



## AN INTEGRATED TWO-STAGE METHODOLOGY FOR OPTIMISING THE ACCURACY OF PERFORMANCE CLASSIFICATION MODELS

Adrian COSTEA<sup>a</sup>, Massimiliano FERRARA<sup>b</sup>, Florentin ȘERBAN<sup>c</sup>

<sup>a</sup>*Department of Statistics and Econometrics and Data Mining Group Laboratory, Faculty of Cybernetics, Statistics and Informatics in Economy, Bucharest University of Economic Studies, 15-17 Calea Dorobanti St., “Ion Angelescu” Building, 6<sup>th</sup> Floor, Room 6207, Sector 1, Bucharest, Romania, 010552*

<sup>b</sup>*CRIOS – Centre for Research on Innovation, Organization and Strategy, Department of Management and Technology, Bocconi University and Department of Law and Economics, University Mediterranea of Reggio Calabria, Via dei Bianchi, 2 – Palazzo Zani – 89127 Reggio Calabria, Italy*

<sup>c</sup>*Department of Applied Mathematics, Faculty of Cybernetics, Statistics and Informatics in Economy, Bucharest University of Economic Studies, 15-17 Calea Dorobanti St., “Ion Angelescu” Building, 6<sup>th</sup> Floor, Room 2625, Sector 1, Bucharest, Romania, 010552*

<sup>d</sup>*Doctoral School of Mathematics, University of Bucharest, 14 Academiei St., Sector 1, Bucharest, Romania, 010014*

Received 18 February 2015; accepted 22 May 2016

**Abstract.** In this paper we propose a two-stage methodology to classify the non-banking financial institutions (NFIs) based on their financial performance. The first stage of the methodology consists of grouping the companies in similar financial performance classes (e.g.: “good”, “average”, “poor” performance classes). We optimise the allocation of the observations within the performance clusters by applying an enhanced version of an observation re-allocation procedure proposed in our previous work. Next, based on the result of the grouping phase, we construct a performance class variable by attaching a performance label to each data row. Then, in the second phase of our methodology, we propose a feed-forward neural-network classification model that maps the input space to the newly-constructed performance class variable. This model allows us to forecast the performance of new companies as data become available.

**Keywords:** knowledge-based systems, uncertainty modelling, applications of fuzzy sets, classification, artificial intelligence, performance evaluation, non-banking financial institutions.

**JEL Classification:** C38, C45, D81, G23.

## Introduction

Assessing comparatively the financial performance of the entities within an industry sector is beneficial for all parties on that business stage. In particular, central banks are interested in financial performance classification models that would be used to identify those entities that present a poor financial performance in order to supervising them more closely and maintaining the financial stability of the entire industry. Moreover, the time and personnel resources would be better allocated by targeting entities with deteriorating performance. The Uniform Assessment System or CAAMPL (Cerna *et al.* 2008), was developed in Romania, by the Supervision Department at National Bank of Romania (NBR). It functions as an efficient tool for assessing the performance of credit institutions. Except credit institutions, the NBR prudentially supervises non-banking financial institutions (NFIs) which are financial institutions, other than banks, that engage in lending activities such as granting of credits, financial leasing, issuance of guarantees, etc. One main difference between credit institutions and NFIs is that NFIs do not attract retail deposits and, as such, their concern is more related to their solvency (long-term liquidity requirements) and not short-term liquidity, since they do not face “break the buck” event that could lead to a run. The NFIs should not be confused with non-bank institutions which operate mainly in the capital markets and, at least in Romania, are prudentially supervised by the Financial Supervisory Authority, the regulator and supervisor of capital markets, insurance companies and pension funds. The CAAMPL system above examines the financial reports of credit institutions to evaluate six components which reflect in a comprehensive and consistent manner the performance of banks, to ensure the concordance with existing banking laws and regulations: capital adequacy (C), quality of ownership (A), assets’ quality (A), management (M), profitability (P), liquidity (L). However this system is not applicable for performance evaluation of non-banking financial institutions (NFIs). The CAAMPL system presents a number of disadvantages related to its application in evaluating the performance of banks and, possibly, NFIs:

- it uses simple linear techniques to discriminate the  $p$ -multidimensional input space that represents the performance of financial institutions. Actually, the discrimination model is a sequential combination of univariate models (i.e.: a model that takes into consideration one discriminating variable at a time);
- the selection of input variables (financial performance ratios) that determine a specific performance class is based on the experience of members of the supervisory authority, and not on scientific rigor (e.g., it is not based on the validity and reliability of financial ratios in international comparisons);
- the heuristic selection of the financial performance ratios leads to difficulties in substantiating their limits which increase the subjective involvement of the analysts in establishing them;
- does not imply any form of quantitative methods for assessing the performance, such as statistical and econometric methods and/or the so-called computational-intelligence (CI) methods, being based mainly “on rules” as it was emphasized by the IMF in IMF (2010). The CI methods come from different research fields, such as fuzzy logic, machine learning and artificial intelligence.

In this paper, we try to overcome the above disadvantages as follows:

- we use multivariate models to discriminate the input space: firstly, we apply a clustering technique to group the NFIs as to their financial performance and, then, we build a classification model that maps the input space to a newly constructed class variable obtained in the clustering step. As such, our methodology consists of two steps: the clustering and the classification steps;
- the selection of input variables for each performance dimension is based on scientific rigor: it is based on Lehtinen's study (Lehtinen 1996) about the validity and reliability of financial ratios in international comparisons. The data are collected with the help of the personnel of the supervisory authority;
- the limits that define a performance ratio as being low, high, average, etc., are constructed based on the values of the ratio using fuzzy numbers. Each linguistic term (low, average, high, etc.) is modeled using a trapezoidal fuzzy number. In this way we eliminate the subjective approach of the CAAMPL system in establishing these limits. Moreover, these fuzzy numbers allow us to select relevant variables for each performance cluster (class) and to characterize linguistically each performance class;
- our approach in assessing NFIs' performance resorts to CI methods (Fuzzy C-Means for the clustering step and artificial intelligence techniques for the classification step), therefore overcoming the shortcomings of CAAMPL system which is based "on rules".

Recently, a lot of models for assessing the financial performance of entities were proposed in the literature. These models apply, in particular, to credit institutions.

Erdogan (2013) applies Support Vector Machines (SVMs) with Gaussian kernel for bankruptcy prediction of banks. The author uses a dataset from Turkish commercial banks and finds out that the SVMs are capable of extracting useful information from financial data, possibly as a part of an early warning system.

Abad González and Gutiérrez López (2015) used logit and regression models in order to predict the bank's capital shortfall/surplus based on a number of financial ratios taken from the financial statements. The results of the predictive models are compared with the results obtained based on the stress tests that were conducted both in European Union in 2011 and in Spain in 2012. The dataset consisted of 23 Spanish credit institutions that participated in two stress test exercises, one that analyzed the resilience of each bank individually and another one that analyzed the financial resilience of the entire banking sector. The authors argue that their models that included equity-to-debt ratio as an explanatory variable provided the best result in terms of goodness-of-fit in both scenarios.

A new classification system of the financial institutions as a support-tool for analysts from the National Bank of Greece was proposed by Doumpos and Zopounidis (2009). The system developed by these authors provides a large set of assessment, visualization and reporting options.

Several methods for classifying credit institutions based on 20 performance indicators grouped into six dimensions (CAMELS) were proposed by Boyacioglu *et al.* (2009). They used four financial data sets and found out that among the clustering/classification techniques which were tested, the best in terms of accuracy rates are neural networks. Șerban *et al.* (2012) and Ștefănescu *et al.* (2008) developed classification algorithms with applications to portfolio management. Toma and Dedu (2014) and Tudor and Dedu (2012) developed qu-

antitative techniques for financial risk assessment. Ionita and Stancu (2015) used regression techniques in order to find relevant leading indicators of the debt crisis in Romania using quarterly data from first quarter 1999 to fourth quarter 2013. The authors have constructed an early-warning index of debt crisis by identifying the most statistically significant macroeconomic indicators. Cimpoeru (2015) used neural networks in the form of Self-Organising Maps algorithm in order to detect the economic differences among world economies and find specific early-warning signals for a financial/economic crisis. The author collected the data, among others, from the World Bank database and, according to Iturriaga and Sanz (2013), she used 15 macroeconomic variables as leading indicators for the crisis. She found out that the developed economies were as vulnerable as the emerging ones in terms of macroeconomic indicators performance prior to the beginning of the global financial crisis in 2007.

A consistent literature review regarding the research on the application of statistical and computational intelligence methods on banks' and firms' bankruptcy prediction problem during 1968–2005 is presented in Kumar and Ravi (2007). They provide for each study the data source, the financial ratios used, the country of origin and the data collection period.

Pivk *et al.* (2013) apply Data Mining techniques in order to improve business processes in banking, telecom and retail sectors. While still requiring a lot of coordination and manual adjustment, the proposed Data Mining solution is successfully applied in eight commercial companies. Hajek *et al.* (2014) use annual reports' sentiment to predict corporate financial performance. The authors employ support vector machines and neural networks in order to forecast the financial performance. While support vector machines provided the best accuracy, the authors argue that there is a non-linear relationship between the sentiment and financial performance.

