A Statistical-Based Approach to Assessing Comparatively the Performance of Non-Banking Financial Institutions in Romania

Adrian Costea

Abstract In this paper we construct a framework that enables us to make class predictions about the performance of non-banking financial institutions (NFIs) in Romania. Our objective is to create a classification model in the form of a logistic regression function that can be used to assess the performance of NFIs based on different performance dimensions, such as capital adequacy, assets' quality and profitability. Our methodology consists of two phases: a clustering phase, in which we obtain several clusters that contain similar data-vectors in terms of Euclidean distances, and a classification phase, in which we construct a class predictive model in order to place the new row data within the clusters obtained in the first phase as they become available. Our goal is two-fold: to validate the dimensionalities of the map used to represent the performance clusters and the quantisation error associated with it and to use the obtained model to analyze the movements of three largest NFIs during the period 2007–2010. Using our validation procedure that is based on a bootstrap technique, we are now able to find the proper map architecture and training–testing dataset combination for a particular problem. At the same time, using the visualization techniques employed in the study, we understand how different financial factors can and do contribute to the companies' movements from one group/cluster to another. Furthermore, the classification model is validated based on high training and testing accuracy rates.

Keywords Non-banking financial institutions • Performance evaluation • Logistic regression • Class prediction

JEL Classification Codes C38 • C81 • G23

A. Costea (✉)
Academy of Economic Studies, Virgil Madgearu Building, Calea Dorobantilor No. 15-17, sector 1, Bucharest 010552, Romania
e-mail: adrian.costea@csie.ase.ro

A. Karasavvoglou and P. Polychronidou (eds.), *Economic Crisis in Europe and the Balkans*, Contributions to Economics, DOI 10.1007/978-3-319-00494-5_11,

1 Introduction

The aim of this paper is to analyze comparatively the financial performance of a number of non-banking financial institutions (NFIs) in Romania by the means of Data Mining techniques. This type of analysis could support the Supervision Department of the National Bank of Romania in its current activity: the supervision authority can identify those institutions that present a lower than average level of financial stability, thus concentrating its scarce resources (time and personnel) on these particular entities. At the same time, an analysis of the biggest NFIs in terms of total assets would be of interest for judging the stability of the entire sector. Other stakeholders (decision-makers, creditors, investors) can benefit from this type of analysis. Decision-makers in the companies involved in the analysis can understand the causes of their business problems by learning from others' achievements/ mistakes. Creditors can obtain a general picture about the financial situation of different companies that would help them manage their credit exposure. Using our models, investors would be able to weigh the different investment opportunities.

Currently, in Romania, NFIs performance evaluation is done manually by consulting their prudential reporting. Periodic financial statements (PFSs) contain a number of raw indicators for NFIs' performance which are analyzed manually by inspectors. Until now it is not possible to perform a comparative analysis of several NFIs or a dynamic analysis of one of these entities based on the indicators of the PFSs, except by considerable effort from the inspectors of the Supervision Department. This is due to the complexity of the problem involving dynamic analysis (for a considerable number of quarters) of all NFIs included in the Special Register (about 65) in terms of a set of 10–15 performance indicators. However, unlike the NFIs' performance evaluation (rating) models (which are non-existent), the Supervision Department developed the Uniform Assessment System – CAAMPL (Cerna et al. 2008) for evaluating the credit institutions (banks). CAAMPL system assesses the performance of credit institutions based on six dimensions: capital adequacy (C), shareholders' quality (A), assets' quality (A), management (M), profitability (P) and liquidity (L). Each performance dimension is evaluated based on a number of indicators and a composite rating is calculated. Except from being inapplicable for assessing the performance of NFIs, the CAAMPL rating system presents some disadvantages, such as:

- It uses simple linear techniques for discriminating the multidimensional space represented by the independent variables (financial performance ratios). In fact, the discrimination model is not a multivariate discrimination model (i.e.: a model that takes into consideration more than one discriminating variable at a time), but a sequential combination of univariate models;
- The selection of independent variables (performance criteria) that determine a rating (a specific class of performance) is not based on scientific rigour, but on the practical experience of the members of the supervision authority;
- As a result of this heuristic selection it is difficult to substantiate the various limits for the independent variables that determine the performance indicator

(rating), which leads to a significant increase in the analyst's subjective involvement in establishing it;
- CAAMPL evaluation system by which the performance of the credit institutions is assessed (the ratings are established) is based mainly "*on rules*" as emphasized by the IMF in IMF (2010) and does not involve quantitative methods for assessing the performance.

