forecasting*, 27: 493-506.

Ohlson, J. A. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, 18:109-131.

Popescu, M.E. (2015) - Proposal for a decision support system to predict financial distress, *Review of International Comparative Management*, Vol. 16(1): 112-118

Psillaki,M., Tsolas, I.E., Margaritis, D. (2008), “Evaluation of credit risk based on firm performance”, (EFMA) European Financial Management Association 2008 Annual Meetings June 25- 28, 2008 Athens, Greece

Salchenberger, L.M., Cinar, E.M. Lash, N.A. (1992), „Neural networks: A new tool for predicting thrift failures”, *Decision Sciences* 23: 899 – 916.

Shumway, T. (2001), "Forecasting bankruptcy more accurately: A simple hazard model", *Journal of Business*, 74 (1): 101-124.

Tam, K.Y., Kiang, M.Y (1992), „Managerial applications of neural networks: The case of bank failure predictions”, *Management Science* 38: 926 – 947.

Yim J., Mitchell, H. (2005), “A comparison of corporate distress prediction models in Brazil: hybrid neural networks, logit models and discriminant analysis”, *Nova Economia Belo Horizonte* 15: 73-93

Zheng, Q., Yanhui, J. (2007), “Financial Distress Prediction on Decision Tree Models”, IEEE.

Zmijewski, M.E.(1984), "Methodological issues related to the estimation of financial distress prediction models", *Journal of Accounting Research*, 22:59-86.