Currently, unlike the case of credit institutions, at the level of the Supervision Department, there are no performance classification models for the NFIs' sector. The NFIs performance evaluation is done manually by consulting their prudential reporting, such as quarterly periodic financial statements (PFSs). A number of raw indicators for NFIs' performance are periodically recorded into financial statements and manually analyzed by inspectors. A powerful comparative analysis of several NFIs or a dynamic analysis of one of these NFIs based on the indicators of the PFSs require considerable effort from the inspectors of the Supervision Department. This problem is caused by the complexity of the problem involving the dynamic analysis, for a large number of quarters, of all prudentially supervised NFIs (around 65 in the data collection period) in terms of a set of many (10 to 20) performance ratios. Therefore, there is a need for new models to assess comparatively the performance of financial institutions, in general, and the performance of non-banking financial institutions, in particular.

The objective of this paper is two-fold: firstly, it fills the gap in the literature by proposing a two-stage methodology for developing performance classification models for the NFIs; secondly, unlike the traditional performance assessment systems (e.g. the CAAMPL system), the proposed methodology uses CI methods for developing these models and it discriminates the input space into more than two classes, gaining more insights into the dataset.

In conducting our research we have followed Hevner *et al.* (2004) guidelines to understand how to perform effective design science (constructive) research, one of which (sixth guideline) is to perform the research as a search process: test a *feasible* number of solutions for a given problem, choosing the best one in terms of some qualitative and/or quantitative

criteria. We investigated different methods from different fields (artificial intelligence, fuzzy logic, statistics, machine learning) for assessing the NFIs' performance in Romania. In this paper we have combined two such methods, one from the field of fuzzy logic (Fuzzy C-Means clustering) and one from the field of artificial intelligence (feed-forward artificial neural networks training algorithms) in order to build a classification model for the NFIs' performance. In this sense, the current work is complementary to the previously proposed models: fuzzy logic techniques in Costea and Bleotu (2012), feed-forward neural network training techniques in terms of backpropagation (BP) and BP-variants algorithms in Costea (2012b) and artificial neural networks training techniques in terms of Self-organising Maps (SOM) algorithm in Costea (2013a). In Costea (2013a), the author used SOM algorithm to find clusters of NFIs with similar financial performance. The dataset consisted of 3 financial performance ratios for 11 NFIs. The data were collected annually for 2007 to 2010, in total 44 observations. In Costea and Bleotu (2012), the authors used Fuzzy C-Means algorithm to evaluate the performance of NFIs in Romania. The proposed model outperformed the SOM model both in terms of pattern allocation and cluster characterization. The authors extended the experiment by including in their analysis 65 NFIs and 8 financial ratios. The data were collected quarterly between 2007 and 2010, obtaining a total of 769 observations. Costea (2012b) extended the NFIs' performance evaluation experiment by collecting quarterly data for 68 NFIs for the 2007 to 2012 period, totalling a dataset of 990 observations. At the same time, the author increased the number of financial ratios to 11, adding new information for benchmarking the NFIs. Moreover, in Costea (2012b), the author used feed-forward artificial neural networks to build the NFIs' performance classification models. The networks were trained using gradient-descent-like algorithms, such as BP and BP variants.

In this paper, we combine the descriptive clustering (in terms of a modified fuzzy clustering algorithm) with a prescriptive classifier (in terms of artificial neural network).

The paper is organized as follows: in the next section we present our two-stage methodology and the implications of parameter settings based on what has been reported in the literature on using the proposed techniques for similar problems. Then, we give a brief description of the NFIs' performance dataset and present our experiments. In the concluding section we discuss the empirical validation of our objectives.

## **1. The two-stage methodology for NFIs' performance classification**

Our methodology consists of two stages. In the first stage we apply a clustering technique, namely Fuzzy C-Means (FCM) algorithm, in order to find performance clusters within the data for the NFIs' sector. According to Costea and Bleotu (2012), FCM algorithm performed better than other clustering algorithms in terms of the formed clusters when applied on a similar performance dataset. At this stage, each cluster is described in terms of average characteristics of the observations allocated to it and a label which stands for the performance class variable is attached to each observation. This label identifies the observation as belonging to the cluster. In the second stage we apply feed-forward artificial neural networks (ANNs) algorithms in order to map the input space to the newly-created performance class variable so that we might be able to forecast the performance of different NFIs as data become available. Next, we present the proposed algorithms in greater detail.

### 1.1. First stage of the methodology: forming the performance clusters

There are many clustering algorithms that could be used to group the NFIs' in performance clusters. However, in our previous studies (e.g., Costea 2005; Costea, Bleotu 2012) we showed that the Fuzzy C-Means (FCM) algorithm outperformed other algorithms, such as the Self-Organising Maps (SOM) algorithm (Kohonen 1997), in terms of both allocation of observations and cluster characterization. In comparing the performance of these clustering algorithms we used as criterion the number of coincidences in linguistic terms of each observation with the linguistic terms that characterize the cluster to which that observation belongs. The higher the number of coincidences in the linguistic terms the better the clustering algorithm. Recently, some authors proposed multiple-criteria decision-making methods for evaluating the clustering algorithms (e.g., Kou *et al.* 2014).

Using fuzzy clustering and, in particular, the Fuzzy C-Means (FCM) algorithm for a comparative assessment of the financial performance of companies is relatively scarce. The fuzzy logic based approach can handle multi-dimensional data and model non-linear relationships among input variables. This approach has been applied for performing the financial analysis of companies. For example, Lindholm and Liu (2003) used it for evaluating early-warning indicators of financial crises and Drobics *et al.* (2000) employed it for developing fuzzy rules out of a clustering obtained with self-organizing maps algorithms. A model for selecting the suppliers based on fuzzy method (TOPSIS) was described by Wang *et al.* (2009). In Stanujkic (2013), the author proposes an extension of the MOORA method by using triangular fuzzy numbers. The author discusses different methods for defuzzifying and calculating the distance between two fuzzy numbers and argues that the MOORA method can be used in fuzzy environment given the discussed methods are used. The MULTIMOORA-2T (Multi-Objective Optimization by Ratio Analysis plus the Full Multiplicative Form – Two Tuples) method for group multi-criteria decision making under linguistic environment was extended by Baležentis, A. and Baležentis, T. (2011). They used two-tuples to represent, convert and map various crisp and fuzzy numbers into the basic linguistic term set.

Traditional clustering methods focus on identifying patterns in the dataset and creating partitions with different structures, called clusters (Jain *et al.* 1999). Each cluster contains elements with similar characteristics. Generally, every element belongs to only one cluster, but there may be observations in the dataset which are difficult to be assigned to a certain cluster. In order to allocate these observations, subjective decisions have to be made.

As a different approach, fuzzy clustering methods assign different membership degrees to the elements in the dataset, that indicate in which degree the observation belongs to every cluster. Fuzzy C-Means (FCM) clustering method represents a traditional fuzzy clustering method (Bezdek 1981). The FCM method minimizes the following objective function,  $F(n, C, m)$ :

$$F(n, C, m) = \sum_{i=1}^n \sum_{j=1}^C (u_{ij})^m (D_{ij})^2, \quad (1)$$

where  $n$  is the number of observations in the dataset,  $C$  is the number of clusters,  $m$  is a parameter that controls the extent of membership sharing between the clusters,  $u_{ij}$  and  $D_{ij}$  are the membership degree and Euclidean distance, respectively, of observation  $i$  to cluster  $j$ .

The membership degrees  $u_{ij}$  are randomly initialized so that two conditions hold: (1)  $u_{ij} > 0$  and (2)  $\sum_{j=1}^C u_{ij} = 1$ . The Euclidean distance  $D_{ij}$  is calculated using the following formula:

$$D_{ij} = \sqrt{\sum_{k=1}^p (O_{ik} - X_{jk})^2}, \tag{2}$$

where  $p$  is the number of input variables,  $O_i$  represents the  $i^{\text{th}}$  observation and  $X_j$  is the center of cluster  $j$ .