While still in place and useful, the CAAMPL system need to be challenged. This challenge is provided by Computational-Intelligence (CI) methods which come from different fields: *machine learning*, *artificial intelligence, evolutionary computation* and *fuzzy logic*. The Knowledge Discovery in Databases (KDD) process (Fayyad et al. 1996) and its engine—Data Mining (DM)—represent the umbrella under which the CI methods operate. In a previous paper (Costea 2011) we formalized the process of NFIs' financial benchmarking by considering this real-world application as a knowledge discovery problem and by following the formal steps of the KDD process. Each business problem (real-world application) can be matched by many data-mining tasks depending on how we approach the problem. We match our real-world application (assessing comparatively the performance of NFIs) with both DM clustering and classification tasks. We use clustering methods in order to find patterns (models) that describe the financial situation of NFIs and classification methods for financial (class) predictions.

Here we analyze only those NFIs registered in the Special Register that have as main activity financial leasing and have been active since the introduction of the regulatory framework for these institutions in Romania. The algorithms used to perform DM tasks mentioned above are numerous and they come from different research fields. In this paper, we use an heuristic method (neural networks with unsupervised learning algorithm known as Self-Organizing Map algorithm) for the DM clustering task, and a statistical approach (multinomial logistic regression) for performing the DM classification task.

The scientific literature in applying DM techniques for financial performance benchmarking is relatively rich. In the next section we engage in a thorough literature review regarding the application of CI methods in assessing comparatively companies' financial performance. Then, we present our methodology and data. Finally, we perform an experiment by analysing the movements of three largest NFIs in terms of total assets during the period 2007–2010 and present our concluding remarks.

2 Literature Review

We found several models for evaluating the performance of financial entities, applicable mainly to the credit institutions. In Collier et al. (2003) the authors described the characteristics of the off-site monitoring instrument of the FDIC (Federal Deposit Insurance Corporation) and the data used in its development.

Doumpos and Zopounidis (2009) proposed a new classification system for the credit institutions as a support-tool for the analysts from the National Bank of Greece. The system provides a rich set of assessment, visualization and reporting options. Swicegood and Clark (2001) compare three models (based on discriminant analysis, neural networks and professional human judgment) used to predict underperformance of commercial banks. Neural networks based model showed better predictive capacity than the other two models.

Boyacioglu et al. (2009) proposed several methods for classifying credit institutions based on 20 performance indicators grouped into six dimensions (CAMELS). They used four sets of financial data, the results showing that among the clustering and classification techniques tested, the best in terms of accuracy rates were neural networks.

Ravi Kumar and Ravi (2007) makes a literature review for research conducted during 1968–2005 on the application of statistical and computational intelligence methods in banks' or firm's bankruptcy prediction. For each source of data, the authors show the indicators used, the country of origin and the period of data collection. Şerban et al. (2011) apply computational intelligence methods (e.g. clustering techniques) to classify the shares from Bucharest Stock Exchange which had profit during the last 2 years, in order to find similarities and differences between these shares and build a diversified portfolio.

The SOM algorithm was used extensively in assessing comparatively companies' financial performance. There are two pioneer works applying the SOM to companies' financial performance assessment. One is Martín-del-Brío and Serrano Cinca (1993) followed by Serrano Cinca (1996, 1998a, b). Martín-del-Brío and Serrano Cinca (1993) propose SOM as a tool for financial analysis. The sample dataset contains 66 Spanish banks, of which 29 went bankrupt. Martín-del-Brío and Serrano Cinca (1993) use 9 financial ratios, among which there are 3 liquidity ratios: current assets/total assets, (current assets – cash and banks)/total assets, current assets/loans, 3 profitability ratios: net income/total assets, net income/total equity capital, net income/loans, and 3 other ratios: reserves/loans, cost of sales/sales, and cash flows/loans. A solvency map is constructed, and different regions of low liquidity, high liquidity, low profitability, high cost of sales, etc. are highlighted on the map. Serrano Cinca (1996) extends the applicability of SOM to bankruptcy prediction. The data contain five financial ratios taken from Moody's Industrial Manual from 1975 to 1985 for a total of 129 firms, of which 65 are bankrupt and the rest are solvent. After a preliminary statistical analysis, the last ratio (sales/total assets) is eliminated because of its poor ability to discriminate between solvent and bankrupt firms. Again, a solvency map is constructed and, using a procedure to automatically extract the clusters, different regions of low liquidity, high debt, low market values, high profitability, etc. are revealed. Serrano Cinca (1998a, b) extends the scope of the Decision Support System proposed in the earlier studies by addressing, in addition to corporate failure prediction, problems such as: bond rating, the strategy followed by the company in relation to the sector in which it operates based on its published accounting information, and comparison of the financial and economic indicators of various countries.