The FCM method minimizes  $F$  based on the Lagrange multipliers method. Firstly, the Lagrangian is constructed taking into account the function to minimize and the equality constraint for the membership degrees:

$$\Lambda(n, C, m, \lambda) = \sum_{i=1}^n \sum_{j=1}^C (u_{ij})^m \sum_{k=1}^p (O_{ik} - X_{jk})^2 - \sum_{i=1}^n \lambda_i \left( \sum_{j=1}^C u_{ij} - 1 \right). \tag{3}$$

Then, in order to minimize the Lagrangian, partial derivatives are equalized to zero:

$$\frac{\partial \Lambda(n, C, m, \lambda)}{\partial u_{ij}} = 0 \rightarrow m(u_{ij})^{m-1} (D_{ij})^2 - \lambda_i = 0 \rightarrow u_{ij} = \left[ \frac{\lambda_i}{m(D_{ij})^2} \right]^{1/(m-1)}; \tag{4}$$

$$\frac{\partial \Lambda(n, C, m, \lambda)}{\partial \lambda_i} = 0 \rightarrow \sum_{j=1}^C u_{ij} - 1 = 0 \rightarrow \sum_{j=1}^C u_{ij} = 1. \tag{5}$$

If we sum up the membership degrees in Eq. (4) from 1 to  $C$  and we equal the result to 1 according to Eq. (5), we obtain:

$$\left( \frac{\lambda_i}{m} \right)^{1/(m-1)} \sum_{r=1}^C \left[ \frac{1}{(D_{ir})^2} \right]^{1/(m-1)} = 1 \rightarrow \left( \frac{\lambda_i}{m} \right)^{1/(m-1)} = \left\{ \sum_{r=1}^C \left[ \frac{1}{(D_{ir})^2} \right]^{1/(m-1)} \right\}^{-1}. \tag{6}$$

By introducing Eq. (6) in Eq. (4), we derive the formula for calculating the membership degrees ( $u_{ij}$ ):

$$u_{ij} = \left[ \frac{\lambda_i}{m(D_{ij})^2} \right]^{1/(m-1)} = \frac{\left[ \frac{1}{(D_{ij})^2} \right]^{1/(m-1)}}{\sum_{r=1}^C \left[ \frac{1}{(D_{ir})^2} \right]^{1/(m-1)}} = \left[ \sum_{r=1}^C \left( \frac{D_{ij}}{D_{ir}} \right)^{2/(m-1)} \right]^{-1}. \tag{7}$$

In order to derive the clusters' centers ( $X_j$ ), we equalize to zero the partial derivative:

$$\frac{\partial \Lambda(n, C, m, \lambda)}{\partial X_{jk}} = 0 \rightarrow -2 \sum_{i=1}^n (u_{ij})^m (O_{ik} - X_{jk}) = 0 \rightarrow X_{jk} = \frac{\sum_{i=1}^n (u_{ij})^m O_{ik}}{\sum_{i=1}^n (u_{ij})^m}. \tag{8}$$

When  $m$  approaches 1, the Fuzzy C-Means algorithm is identical to the classical C-Means algorithm (the Hard C-Means - HCM - algorithm). The partition becomes fuzzier when

the value of the parameter  $m$  increases. When  $m \rightarrow \infty$ , then  $u_{ij} \rightarrow 1/C$  and the centers approach the centroid of the dataset, so the centers tend to be equal. Except the number of clusters  $C$  and  $m$ , the FCM method has two more parameters:  $N$  which represents the maximum algorithm iterations and  $\varepsilon$  which is a small number that is used as a stopping criterion for the algorithm.

Given the above parameters ( $C$ ,  $m$ ,  $N$  and  $\varepsilon$ ), the algorithm can be described as follows:

1. Initialize randomly the membership degrees  $U^{(0)} = \{u_{ij}\}$ , so that:  $u_{ij} > 0$  and  $\sum_{j=1}^C u_{ij} = 1$ ;
2. Then, for each iteration  $s, s = \overline{0, N}$ , repeat the following steps:
  - a. Allocate the observations into the  $C$  clusters according to  $U^{(s)}$  and calculate the clusters' centers  $\left( X_j^{(s)} \right)$  with Eq. (8) and using  $U^{(s)}$ ;
  - b. Calculate the new membership degrees ( $U^{(s+1)}$ ) with Eq. (7) and using  $X_j^{(s)}$ ;
  - c. Stop if  $U^{(s+1)} - U^{(s)} \leq \varepsilon$  or  $s = N$ .

The output of the FCM algorithm is final matrix  $U$ : a set of membership degrees ( $u_i$ ) is assigned to each observation, which indicates that observations may contain, with different strengths, characteristics of more than one cluster. In this case, the FCM method assigns each observation to the cluster corresponding to the highest membership degree. In spite of the additional information provided by the FCM method, there is a problem with the observations for which similar two (or more) highest membership degrees are obtained. We label these observations as being "uncertain". We consider that an observation is uncertain if the difference between the two highest membership degrees is less than  $1/C$ . We chose this upper limit due to the fact that a *perfectly* uncertain observation would have all membership degrees equal to  $1/C$  (the observation would belong equally to each cluster). These observations are uncertain in the sense that there is an uncertainty regarding their allocation into a particular cluster. The other observations (which do not satisfy the above condition) are considered "certain".

Next, based on our previous work (Costea 2005; Costea, Bleotu 2012), we propose a procedure to identifying and re-allocating the uncertain observations by introducing some weights to the FCM algorithm. These weights are calculated based on the some linguistic variables that we define. For each financial variable we define a linguistic variable using five linguistic terms: *very low* (VL), *low* (L), *average* (A), *high* (H), *very high* (VH). Each linguistic term is represented as a trapezoidal fuzzy number. For each variable and for each linguistic term, we define the trapezoidal fuzzy number following the procedure in Costea (2005). Basically, for each variable, we apply the FCM algorithm and group the values of that variable in five fuzzy classes (e.g., a specific value for a particular ratio is *very low*, *low*, *average*, *high* or *very high* with different membership degrees). We define the linguistic terms as follows (see Fig. 1):

- the linguistic term VL is defined by three points: a minimum point (A), a maximum point (B) and the minimum point for the linguistic term L (C);
- the linguistic terms L, A, H are defined by four points: the maximum point for the previous linguistic term (e.g., point B for the linguistic term L), a minimum point (e.g., point C for the linguistic term L), a maximum point (e.g., point D for the linguistic term L) and the minimum point for the next linguistic term (e.g., point E for the linguistic term L);

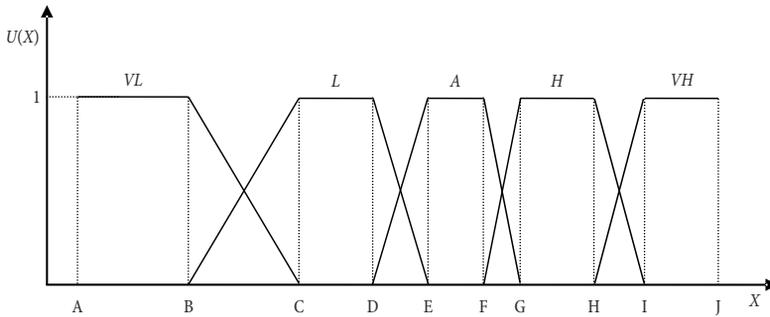


Fig. 1. The trapezoidal representation of the five linguistic terms (VL, L, A, H, VH) for a generic variable X

- the linguistic term *VH* is defined by three points: the maximum point for the linguistic term *H* (H), a minimum point (I) and a maximum point (J).

In other words, in order to define all of the five fuzzy numbers, we need to define a minimum and a maximum point for each linguistic term. The minimum/maximum point for the linguistic term *LT* is defined as the minimum/maximum value for that ratio in the entire data set for which the membership degree in the class of linguistic term *LT* is greater or equal to 0.99. For each ratio and for each linguistic term, the minimum point is initialized to *+inf* and the maximum point is initialized to *-inf*. It is possible that there is no observation that has a membership degree greater than or equal to 0.99, even if this case is unlikely. However, 0.99 is a parameter for our model so that it can be changed to accommodate highly heterogeneous data.

In the literature there are other methods for generating the linguistic variables, including the approach presented in Massanet *et al.* (2014), where, instead of using continuous membership functions, the authors use discrete fuzzy numbers to define more complex experts' qualitative expressions, such as: “between average and high”. Another approach of dealing with these expressions and more complex ones was introduced in Rodriguez *et al.* (2012) where hesitant fuzzy linguistic terms are used to model decision making problems. In order to overcome the limitations of the use of fuzzy linguistic approach regarding information modelling and computational processes, researchers proposed several linguistic models (e.g., Herrera *et al.* 1994, 2008). The difference here is that we do not work with a panel of experts whose opinions we are trying to reconcile, but, rather, we transform our financial ratios (quantitative data) in a qualitative form building linguistic variables. We use the linguistic variables to construct the weights ( $w_{jk}$ ), based only on the certain observations. A weight ( $w_{jk}$ ) is the standardized variation coefficient of the vector  $NOLT_{jk}$ , where  $NOLT_{jk}$  is a vector calculated as follows:

$$NOLT_{jk}(i) = \frac{\text{number of occurrences of linguistic term } i \text{ for ratio } k \text{ in cluster } j}{\text{number of observations in cluster } j}, \quad (9)$$

where  $i = VL, L, A, H, VH$ .

The variations coefficients and the weights are given below:

$$V_{jk} = \frac{\text{standard\_deviation}(NOLT_{jk})}{\text{average}(NOLT_{jk})}; \quad (10)$$

$$w_{jk} = \frac{V_{jk}}{\sum_{k=1}^p V_{jk}}. \quad (11)$$

Clearly, for one particular ratio  $k$  the weights sum up to 1. A high weight for a particular cluster  $j$  and a particular ratio  $k$  would mean that the ratio is a good definer for that cluster (it concentrates its values in few linguistic terms). Conversely, a low weight indicates that the ratio is not a good definer for the cluster. In our experiments we stated that one linguistic term characterizes one ratio for one cluster if it represents at least 40% from the total number of values corresponding to that ratio in that cluster. We have chosen 40% in order to allow maximum two linguistic terms to characterize a cluster for each ratio (any percentage greater than 33.33% would have the same outcome). Using this rule, we can characterize each cluster as having, for example,  $VL$  values for a particular ratio,  $H$  and  $VH$  values for another ratio, etc, depending on which linguistic term(s) bypass the 40% hurdle. Moreover, based on the above characterization of the clusters and studying the financial ratios in each cluster, we can order them as being “good”, “average” and “bad” performance clusters. It is possible that, for one specific variable, no linguistic term bypasses the 40% hurdle (e.g., see in Table 3 the *Costs to revenues* financial ratio for Cluster 3). In this case we say that the variable is not a good definer for that particular cluster. The granularity of the linguistic terms should be set carefully: the higher the granularity of the linguistic term (e.g., 7 or 9 instead of 5), the smaller the probability that the linguistic terms bypass the hurdle, but the richer the information we save from the original data, while a small granularity (e.g., 3 or 4 instead of 5) would increase the chance that one linguistic term bypasses the hurdle and, at the same time, would limit the information that is saved from the original data. In any case, the granularity of the linguistic terms is a parameter for our models. We chose to set a moderate level of granularity for the linguistic terms (5) instead a higher one (e.g., 7 or 9), because, after we generate the qualitative variables, we are interested in characterizing each cluster as being a “good”, “average” or “bad” performance clusters and a higher granularity would prevent us to do so by generating many variables that are not good definers for the clusters: we prefer that linguistic terms bypass the hurdle, at the expense of loosing some information from the original data.

The above weights ( $w_{jk}$ ) are introduced in the new objective function by changing the distances, as follows:

$$D_{ij}^{new} = \sqrt{\sum_{k=1}^p (O_{ik} - X_{jk})^2 w_{jk}} \quad (12)$$

and the objective function to minimize at each iteration  $s$ , becomes:

$$F^{new}(n, C, m) = \sum_{i \in T} \sum_{j=1}^C \left( u_{ij}^{(s)} \right)^m \left( D_{ij}^{new(s)} \right)^2 \left( 1 - u_{ij}^{(s-1)} \right), \quad (13)$$

where  $T$  represents the set of certain observations in iteration  $s$  and  $u_{ij}^{(s-1)}$  stands for the

membership degrees of the certain observations corresponding to cluster  $j$  in the previous iteration. This term is used in order to avoid that lower membership degrees from the uncertain observations become more important in the new allocation. A higher previous membership degree value  $u_{ij}^{(s-1)}$  should lead to a lower (better) recalculated distance from that uncertain observation to the centre of that cluster. Consequently, the  $\left(1 - u_{ij}^{(s-1)}\right)$  factor is used when defining the objective function.

Again, we apply the Lagrange multipliers method in order to find the new formulas for the membership degrees and the clusters' centers, by minimizing the new objective function ( $F^{new}$ ). Firstly, the Lagrangian to minimize is written as follows:

$$\Lambda^{new}(n, C, m, \lambda) = \sum_{i \in T} \sum_{j=1}^C \left(u_{ij}^{(s)}\right)^m \left(1 - u_{ij}^{(s-1)}\right) \sum_{k=1}^p \left(O_{ik} - X_{jk}^{(s)}\right)^2 w_{jk} - \sum_{i \in T} \lambda_i \left(\sum_{j=1}^C u_{ij}^{(s)} - 1\right). \quad (14)$$

The partial derivatives are equalized to zero:

$$\frac{\partial \Lambda^{new}(n, C, m, \lambda)}{\partial u_{ij}^{(s)}} = 0 \rightarrow m \left(u_{ij}^{(s)}\right)^{m-1} \left(D_{ij}^{new(s)}\right)^2 \left(1 - u_{ij}^{(s-1)}\right) - \lambda_i = 0 \rightarrow u_{ij}^{(s)} = \left[ \frac{\lambda_i}{m \left(D_{ij}^{new(s)}\right)^2 \left(1 - u_{ij}^{(s-1)}\right)} \right]^{1/(m-1)}; \quad (15)$$

$$\frac{\partial \Lambda^{new}(n, C, m, \lambda)}{\partial \lambda_i} = 0 \rightarrow \sum_{j=1}^C u_{ij}^{(s)} - 1 = 0 \rightarrow \sum_{j=1}^C u_{ij}^{(s)} = 1. \quad (16)$$

If we sum up the membership degrees in Eq. (15) from 1 to  $C$  and we equal the result to 1 according to Eq. (16), we obtain:

$$\left(\frac{\lambda_i}{m}\right)^{1/(m-1)} \sum_{r=1}^C \left[ \frac{1}{\left(D_{ir}^{new(s)}\right)^2 \left(1 - u_{ir}^{(s-1)}\right)} \right]^{1/(m-1)} = 1 \rightarrow \left(\frac{\lambda_i}{m}\right)^{1/(m-1)} = \left\{ \sum_{r=1}^C \left[ \frac{1}{\left(D_{ir}^{new(s)}\right)^2 \left(1 - u_{ir}^{(s-1)}\right)} \right]^{1/(m-1)} \right\}^{-1}. \quad (17)$$

By introducing Eq. (17) in Eq. (15), we derive the formula for calculating the membership degrees  $\left(u_{ij}^{(s)}\right)$ :

$$u_{ij}^{(s)} = \left[ \frac{\lambda_i}{m \left( D_{ij}^{new(s)} \right)^2 \left( 1 - u_{ij}^{(s-1)} \right)} \right]^{1/(m-1)} = \frac{\left[ \frac{1}{\left( D_{ij}^{new(s)} \right)^2 \left( 1 - u_{ij}^{(s-1)} \right)} \right]^{1/(m-1)}}{\sum_{r=1}^C \left[ \frac{1}{\left( D_{ir}^{new(s)} \right)^2 \left( 1 - u_{ir}^{(s-1)} \right)} \right]^{1/(m-1)}} = \quad (18)$$

$$\left\{ \sum_{r=1}^C \left[ \frac{\left( D_{ij}^{new(s)} \right)^2 \left( 1 - u_{ij}^{(s-1)} \right)}{\left( D_{ir}^{new(s)} \right)^2 \left( 1 - u_{ir}^{(s-1)} \right)} \right]^{1/(m-1)} \right\}^{-1}.$$

In order to derive the clusters' centers  $\left( X_j^{(s)} \right)$ , we equalize to zero the partial derivative:

$$\frac{\partial \Lambda^{new}(n, C, m, \lambda)}{\partial X_{jk}} = 0 \rightarrow -2 \sum_{i \in T} \left( u_{ij}^{(s)} \right)^m \left( 1 - u_{ij}^{(s-1)} \right) \left( O_{ik} - X_{jk}^{(s)} \right) w_{jk} =$$

$$0 \rightarrow X_{jk}^{(s)} = \frac{\sum_{i \in T} \left( u_{ij}^{(s)} \right)^m \left( 1 - u_{ij}^{(s-1)} \right) O_{ik}}{\sum_{i \in T} \left( u_{ij}^{(s)} \right)^m \left( 1 - u_{ij}^{(s-1)} \right)}. \quad (19)$$

Based on the above new formulas for updating the membership degrees and clusters' centers, we refine our approach in Costea (2005) and Costea and Bleotu (2012) and propose the following procedure to identify and reallocate the uncertain observations into the clusters:

1. Given a  $n$  by  $p$  dataset and the other parameters ( $m$ ,  $C$ ,  $N$  and  $\varepsilon$ ), generate the  $n$  by  $p$  dataset of linguistic variables as described above by applying the FCM algorithm individually for each financial ratio;
2. Given the same  $n$  by  $p$  dataset and the other parameters ( $m$ ,  $C$ ,  $N$  and  $\varepsilon$ ), apply the FCM algorithm in order to find the set with certain ( $T$ ) and uncertain ( $T'$ ) observations and a set of membership degrees  $U^{(certain)}$ ;
3. Characterize each obtained cluster using the linguistic variables generated in step 1 and only the set with certain observations ( $T$ );
4. Calculate the weights ( $w_{jk}$ ) using Eq. (11), linguistic variables generated in step 1 and only the set with the certain observations ( $T$ ). These weights will be constant to all of the following iterations;
5. Then, if at least one uncertain observation was reallocated and the set of uncertain observations is not empty ( $T' \neq \emptyset$ ), we perform another iteration  $s$  that consists of following steps:
  - a. Calculate the clusters' centers using Eq. (19) and the sets of membership degrees  $u_{ij}^{(s)}$  and  $u_{ij}^{(s-1)}$  corresponding to the certain observations of the current and previous iterations, respectively. For the first iteration ( $s = 1$ ), we have  $u_{ij}^{(1)} = U^{(certain)}$  (obtained in step 2) and  $u_{ij}^{(0)} = 0$ ;

- b. For the uncertain observations ( $O_i \in T'$ ), calculate the new membership degrees  $u_{ij}^{(s+1)}$  using Eq. (18), the clusters' centers obtained in step (a) and the previous membership degrees  $u_{ij}^{(s)}$ ;
- c. Based on the new membership degrees ( $u_{ij}^{(s+1)}$ ) identify which uncertain observations became certain observations and reallocate them into the clusters. Change, accordingly, the two sets of observations: increase the set with certain observations ( $T$ ) and decrease the set with uncertain ( $T'$ ) observations, and start a new iteration  $s$  from step 5.