The other major SOM financial application is Back et al. (1998), which is an extended version of Back et al. (1996). Back et al. (1998) analyse and compare more than 120 pulp-and-paper companies between 1985 and 1989 based on their annual financial statements. The authors used 9 ratios, of which 4 are profitability ratios (operating margin, profit after financial items/total sales, return on total assets, return on equity), 1 is an indebtedness ratio (total liabilities/total sales), 1 denotes the capital structure (solidity), 1 is a liquidity ratios (current ratio), and 2 are cash flow ratios (funds from operations/total sales, investments/total sales). The maps are constructed separately for each year and feature planes are used to interpret them. An analysis over time of the companies is conducted by studying the position each company has in every map.

One of the pioneer works in applying *discriminant analysis* (DA) to assessing comparatively companies' financial performance is Altman (1968). Altman calculated discriminant scores based on financial statement ratios such as working capital/total assets; retained earnings/total assets; earnings before interest and taxes/total assets; market capitalisation/total debt; sales/total assets. Ohlson (1980) is one of the first studies to apply *logistic regression* (LR) to predicting the likelihood of companies' bankruptcy. Since it is less restrictive than other statistical techniques (e.g. DA) LR has been used intensively in financial analysis. De Andres (2001, p. 163) provides a comprehensive list of papers that used LR for models of companies' financial distress.

3 Methodology and Data

Our methodology consists of two phases: a clustering phase, in which we obtain several clusters that contain similar data-vectors in terms of Euclidean distances, and a classification phase, in which we construct a class predictive model in order to place the new row data within the clusters obtained in the first phase as they become available.

In the first phase, we employ unsupervised neural networks in terms of Kohonen' Self-Organizing Maps (SOM) algorithm, in order to build clusters that include NFIs with similar performance (in terms of financial ratios). Based on the SOM, we construct a two-dimensional unified-distance matrix map (a two-dimensional representation technique for the distance between neurons). Then, we characterize each cluster as containing NFIs with good, average or poor performance by looking at the feature planes for each individual input variable. Based on this characterization, we build the "class performance" variable by attaching to each data row a class label depending onto which cluster it belongs.

In the second phase, we employ a statistical technique, namely multinomial logistic regression, in order to build a classification model that links the newly constructed "class performance" variable to the input variables (financial performance ratios). We build this classification model in order to avoid the problems associated with adding new data to an existing SOM cluster model. Inserting new data into an existing SOM model becomes a problem when the data have been standardized, for example, within an interval like [0,1]. Also, the retraining of maps requires considerable time and expertise.

We applied our methodology on NFIs' performance dataset. The data were collected annually from 2007 to 2010 for the NFIs registered in the Special Register that have as main activity financial leasing.

3.1 The SOM

The SOM (Self-Organising Map) algorithm is a well-known unsupervised-learning algorithm developed by Kohonen in the early 80's and is based on a two-layer neural network (Kohonen 1997). The algorithm creates a two-dimensional map from *n*-dimensional input data. After training, each neuron (unit) of the map contains input vectors with similar characteristics, e.g. NFIs with similar financial performance. The result of SOM training is a matrix that contains the codebook vectors (weight vectors). The SOM can be visualised using the *U-matrix* method proposed by Ultsch (1993). The unified distance matrix or U-matrix method computes all distances between neighbouring weights vectors. The borders between neurons are then constructed on the basis of these distances: dark borders correspond to large distances between two neurons involved, while light borders correspond to small distances. In this way, we can visually group the neurons ("raw" clusters) that are close to each other to form supra-clusters or "real" clusters (Fig. 1a).

In addition to the U-matrix map, a *component plane* or *feature plane* can be constructed for each individual input variable. In the feature planes light/"warm" colours for the neurons correspond to high values, while dark/"cold" colours correspond to low values (Fig. 1b). The component plane representation can be considered a "sliced" version of the SOM, where each plane shows the distribution of one weight vector component (Alhoniemi et al. 1999, p. 6). Also, *operating points* and *trajectories* (Alhoniemi et al. 1999, p. 6 and Fig. 1a gray line) are used to find how different points (observations) move around on the map (e.g. how the countries evolved over time with respect to their economic performance).

Many researchers have focused on applying SOM to perform the DM clustering task in general, and economic/financial performance benchmarking in particular. Oja et al. (2003) cites 5384 scientific papers – published between 1981 and 2002 – that use the SOM algorithms, have benefited from them, or contain analyses of them. However, relatively few of them (73) have applied SOM to business-related issues (Oja et al. 2003).