The solution of proposed procedure to (re)allocate the observations into performance clusters consists of the set of the final clusters' centers ( $X_j$ ). After we obtain these cluster centers and the linguistic characterization of each cluster, we construct the so-called “performance class” variable by associating a performance class (label) to each observation depending to which cluster it belongs (1-for the cluster labeled “Best”, 2-for the cluster labeled “Average”, 3-for the cluster labeled “Bad” and 4-for the cluster labeled “Worst”). Next, we describe the second stage of our methodology.

## 1.2. Second stage of the methodology: building the performance classification model

In order to map the  $p$ -dimensional input space to the newly-obtained performance class variable we use feed-forward artificial neural networks (ANNs) models. ANNs have been extensively used in financial applications. Koskivaara (2004) used ANNs as classifiers in order to address problems that can arise in the audit process: material error, going concern, financial distress, control risk assessment, management fraud, and audit fee. Costea and Eklund (2004) found out that the ANN performed similarly to statistical and induction techniques in terms of accuracy rates when applied to assessing comparatively the companies in the telecom industry. Other papers (e.g., Swicegood, Clark 2001) compared different performance classification models, such as decision trees models, logistic regression models and ANN-based models and found out that the neural networks models performed the best in terms of training and testing accuracy. However, the explanatory capabilities of the decision trees have to be taken into account in the process of choosing the best model.

A different way of learning the connection weights of an ANN is by using a genetic algorithm. The set of ANN connection weights constitute a solution (a chromosome). Starting with a population of solutions, the genetic algorithm derives potentially better solutions by applying specific operators. Many authors have conducted research in this area. The possible benefits of combining ANNs and evolutionary algorithms (EAs) are examined by Yao (1999). EAs constitute a class of population-based stochastic search algorithms based on principles of natural evolution (Yao 1999). EAs consists of evolution strategies (ESs), evolutionary programming (EP) and genetic algorithms (GAs). Yao offers different combinations between ANNs and EAs such as: the evolution of ANN connection weights, the evolution of ANN architectures and the evolution of ANN learning rules. Based on an extensive literature review, the author proves that combinations of ANNs and EAs can lead to improved models and systems rather than relying on ANNs or EAs alone. Yao (1999) examines a large number of papers where one of the two training mechanisms (evolutionary algorithms and gradient-descent-like algorithms, such as backpropagation – BP – algorithm and its variations)

was found to achieve better results than the other. These contradictory results are attributed by Yao to the type of EA and BP algorithms used: whether the comparison is made between a classical binary GA and a fast BP algorithm, or between a fast EA and a classical BP algorithm. Some authors (e.g., Schaffer 1994) reported that training the ANN using GAs is not as efficient as the classical ANN training (the training that uses gradient-descent algorithms).

Studying the application of a non-binary GA for learning the connection weights of an ANN under various structural design and data distributions, Pendharkar (2002) found out that additive noise, size and data distribution features play an important role in learning, reability and predictive performance of ANNs. The implications of data distributions determined through kurtosis and variance-covariance homogeneity on the predictive performance of GA-based and gradient-descent-based ANN for classification were examined by Pendharkar and Rodger (2004). They also investigated the implication of three types of crossover operator, such as uniform, arihtmetic and one-point crossover, on the prediction performance of GA-based ANN and found no significant difference between these different crossover operators. In Costea and Nastac (2005), the authors used GA-based ANN training to improve the performance of an already trained ANN classification model, but the genetic algorithm was unable to improve the accuracy of the ANN model. This result might be due to the high accuracy rate of the trained network that the genetic algorithm had to improve (in excess of 90%).

The generic classification model based on a neural approach, adapted from Costea and Nastac (2005) is based on some preliminary steps: pre-processing data and separating data into training (*TR*) and test (*TS*) sets. After that, the proper ANN architecture is constructed, by determining the proper number of hidden layers and the appropriate number of neurons in each hidden layer. The class variable is coded using one neuron or a number of neurons equal to the performance clusters. The latter approach would enable us to eliminate the risk that one class, corresponding to a higher value of the output variable (e.g., 4) is more important than classes with lower values (e.g., 1, 2 and 3). Finally, the ANN is trained and tested using different values of the parameters involved in the training process.

Choosing the number of hidden layers and the number of neurons in each hidden layer depends on the size of input/output vector, the size of training and test subsets, and on the problem of non-linearity, as Basheer and Hajmeer (2000) point out. At the same time, these authors present a number of papers that provide different rules of thumb, which refer to the correspondence between the number of hidden neurons (*NH*) and the number of input neurons (*NI*) and output neurons (*NO*) or the number of training samples ( $N_{TRN}$ ). In this way, Lachtermacher and Fuller (1995) relates the number of input neurons (*NI*) and hidden neurons (*NH*) for one output ANN with the number of training samples  $N_{TRN}$  by the formula:  $0.11N_{TRN} < NH(NI+1) < 0.30N_{TRN}$ . Upadhyaya and Eryurek (1992) express the relation between the total number of weights ( $N_w$ ) and the number of training samples using the formula:  $N_w = N_{TRN} \log_2(N_{TRN})$ . Masters (1994) estimates the number of hidden neurons in the hidden layer by the geometric mean of the number of inputs (*NI*) and the number of outputs (*NO*). The choice of these parameters is mostly based on art rather than science, as Basheer and Hajmeer's (2000) highlight that "the most popular approach to finding the optimal number of hidden nodes is by trial and error with one of the above rules". In order

to illustrate this idea, we choose the Lachtermacher and Fuller (1995) rule and vary the  $NH$  according to the size of the training set. Several experiments for ANN architectures with one and two hidden layers are performed in each case in order to determine the appropriate number of hidden layers. The conclusions obtained show that an ANN with two hidden layers perform better in terms of training mean squared error. The three hidden layer case may increase the network complexity and it is dropped off given that high training accuracy rates were obtained for less complex ones.

Demuth and Beale (2001), Nastac and Koskivaara (2003) and Costea (2003) compared different ANN gradient-descent-like training algorithms in order to choose the best algorithm for a specific problem. Using also these results, our approach is based on combining the sigmoid and linear activation functions for the hidden and output layers respectively, as it provided the best results in our experiments. Comparing four training algorithms in terms of error rates and convergence speed, Costea (2003) finds out that there exists a negative correlation between error rates and convergence speed. Therefore, the selection of the training algorithm is based on a trade-off between these two factors. Previous studies show that Scaled Conjugate Gradient (SCG) algorithm (Moller 1993) performs well when applied to a wide variety of problems. Although this is not the fastest algorithm, it has a good convergence rate and does not need large computational memory. In order to avoid the network over-fitting the training samples, we use the so-called validation-stop method which consists of separating the training data into effective training ( $TRe$ ) and validation ( $VAL$ ) datasets. The training process ends when the difference between the effective training error and the validation error exceeds a small value, which is given as a parameter. Hagan *et al.* (1996) and Demuth and Beale (2001) state that when using the validation stop attention must be paid in choosing the training algorithm so that it does not converge too rapidly. From this point of view, the SCG algorithm has a slow convergence and, therefore, works well when the validation stop method is applied.

The solution obtained when applying the ANN-based model for performance classification is represented by the set of the final connection weights of the network. The solution is evaluated using four accuracy measures: effective training dataset accuracy rate ( $ACR_{TRe}$ ), validation dataset accuracy rate ( $ACR_{VAL}$ ), total training dataset accuracy rate ( $ACR_{TR}$ ) and testing dataset accuracy rate ( $ACR_{TS}$ ). The higher these accuracy rates and the small the differences among them, the better the solution.

Except the training accuracy rate  $ACR_{TR}$ , in the literature there are different methods used to evaluate the performance of classification algorithms. For example, Kou *et al.* (2012), Peng *et al.* (2008, 2011a, 2011b) use multiple-criteria decision-making methods to rank the classification algorithms. In our experiments we used out-of-sample accuracy rate ( $ACR_{TS}$ ) to evaluate more rigorously our models.

In the next section, we present the dataset with the NFIs' financial performance indicators.

## 2. The NFIs' financial performance dataset

Unlike the CAAMPL system (Cerna *et al.* 2008), which consists of six performance dimensions and is applicable on credit institutions, in our approach regarding NFIs' industry, the number of the performance dimensions will be restricted to three quantitative dimensions, namely: capital adequacy (C), assets' quality (A) and profitability (P). The other quantitative dimension used in evaluating the credit institutions, namely liquidity dimension (L) is not applicable to NFIs, since these institutions do not attract retail deposits. The qualitative dimensions, namely quality of ownership (A) and management (M) were also excluded from our experiment because they involve a distinct approach and it does not fall into the scope of the present study to consider them.

After selecting the performance dimensions, we choose different indicators for each dimension, based on the analysis of the periodic financial statements of the NFIs. These financial ratios are presented in Table 1.

Table 1. The performance dimension and the corresponding financial ratios

Dimension	Indicators
Capital adequacy	Equity ratio (Leverage) = own capital / total assets (net value) Own capital / equity Indebtedness sources = borrowings / own capital
Assets' quality	Loans granted to clients (net value) / total assets (net value) Loan granted to clients (net value) / total borrowings Past due and doubtful loans (net value) / total loans portfolio (net value) Past due and doubtful claims (net value) / total assets (net value) Past due and doubtful claims (net value) / own capital
Profitability	Return on assets (ROA) = net income / total assets (net value) Return on equity (ROE) = net profit / own capital The rate of profit = gross profit / total revenues Activity cost = total costs / total revenues

The choice of the financial ratios of each dimension was based on Lehtinen's study (Lehtinen 1996) about the validity and reliability of financial ratios in international comparisons and they were collected with the help of the experts within the Supervision Department at the National Bank of Romania. The data for the financial ratios was gathered quarterly, from 2007 to 2010, from the NFIs' periodic financial statements. The NFIs were selected from the Special register, which includes only those NFIs from the General register which meet certain criteria of performance in terms of loans and borrowings. NFIs which meet these criteria corresponding to three reporting-periods in a row or three quarters are entered in the Special register. Conversely, if a NFI from the Special register does not match the criteria for three consecutive quarters, it will be de-classified in the General register. We chose only the NFIs from the Special register because, even if they are small in number, they represent more than 80% of the entire NFIs' industry in terms of financial value. In total, in the analyzed period, there were 65 active NFIs in the Special register. We collected the data for these 65 NFIs quarterly from 2007 to 2010, obtaining a total of 769 observations. Then, the above twelve financial ratios are computed. Four ratios, namely *Leverage* (for the capital adequacy dimen-

sion), *Loans\_to\_Assets* and *Loans\_to\_Borrowings* (for the assets' quality dimension) and *ROA* (for profitability dimension) were discarded from our analysis, due to the high variation of their values or incorrect values, remaining with eight ratios. Also, the data set composed by 769 observations multiplied by 8 ratios contains quarterly and yearly averages (16 quarterly averages and 4 yearly averages = 20 observations).

Our data set was preprocessed by leveling the outliers to the domain  $[-50, 50]$  and also by normalizing each ratio, i.e. subtracting the mean from each value and dividing the result by the standard deviation of the ratio, in order to avoid our techniques results being affected by these extremes/abnormal financial ratios.

### 3. Applying the two-stage methodology for NFIs' financial performance benchmarking

The aim of this experiment is to evaluate comparatively the performance of 65 Romanian NFIs registered in the Special Register which have been active since 2006, the first year this sector has been regulated in Romania. This analysis can be used by the Supervision Department of the National Bank of Romania to allocate more efficiently its resources. Identifying poorly performing NFIs would give the opportunity to the supervisors to concentrate on a small number of NFIs, which face difficulties. Before we apply our methodology, we present in Table 2 summary statistics related to each financial performance ratio before transforming the data (normalizing each ratio) as described at the end of preceding Section. As we presented in Section 2, we have discarded four financial ratios (financial ratio 1, 4, 5, and 9 in Table 2) due to their associated large kurtosis and skewness (the kurtosis and skewness for these financial ratios depart substantially from the values encountered for these variability measures in the case of standard normal distribution: 3 and 0, respectively). Next, we individually applied normalization on each of the remaining eight financial ratios.

The methodology was implemented using Matlab environment, by building different scripts based on the existing functions or entirely-new ones.

As it is presented in the previous sections, the first step of the methodology consists of applying the procedure based on the FCM algorithm in order to build clusters with similar performance.

Firstly, we generate one linguistic variable for each financial performance ratio according to Section 1.1. Consequently, we generated eight linguistic variables. We represent each linguistic variable with the help of five trapezoidal-form fuzzy numbers: *VL*, *L*, *A*, *H* and *VH*. In Figure 2 we show the graphical representation of the linguistic variables and their trapezoidal approximations.

This representation of our financial variables using linguistic terms allow us to characterize each performance cluster as having observations with *VL*, *L*, *A*, etc values. In fact, it allows us to order the clusters in terms of financial performance as being "worst", "bad", "average" or "best". In the implementation stage of the algorithm we have used  $m = 1.5$  for generating the linguistic variables and  $c = 4$  to make the results comparable with those from our previous work (Costea 2013a). The other parameters of FCM were set to  $N = 10000$  and  $\varepsilon = 0.00001$ .

The linguistic characterization of each cluster is presented in Table 3.

Table 2. Summary statistics for the twelve financial performance ratios

	1	2	3	4	5	6	7	8	9	10	11	12
Equity ratio (Leverage) = own capital / total assets (net value)	0.004	0.900	10.356	0.958	1.132	0.043	0.044	0.543	-0.044	-0.185	-0.085	1.085
Own capital / equity	0.047	0.507	0.713	0.063	0.044	0.003	0.003	0.247	0.025	0.162	0.017	0.017
Indebtedness sources = borrowings / own capital	0.067	1.001	9.583	0.897	1.022	0.012	0.014	0.135	-0.002	0.026	-0.007	1.007
Loans granted to clients (net value)	#N/A	-50.000	50.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.024	0.976
Loan granted to clients (net value) / total borrowings	1.311	14.069	19.778	1.760	1.224	0.097	0.096	6.841	0.701	4.488	0.460	0.460
Past due and doubtful loans (net value) / total loans	1.720	197.924	391.150	3.098	1.498	0.009	0.009	46.796	0.491	20.146	0.212	0.212
Past due and doubtful claims (net value) / own capital												
Past due and doubtful claims (net value) / total assets (net value)												
Return on assets (ROA) = net income / total assets (net value)												
Return on equity (ROE) = Net profit / own capital												
The rate of profit = gross profit / total revenues												
Activity cost = total costs / total revenues												
Mean	0.004	0.900	10.356	0.958	1.132	0.043	0.044	0.543	-0.044	-0.185	-0.085	1.085
Standard Error	0.047	0.507	0.713	0.063	0.044	0.003	0.003	0.247	0.025	0.162	0.017	0.017
Median	0.067	1.001	9.583	0.897	1.022	0.012	0.014	0.135	-0.002	0.026	-0.007	1.007
Mode	#N/A	-50.000	50.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.024	0.976
Standard Deviation	1.311	14.069	19.778	1.760	1.224	0.097	0.096	6.841	0.701	4.488	0.460	0.460
Sample Variance	1.720	197.924	391.150	3.098	1.498	0.009	0.009	46.796	0.491	20.146	0.212	0.212
Kurtosis	543.323	7.800	2.046	507.956	158.533	15.356	20.371	38.282	712.499	96.967	61.297	61.297
Skewness	-22.362	-0.901	-0.442	21.437	11.436	3.772	4.108	0.279	-26.264	-1.304	-6.049	6.049
Range	34.047	100.000	100.000	44.613	22.531	0.764	0.910	100.000	20.418	100.000	7.951	7.951
Minimum	-33.103	-50.000	-50.000	0.000	0.000	-0.050	-0.055	-50.000	-19.078	-50.000	-5.976	-0.975
Maximum	0.944	50.000	50.000	44.613	22.531	0.714	0.856	50.000	1.340	50.000	1.975	6.976
Sum	3.229	691.754	7963.619	736.659	870.808	32.708	33.814	417.314	-33.538	-141.932	-64.994	833.994
Count	769.000	769.000	769.000	769.000	769.000	769.000	769.000	769.000	769.000	769.000	769.000	769.000
Largest(1)	0.944	50.000	50.000	44.613	22.531	0.714	0.856	50.000	1.340	50.000	1.975	6.976
Smallest(1)	-33.103	-50.000	-50.000	0.000	0.000	-0.050	-0.055	-50.000	-19.078	-50.000	-5.976	-0.975
Confidence Level(95.0%)	0.093	0.996	1.400	0.125	0.087	0.007	0.007	0.484	0.050	0.318	0.033	0.033