There are two main differences between our study and those referred to in terms of using the SOM as a performance-benchmarking tool. One difference comes from the limitation that techniques such as the SOM have: in essence they constitute

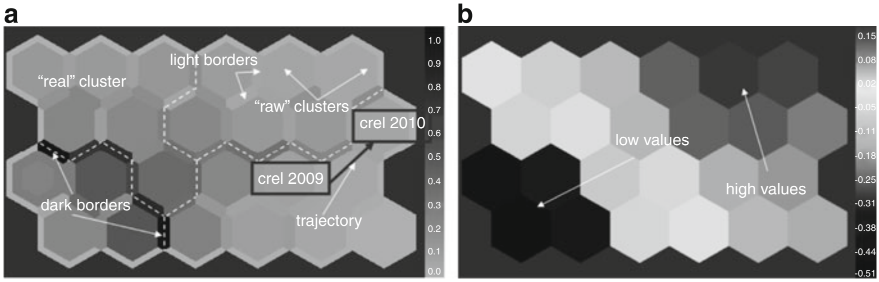

Fig. 1 (**a**) The U-matrix representation with Nenet v1.1a software program and (**b**) some variable component plane

descriptive data analysis techniques and aim at summarising the data by transforming it into a two-dimensional space and preserving the dissimilarities between observations. Employing the SOM does not imply that the use of other well-known techniques is renounced; rather, it is very productive to complement it with other tools (Serrano Cinca 1998a). Consequently, in this study, we go one step further and use the output of the SOM as the input for the classification models. Moreover, another distinction with the other studies is that, in our research, we answer some technical questions related to the practical implementation of the SOM as a performance-benchmarking tool. We have addressed two technical SOM problems: the validation of map topology and quantisation error.

3.2 Multinomial Logistic Regression

Multinomial Logistic Regression (MLR) classifies cases by calculating the likelihood of each observation belonging to each class. The regression functions have a logistic form and return the likelihood (the odds) that one observation (x) belongs to a class (C):

$$odds(x \in C) = \frac{1}{1 + e^{-logit}} = \frac{1}{1 + e^{-(w_0 + w_1 v_1 + \ldots + w_p v_p)}} \quad (1)$$

where $v_1, \ldots v_p$ are the input variables, and $w_0, \ldots, w_p$ are the regression coefficients (weights).

MLR calculates the estimates ($\hat{w}_i,\ i = 0, \ldots, p$) for the coefficients of all regression equations using the maximum likelihood estimation (MLE) procedure. If there are c classes, MLR builds c-1 regression equations. One class, usually the last one, is the reference class.

MLR calculates the standard errors for the regression coefficients, which show the potential numerical problems that we might encounter. Standard errors larger than 2 can be caused by multicolinearity between variables (not directly handled by

SPSS or other statistical packages) or dependent variable values that have no cases, etc. (Hosmer and Lemeshow 2000).

Next, MLR calculates the *Wald* statistic, which tests whether the coefficients are statistically significant in each of the c-1 regression equations. In other words, it tests the null hypothesis that the logit coefficient is zero. The Wald statistic is the ratio of the unstandardised logit coefficient to its standard error (Garson 2005).

Next, MLR shows the degree of freedom for the Wald statistic. If "sig." values are less than the 1 – confidence level (e.g. 5 %) then the coefficient differs significantly from zero. The signs of the regression coefficients show the direction of the relationship between each independent variable and the class variable. Positive coefficients show that the variable in question influences positively the likelihood of attaching the specific class to the observations.

Values greater than 1 for $e^{\hat{w}_i}$ show that the increase in the variable in question would lead to a greater likelihood of attaching the specific class to the observations. For example, if $e^{\hat{w}_1} = 3$ for class c_1 and variable v_1, we can interpret this value as follows: for each unit increase in v_1 the likelihood that the observations will be classified in class c_1 increases by approximately three times.

Finally, MLR shows the lower and upper limits of the confidence intervals for the $e^{\hat{w}_i}$ values at the 95 % confidence level.

Statistical techniques were deployed first to tackle the classification task: univariate statistics for prediction of failures introduced by Beaver (1966), linear discriminant analysis (LDA) introduced by Fisher (1936), who first applied it to Anderson's iris dataset (Anderson 1935), multivariate discriminant analysis (MDA) – Altman (1968), Edmister (1972), Jones (1987), and probit and logit (logistic) models – Ohlson (1980), Hamer (1983), Zavgren (1985), Rudolfer et al. (1999).

3.3 The Dataset

In this paper we assess comparatively the performance of different NFIs. We base our variables choice on the existing Uniform Evaluation Systems – CAAMPL (Cerna et al. 2008) applicable in the case of credit institutions or banks. The CAAMPL system uses the financial reports of credit institutions and evaluates six components that reflect in a consistent and comprehensive manner the performance of banks in concordance with the banking laws and regulations in force: capital adequacy (C), quality of ownership (A), assets' quality (A), management (M), profitability (P), liquidity (L). In this application we have restricted the number of the performance dimensions to three quantitative dimensions, namely: capital adequacy (C), assets' quality (A) and profitability (P). The other quantitative dimension used in evaluating the credit institutions (liquidity dimension) is not applicable to NFIs, since they do not attract retail deposits. We have also eliminated the qualitative dimensions from our experiment (quality of ownership and management) because they involve a distinct approach and it was not the scope of this study to take them into account.