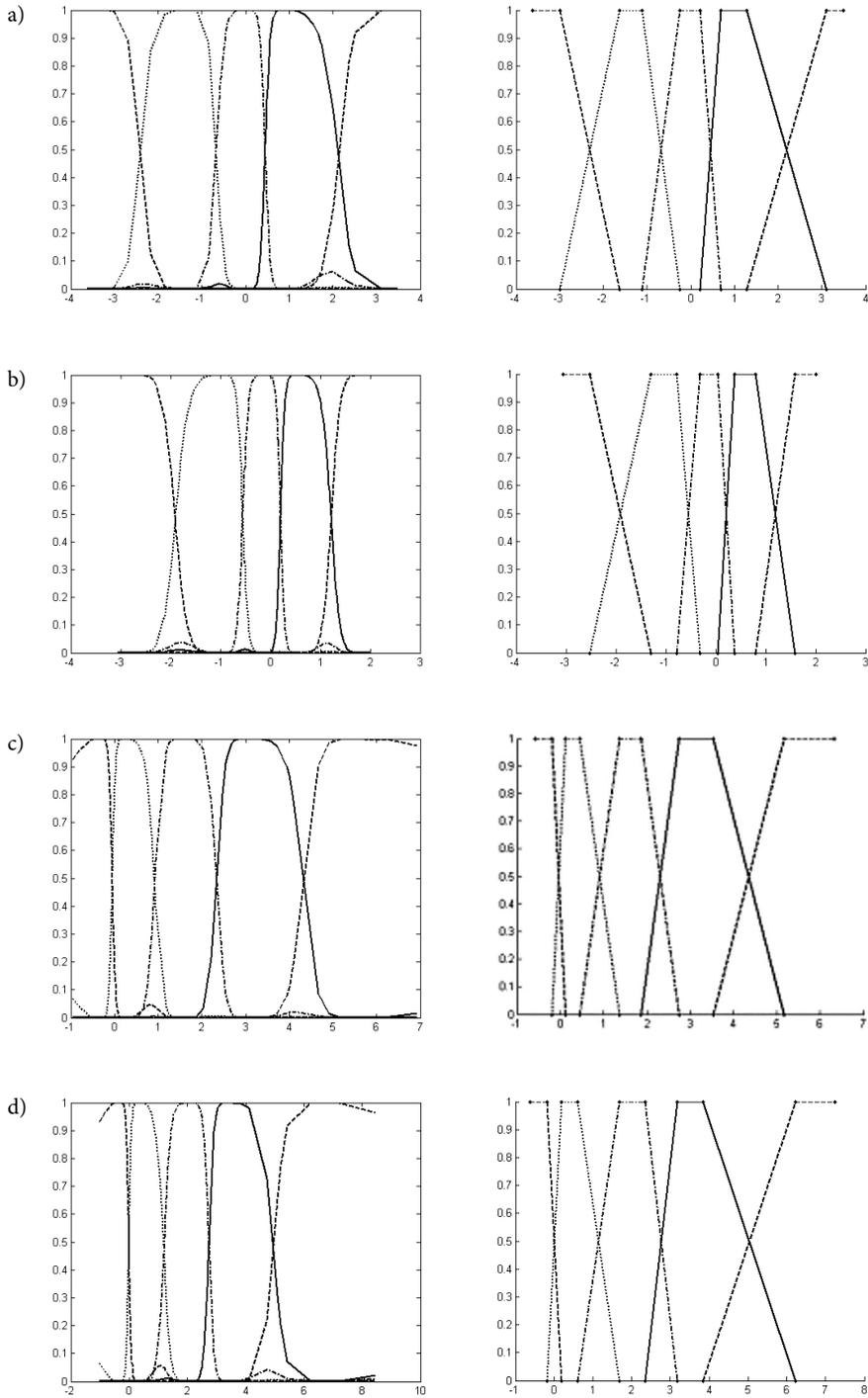


Fig. 2 (To be Continued)

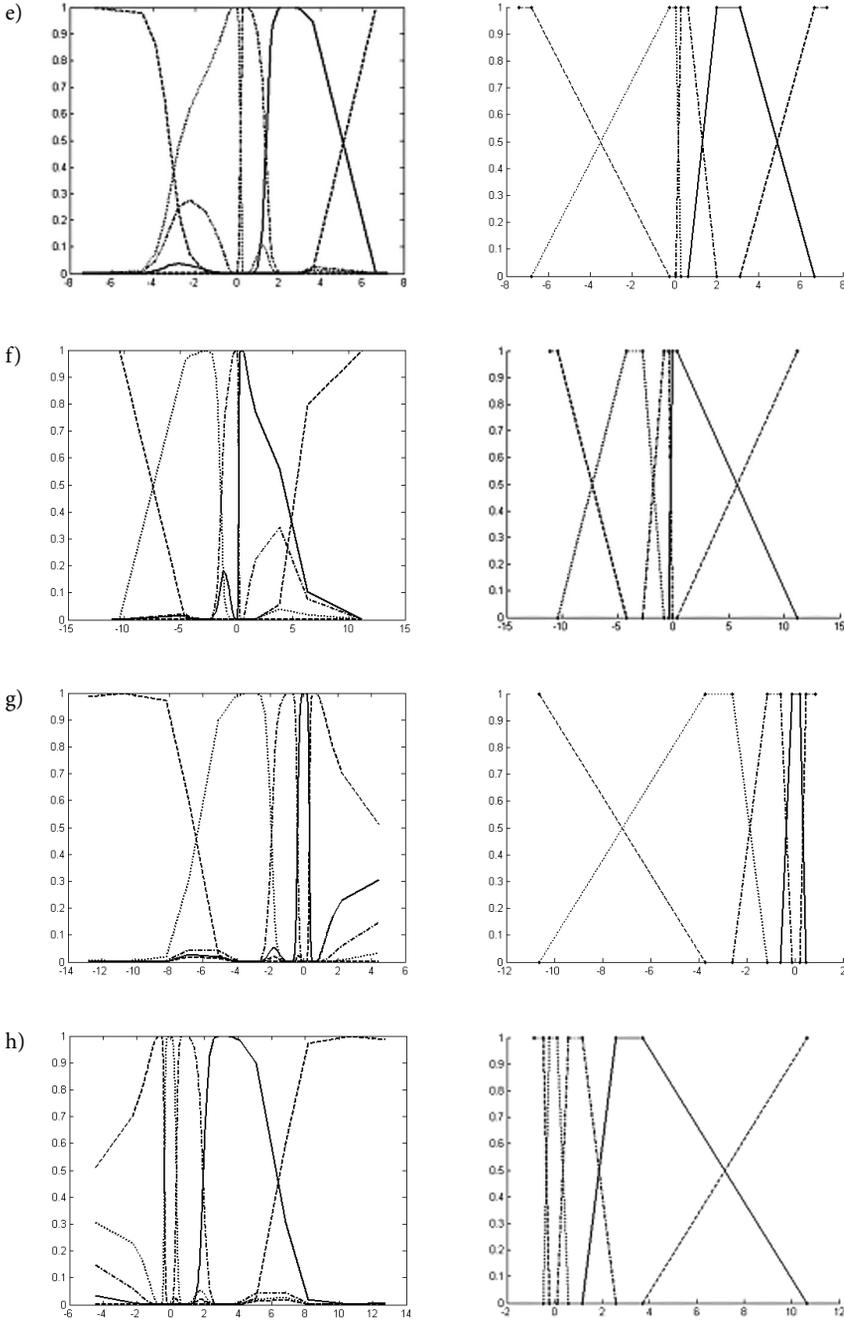


Fig. 2. Linguistic representation of eight financial variables (a) Own capital/equity, (b) Indebtedness sources, (c) Past due and doubtful loans (net value)/total loans portfolio (net value), (d) Past due and doubtful claims (net value)/total assets (net value), (e) Past due and doubtful claims (net value)/own capital, (f) ROE, (g) The rate of profit, (h) Activity cost and their trapezoidal approximations. Each trapezoid correspond to one linguistic term (very low - VL, low - L, average - A, high - H and very high - VH)

Table 3. Characterization of clusters

	Own capital to equity	Borrowings to own capital	Past due loans to loans	Past due loans to assets	Past due loans to own capital	ROE	Gross profit to revenues	Costs to revenues	Order	No. of certain observations	No. of observations after the allocation of the uncertain observations
Cluster 1	A	VH	VL	VL	L	A	H	L	Average	113	131
Cluster 2	VL	L	VL	VL	L	A	H	L	Worst	62	81
Cluster 3	A	A	H	A	A	A	H	-	Best	44	54
Cluster 4	A	A	VL	VL	L	A	H	L	Bad	451	503
Total										670	769

Each cluster was characterized based only on the certain 670 observations. The rest of 99 observations were considered uncertain in the sense that it is difficult to place them in a particular cluster, as they have similar membership degrees in two clusters. A so-called “perfectly” uncertain observation would have the same membership degree ( $1/C$ ) in each cluster. We considered that one linguistic term characterizes one cluster if it represents more than 40% out of total number of samples for that cluster. We have chosen 40% in order to allow maximum two linguistic terms to characterize a cluster for each ratio. According to the linguistic characterization, we can now label the clusters as being good or bad performance clusters (the third-last column in Table 3). For example, as compared to the other clusters, we can see from Table 3 that NFIs in cluster 3 are the best performers by having for all ratios average or high values. Of course, one can argue that classifying the clusters in this way involves the subjectivity of the analyst, but, first of all, the clusters might even not be classifiable at all (the  $p$ -dimensional input space might not be discriminable to begin with) and, when they are, the matrix in Table 3 helps the subjective analyst to correctly label them. However, the linguistic characterization is useful to the financial analysts, since they use the same expressions (“high rate of return”, “low capital adequacy”, etc.) when they represent the financial situation of the sector or the company.

After we characterize the clusters we re-allocated the uncertain observations within the clusters based on the procedure we presented in Section 1.1 and we obtained the structure of the clusters presented in the final column of Table 3. In this way, we were able to construct the so-called “class performance” variable: 1 – “Best” (54 observations), 2 – “Average” (131 observations), 3 – “Bad” (503 observations) and 4 – “Worst” (81 observations). This corresponds to a long right-tailed distribution of the  $p$ -dimensional input space (a positive skewness or few extreme high observations).

Once we have the “performance class” variable we can proceed with the second phase of our methodology: train a neural-network to model the relationship between the performance class variable and the input space.

We have selected 54 observations from each cluster, which means the number of observations in the smallest cluster, totalling 216 observations, in order to have an uniform number of observations in each cluster to train the classification model. In this stage we have also split the data in training (*TR*) and testing (*TS*) sets, by selecting one testing instance for every nine training instances. Thus, 192 observations for training and the rest of 24 observations for testing were randomly obtained.