After choosing the performance dimensions, we select different indicators for each dimension based on the analysis of the periodic financial statements of the NFIs: Equity ratio (Leverage) = own capital/total assets (net value) for the "capital adequacy" dimension, Loans granted to clients (net value)/total assets (net value) for the "assets' quality" dimension and Return on assets (ROA) = net income/total assets (net value) for the "profitability" dimension. The data were collected with the help of the members of the NFIs' Supervision Unit within the Supervision Department of the National Bank of Romania. The data were collected annually from 2007 to 2010 for the NFIs registered in the Special Register that have as main activity financial leasing and have been active since the introduction of the regulatory framework for these institutions in Romania. In total there were 11 NFIs that met the above criteria and 44 observations (11 NFIs $\times$ 4 Years = 44 observations). In the following table we present some descriptive statistics related to the financial ratios used to evaluate the NFIs' performance.

As it can be seen from Table 1, the NFIs with a negative own capital have substantially influenced the mean of Leverage financial ratio which takes a negative value. In average 69.5 % of total assets are used for loans issued by the specific NFIs during the period 2007–2010. The highest variance is encountered for Leverage, and the second highest for the assets' quality indicator. The financial ratio that is closest to the normal distribution is ROA (Kurtosis = 1.72, Skewness = -1.23). Minimum and maximum values for the financial ratios show that the dataset contains companies that are highly indebted (high negative values for the Leverage), have issued a lot of loans (value close to 1 for the Loans/Assets ratio), and have high profitability (maximum value for ROA – 9.5 %).

4 Experiment

We applied our methodology to the NFIs' financial performance dataset. We tried to validate the SOM dimensionalities according to empirical measures presented in DeBodt et al. (2002). For each map dimensionality (4×4, 5×5, 6×6, 7×7, 8×8, 9×9) we used 100 bootstrap datasets to train the SOM. We expected the variation coefficients of the quantisation error vectors to increase with the map dimensionality. However, we obtained very small variation coefficients (approx. 2 %) for all architectures, which did not allow us to reject any architecture. Therefore, a final 6×4 SOM map was chosen based on the ease-of-readability criterion. For this SOM architecture we tested three quantisation errors: one obtained when all the data are used for training and testing the SOM ("100-100" case), another when 90 % of data are used for both training and testing ("90-90" case), and the other when 90 % is used for training, and the remaining 10 % for testing ("90-10" case). Again, for each training–testing dataset combination we extracted 100 bootstrap datasets from the original data and obtained a quantisation error vector for each combination.

Table 1 Descriptive statistics for the financial performance ratios

	Leverage	Loans/assets	ROA
Mean	−0.01467	0.695108	−0.02986
Standard error	0.032453	0.021073	0.012139
Median	0.035598	0.718977	−0.01319
Standard deviation	0.215268	0.139781	0.08052
Sample variance	0.04634	0.019539	0.006483
Kurtosis	8.925482	−0.84266	1.717818
Skewness	−2.82907	−0.45214	−1.22532
Range	1.122737	0.482275	0.353689
Minimum	−0.90823	0.420091	−0.25866
Maximum	0.214509	0.902366	0.095032
Sum	−0.6454	30.58474	−1.31378
Count	44	44	44

Then, we used *t*-tests to compare the means of the three vectors. The *t* statistic is obtained by dividing the mean difference (of the two vectors) by its standard error. The significance of the *t* statistic (p-values < 0.05) tells us that the difference in quantisation error is not due to chance variation, and can be attributed to the way we select the training and testing sets. Even though we found some differences between the quantisation error vectors, the confidence in the results was rather poor (p-value for "100-100" – "90-90" pair was 0.051). Finally, we followed the "100-100" case using the entire dataset to train and test the 6×4 SOM. Even if in this particular case they were not of much help, these empirical validation procedures allow us to choose more rigorously the SOM parameters. Finally, the SOM parameters chosen were: $X = 6$, $Y = 4$, training length $1 - rlen_1 = 1{,}000$, learning rate $1 - \alpha_1(0) = 0.05$, radius $1 - N_1(0) = 6$, training length $2 - rlen_2 = 10{,}000$, learning rate $1 - \alpha_2(0) = 0.02$, radius $2 - N_2(0) = 2$.

The final 6×4 SOM map with the identified "real" clusters (dotted lines) (shown in Fig. 2) was the best in terms of quantization error (0.074522).

We used U-matrix method to group the "raw" clusters into "real clusters". This is done by looking at the borders between neurons in the map, by analysing the component plane for each input variable and the observations that belong to each cluster. In this way we have identified four "real" clusters (clusters A, B, C, and D in Fig. 2) which are described as follows (see Table 2):

Cluster A includes the NFIs with the highest values for the input variables measuring capital adequacy and profitability and second highest values registered for the variable measuring the assets' quality. This "real" cluster contains eight observations. It is the only cluster with positive average profitability ratios. Cluster B is the largest cluster containing half of the total observations (22 observations). It includes NFIs with medium capital adequacy and profitability and highest value for the variable measuring assets' quality. All ratios in cluster C have average values. However, this cluster contains NFIs with a lower performance than those in cluster B. Both cluster B and C contain NFIs that in average have negative profitability

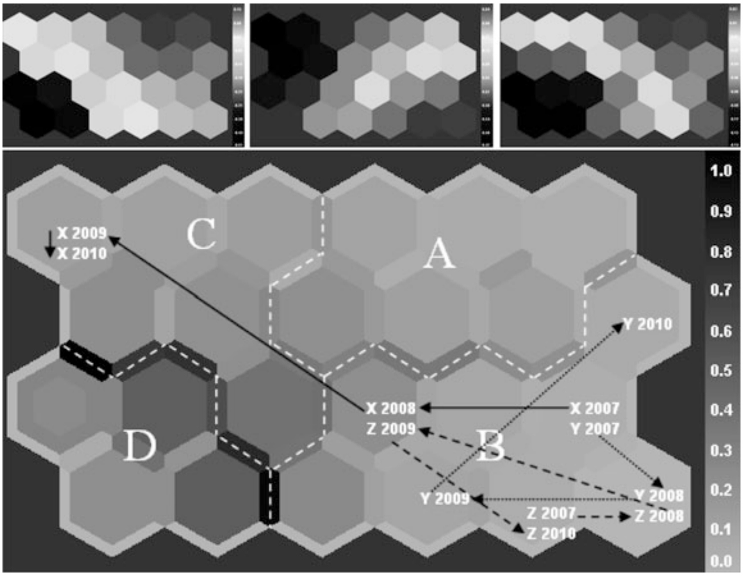

Fig. 2 The final 6 × 4 map with identified "real" clusters and the component planes for the three variables: Equity ratio (Leverage), Loans granted to clients (net value)/total assets (net value) and Return on assets (*ROA*). The trajectories (*black arrows*) between 2007 and 2010 for the largest three NFIs ("*solid-line*" *arrows* for company X, "*dotted-line*" *arrows* for company Y, and "*dashed-line*" *arrows* for company Z)

Table 2 The characterization of the clusters obtained by applying SOM algorithm

Cluster	# of obs.	Leverage	Loans/assets	ROA
A	8	0.147659	0.63482	0.008182
B	22	0.029809	0.811298	−0.02236
C	9	0.013916	0.532348	−0.02504
D	5	−0.52154	0.573298	−0.13241

ratios. Cluster D contains the worst performers. All performance ratios show low values. Again, the profitability ratios are negative in average.

The SOM trajectories can be used to check the financial performance of the different NFIs over time. The trajectories in Fig. 2 show the movements of the three largest NFIs (in terms of average total assets for 4 years – 2007–2010): the largest denoted with X (solid-line), the second largest denoted with Y (dotted-line) and the third largest denoted with Z (dashed-line) between 2007 and 2010.

For example, company X started in cluster B in 2007 and 2008, but dropped to cluster C the following year and remained there in 2010. This was partially due to a greater decrease in own capital as compared to a smaller increase in total assets. At the same time, in 2009 the loans granted by company X have decreased dramatically as compared to 2008, reaching almost a 50 % decrease.

Once we had constructed the "real" clusters, we built the class variable, assigning a class value (1–4) to each observation within a cluster. Next, we applied

MLR to build the classification models by following the methodological steps (Costea 2005). We used SPSS to perform the classification. We used our dataset without preprocessing the data given the values of the ratios are already standardised in a $[-1; 1]$ interval. We validated our models based on the training data by using proportional by-chance and maximum by-chance accuracy rates. Both criteria require the classification accuracy to be 25 % better than the *proportional by-chance accuracy rate* and *maximum by-chance accuracy rate* respectively (Hair et al. 1987, pp. 89–90). The proportional by-chance accuracy rate is calculated by summing the squared proportion of each group in the sample: the square proportion of cases in class 1 + … + the square proportion of cases in class *n*. The maximum by-chance accuracy rate is the proportion of cases in the largest group. For example, the training accuracy rate (100 %) satisfied both proportional by-chance criterion (100 % > 1.25 * 33.78 % = 42.23 %) and maximum by-chance criterion (100 % > 1.25 * 50 = 62.50 %). The significance of the Chi-Square statistic ($p < 0.0001$) and the determination coefficient (Nagelkerke's $R^2 = 100.00$ %) show a very strong relationship between class variable and the input variables.

We interpret the results of MLR by looking at the SPSS output tables. According to "Likelihood Ratio" test, all variables are statistically significant (sig. < 0.05) which gives the evidence that all three independent variables contribute significantly to explaining differences in classification. Some coefficients in the regression equations are not statistically significant (Wald test). Some values in "Std. Error" column are greater than 2, which indicate a multicolinearity problem for our NFIs' performance dataset. Variable "ROA" has a value of 1.21 in column "Exp (B)" for the 2nd regression equation, which means that for each unit increase in this variable the likelihood that the observations will be classified in class B increases by approximately 1.20 times. Next, we validate our models based on the test data using the general procedure described in Sect. 5.2 from Costea (2005). The results are presented in Table 3.

The results of MLR classification technique are rather poor for this experiment. First of all, there are many regression coefficients that are statistically insignificant, due to high standard errors obtained for most of them. Secondly, the MLR models tend to over fit the training data. We obtained 100 % accuracy rates for all three training sessions: one with the entire dataset as training set, the second with half of the observations considered for training (split = 0) and the third with the other half of the observations considered as training instances (split = 1). In these two last cases we used the other half of the instances as test sample. There are major discrepancies between the training and test accuracy rates. More robustness in collecting and preprocessing the data is necessary in order for the classification model to be accurate and useful. In the future work we will handle the multicolinearity problem by adding new training data and more input variables. Also, we will check different preprocessing methods once we have the updated dataset.

Table 3 Accuracy rate validations for the financial MLR classification models. The validation is done according to Sect. 5.2 in Costea (2005)

	Main dataset	Part1 (split = 0)	Part2 (split = 1)
Learning sample	100.00 %	100.00 %	100.00 %
Test sample	No test sample	77.27 %	81.81 %

5 Conclusions

In this paper we presented how Data Mining techniques, namely Self-Organizing Map (SOM) algorithm and Multinomial Logistic Regression (MLR) can be used in performing financial performance benchmarking of different non-banking financial institutions in Romania. We selected only those NFIs that are registered in the Special register, have as main activity financial leasing and have been active since the introduction of the regulatory framework for these institutions in Romania.

We trained several SOMs and selected the best one in terms of quantisation error and ease-of-readability. We validated the map dimensionalities and quantisation error using different training and testing datasets and bootstrap technique. We could not reject any SOM architecture for a given significance level and we chose the dimensionalities of the map with the smallest quatisation error. Although we did not find significant differences for the quatisation errors, based on our empirical procedure we are now able to find the optimal training–testing dataset combination for a particular problem. The final map was used to analyze over time the largest three companies in terms of total assets by studying the cluster where each company was positioned for each period. As a main pattern, we can see that for the analyzed companies there was a sharp drop in their performance in 2009 as compared to 2008. This coincides with the effect of the global financial crisis that materialized in Romania during year 2009 and hardly hit the auto sales industry which in turn affected negatively the performance of the NFIs that engaged in financing this sector.

We obtained a perfect classification in terms of training accuracy rates for all three training sessions, but rather high differences between training and testing accuracy rates. This might be due to the small number of training observations and a possible problem of multicolinearity among input variables. New experiments using other methods to preprocessing the data and adding new observations/input variables to the NFIs' financial performance dataset might yield better results.

This type of analysis can benefit the NFIs involved, Supervision Department from the National Bank of Romania in its monitoring process, business players such as international companies that want to expand their business and individual investors. Using our models, investors would be able to weigh the different investment opportunities by performing the comparisons themselves.

Acknowledgments This work was supported from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007–2013, project number POSDRU/89/1.5/S/59184 “Performance and excellence in postdoctoral research in Romanian economics science domain”. The author would like to thank an anonymous reviewer for his/her constructive comments.

References

Alhoniemi E, Hollmen J, Simula O, Vesanto J (1999) Process monitoring and modeling using the self-organizing map. Integr Comput-Aided Eng 6(1):3–14

Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J Finance 23:589–609

Anderson E (1935) The irises of the Gaspe’ peninsula. Bull Am Iris Soc 59:2–5

Back B, Sere K, Vanharantam H (1996) Data mining accounting numbers using self organising maps. In: Alander J, Honkela T, Jakobsson M (eds) Proceedings of finnish artificial intelligence society conference, Vaasa, 20–23 Aug 1996, pp 35–47

Back B, Sere K, Vanharanta H (1998) Managing complexity in large databases using self-organising maps. Account Manage Inf Technol 8(4):191–210

Beaver WH (1966) Financial ratios as predictors of failure, empirical research in accounting: selected studies. Suppl J Account Res 4:71–111

Boyacioglu MA, Kara Y, Baykan OK (2009) Predicting financial bank failures using neural networks, support vector machines and multivariate statistical methods: a comparative analysis of samples in the savings deposit insurance fund (SDIF) transferred banks in Turkey. Expert Syst Appl 36(2):3355–3366

Cerna S, Donath L, Seulean V, Herbei M, Bărglăzan D, Albulescu C, Boldea B (2008) Financial stability. West University Publishing House, Timişoara

Collier C, Forbush S, Nuxoll D, O’Keefe J (2003) The SCOR system of off-site monitoring: its objectives, functioning, and performance. FDIC Bank Rev 15(3):17–32

Costea A (2005) Computational intelligence methods for quantitative data mining. Turku Centre for Computer Science, Ph.D. thesis, No. 67, Painosalama Oy, Turku, November 2005

Costea A (2011) Assessing the performance of non-banking financial institutions – a knowledge discovery approach. Ann Univ Craiova Econ Ser 39:174–185

De Andres J (2001) Statistical techniques vs. SEES algorithm. An application to a small business environment. Int J Digit Account Res 1(2):153–178

De Bodt E, Cottrell M, Verleysen M (2002) Statistical tools to assess the reliability of self-organizing maps. Neural Netw 15:967–978

Doumpos M, Zopounidis C (2009) A multicriteria bank rating system. Eur Work Group “Mult Criteria Decis Aiding” 3(19):17–19

Edmister RO (1972) An empirical test of financial ratio analysis for small business failure prediction. J Financ Quant Anal 7:1477–1493

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery: an overview. In: Fayyad U, Piatetsky-Shapiro G, Smyth P, Uthurusamy R (eds) Advances in knowledge discovery and data mining. AAAI Press, Menlo Park, pp 1–30

Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7:179–188

Garson GD (2005) PA 765 statnotes: an online textbook. Work in progress. http://www2.chass.ncsu.edu/garson/pa765/statnote.htm. Accessed 15 Sept 2012

Hair JF, Anderson-Jr R, Tatham RL (1987) Multivariate data analysis with readings, 2nd edn. Macmillan, New York

Hamer M (1983) Failure prediction: sensitivity of classification accuracy to alternative statistical method and variable sets. J Account Public Policy 2(Winter):289–307

Hosmer DW, Lemeshow S (2000) Applied logistic regression, 2nd edn. Wiley, New York

International Monetary Fund (2010) Romania: financial sector stability assessment. IMF country report no 10/47

Jones F (1987) Current techniques in bankruptcy prediction. J Account Lit 6:131–164

Kohonen T (1997) Self-organizing maps, 2nd edn. Springer, Heidelberg

Martín-del-Brío B, Serrano Cinca C (1993) Self organising neural networks for the analysis and representation of data: some financial cases. Neural Comput Appl 1(2):193–206

Ohlson JA (1980) Financial ratios and the probabilistic prediction of bankruptcy. J Account Res 18 (1):109–131

Oja M, Kaski S, Kohonen T (2003) Bibliography of self-organizing map (SOM) papers: 1998–2001 addendum. Neural Comput Surv 3:1–156

Ravi Kumar P, Ravi V (2007) Banks and firms in bankruptcy prediction via intelligent and statistical techniques – a review. Eur J Oper Res 180(1):1–28

Rudolfer S, Paliouras G, Peers I (1999) A comparison of logistic regression to decision tree induction in the diagnosis of carpal tunnel syndrome. Comput Biomed Res 32:391–414

Şerban F, Ştefănescu MV, Ferrara M (2011) Portfolio optimization and building of its efficient frontier. Econ Comput Econ Cybern Stud Res 2:125–137, ISSN: 0424 – 267X

Serrano Cinca C (1996) Self organising neural networks for financial diagnosis. Decis Support Syst 17:227–238

Serrano Cinca C (1998a) Self-organising maps for initial data analysis: let financial data speak for themselves. In: Deboeck G, Kohonen T (eds) Visual intelligence in finance using self-organising maps. Springer-Verlag, London

Serrano Cinca C (1998b) From financial information to strategic groups – a self organising neural network approach. J Forecast 17:415–428

Swicegood P, Clark JA (2001) Off-site monitoring systems for predicting bank underperformance: a comparison of neural networks, discriminant analysis, and human professional judgement. Intell Syst Account Finance Manage 10(3):169–186

Ultsch A (1993) Self organized feature maps for monitoring and knowledge acquisition of a chemical process. Gielen S, Kappen B (eds) Proceedings of the international conference on artificial neural networks (ICANN93), Springer, London, pp 864–867

Zavgren C (1985) Assessing the vulnerability to failure of American industrial firms: a logistics analysis. J Bus Financ Account (Spring) 12(1):19–45