The next step of the methodology consists of determining the proper architecture for the ANN-based classification model, which maps the 8-dimensional input space to the newly-constructed performance class variable. In our experiments, which were performed using Matlab's Neural Networks toolbox, all the parameters of the ANNs, like the learning algorithm (SCG), the performance goal of the classifier and the maximum number of epochs were kept constant, except the number of neurons in the hidden layers (*NH* when we had one hidden layer, or  $NH_1$ ,  $NH_2$  when we had two hidden layers).

Next we present the empirical procedure for determining the architecture of an ANN with two hidden layers. Firstly, three trainings were performed, in order to find the best ANN architecture. For each training, the training set (*TR*) was split into the effective training set (*TR<sub>e</sub>*) and the validation set (*VAL*), obtaining each time approximately 168 (9 out of 10) observations for effective training and 24 observations for validation. Validation-stop method was used as stopping criterion. We have applied the Lachtermacher and Fuller (1995) rule and varied *NH* from 3 (this value has to be greater than  $0.11 * 168 / (8 + 1) = 2.05$ ) to 5 (this value has to be lower than  $0.30 * 168 / (8 + 1) = 5.60$ ). For the two-hidden layer case we have selected more neurons on the first hidden-layer than on the second hidden-layer, simulating a rough and a fine-tuning training phases (we varied  $NH_1$  from 2 to 3 and  $NH_2$  from 1 to 2). Then, the network was trained for each ANN architecture, based on the effective training dataset. The best ANN architecture in terms of mean squared error for the effective training dataset ( $MSE_{TR_e}$ ) was saved if the mean squared error based on the validation set ( $MSE_{VAL}$ ) is less than  $6/5 * MSE_{TR_e}$ . This restriction has been imposed in order to avoid saving ANN architectures with large differences of the mean squared error between effective training and validation. The final ANN architecture consists of 3 neurons on the first hidden layer and 2 neurons on the second hidden layer. The following accuracy rates were obtained: effective training dataset accuracy rate ( $ACR_{TR_e}$ ) = 93.41 percent, validation dataset accuracy rate ( $ACR_{VAL}$ ) = 96.00 percent, total training dataset accuracy rate ( $ACR_{TR}$ ) = 93.75 percent and testing dataset accuracy rate ( $ACR_{TS}$ ) = 91.67 percent.

In our previous research we performed experiments on the same dataset using other clustering/classification techniques. In Tables 4 and 5 we compare the results of these different approaches.

As it can be seen from Table 4 the best technique for the first step of our two-step methodology is by far the modified FCM clustering algorithm presented in this paper. Our technique outperformed SOM (Costea 2013a), classical C-Means (Costea 2012a) and classical Fuzzy C-Means (Costea 2012b) in terms of both characterization of the performance clusters and pattern allocation. The modified FCM algorithm came third in terms of ease of implementation, but in comparison to classical C-Means and SOM, it provided a way of determining the relevant input financial ratios for each performance class.

Table 4. The comparison of different clustering techniques for assessing NFIs’ financial performance

Technique	Ease-of-implementation	Determination of attribute importance (relevance)	Linguistic characterization of clusters	Pattern allocation (in terms of correspondence to reality)
C-Means	Best	Not possible	Not possible	Worst
Fuzzy C-Means	Second Best	Possible	Objective: automatic characterization of the clusters based on the linguistic terms.	Second best
Self-Organising Maps	Worst. Highly parametrized.	Not possible	Subjective: based on the individual feature planes, the distances between neurons and analyst interpretation.	Third best
Modified Fuzzy C-Means	Third best	Possible	Objective: automatic characterization of the clusters based on the linguistic terms. A more robust linguistic characterization than Fuzzy C-Means: based only on the certain observations.	Best

Table 5. The comparison of different classification techniques for assessing NFIs’ financial performance

Technique	Ease-of-implementation	Result explanation capability	Pattern allocation (in terms of training accuracy)	Pattern allocation (in terms of testing accuracy)
Multinomial Logistic Regression	Third best	Second best	100.00	77.27 or 81.81
Decision Trees (C4.5 algorithm)	Second best	Best	87.90	81.80 or 88.30
Artificial Neural Networks	Best	Third best	93.75	91.67
Artificial Neural Networks trained with a Genetic Algorithm	Worst	Worst	The GA could not improve the ANN training accuracy	The GA could not improve the ANN testing accuracy

From Table 5 we can see that artificial neural networks performed best in terms of testing accuracy rates (the most rigorous criterion for evaluating classification models) and had the smallest difference between the training and testing accuracy rates. Moreover, ANN was the best technique in terms of ease-of-implementation, but lacked result explanation as compared to decision trees (Costea 2013b) and logistic regression (Costea 2014). Overall, our methodology (the combination between the modified Fuzzy C-Means algorithm and artificial neural networks) is the most appropriate one for assessing comparatively the NFIs’ financial performance.

## Conclusions

In this study we propose a two-stage methodology in order to assess comparatively the performance of non-banking financial institutions (NFIs) in Romania and forecast their future performance. We use a  $8 \times 769$  dataset that consists of 8 performance indicators measuring three performance dimensions: the capital adequacy, the assets' quality and the profitability of around 65 NFIs. The dataset includes quarterly and yearly averages. This type of analysis is beneficial for all parties in the financial industry and, in particular, for the Supervision Department of the central bank which would allocate more efficiently its resources by identifying in advance the NFIs with deteriorating performance. Thus, we consider that the implementation of such a model in the central bank could have a direct effect on the efficiency of supervision activity of these institutions in Romania. Moreover, the NFIs' performance classification models could be extended in the case of credit institutions and incorporated in the central bank's Decision Support System (Filip 2008; Filip *et al.* 2014).

In the first stage of our methodology we use a procedure based on the Fuzzy C-Means clustering that allows us to group the NFIs in similar performance clusters. We use four clusters for the implementation in order to compare the results with what we have obtained previously. We characterize each cluster with the help of linguistic variables with five linguistic terms (very low – *VL*, low – *L*, average – *A*, high – *H* and very high – *VH*). Each linguistic term is represented by a trapezoidal fuzzy number defined according to our previous work (Costea, Bleotu 2012). We characterize each cluster based only on the certain observations (observations that have large differences between the two highest membership degrees). The remaining observations (the uncertain ones) are re-allocated into the clusters based on some weights constructed as standardized variation coefficients (see Section 1.1). Basically, the higher the variation coefficient for one financial ratio and for one cluster, the better definer is the ratio for that particular cluster (many observations have the values concentrated in few linguistic terms). The linguistic characterization of the clusters allows us to order the clusters as containing the best, average, bad and worst performance NFIs. Consequently, we defined a “performance class” variable by attaching a performance class to each observation (1 – “Best performance” for 54 observations, 2 – “Average” for – 131 observations, 3 – “Bad performance” for 503 observations and 4 – “Worst performance” for 81 observations).

In the second stage of our methodology, we build classification models that map the 8-dimensional input space to the newly-constructed “performance class” variable in order to forecast the performance of new companies as data become available. The first phase of the classification approach proposed in this paper consists of applying some preliminary steps. Firstly, we preprocess the data by removing the outliers and standardizing the data using normalization, in order to counteract the effect of extreme values. The second step refers to selecting an even number of observation in each cluster to properly train the classifier and, finally, in the last preliminary step, we select the codification for the output variable. Four neurons were used to represent the class performance variable in order to eliminate the risk that one class with a higher value for the output variable would be more important in training than the others. In the fourth step, the proper ANN architecture was obtained by using the rule from Lachtermacher & Fuller (1995) and varying the number of neurons of first

and second hidden layers. The ANN obtained consists of 8 input neurons, 3 neurons in the first hidden layer, 2 neurons in the second hidden layer and 4 neurons on the output layer. This network was trained using variants of the backpropagation algorithm (e.g., the Scaled Conjugate Gradient algorithm). High accuracy rates and small differences among them were obtained. The final result shows a training accuracy rate of  $ACR_{TR} = 93.75$  percent and a testing accuracy rate of  $ACR_{TS} = 91.67$  percent. As compared to other reported techniques applied on the same dataset, our methodology proved to be the most appropriate one.

## Acknowledgements

This work was supported by the Supervision Department of the National Bank of Romania which kindly provided the periodic financial statements of the analysed non-banking financial institutions under its confidentiality agreement policy. In conducting this research, the third author was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS – UEFISCDI, project number PN-II-RU-TE-2014-4-2905.

## References

- Abad González, J.; Gutiérrez López, C. 2015. Can financial statements predict stress test results? Evidence from the Spanish case, *Economic Computation and Economic Cybernetics Studies and Research* 49(2): 231–248.
- Baležentis, A.; Baležentis, T. 2011. A novel method for group multi-attribute decision making with two-tuple linguistic computing: supplier evaluation under uncertainty, *Economic Computation and Economic Cybernetics Studies and Research* 45(4): 5–30.
- Basheer, I. A.; Hajmeer, M. 2000. Artificial neural networks: fundamentals, computing, design, and application, *Journal of Microbiological Methods* 43: 3–31.  
[https://doi.org/10.1016/S0167-7012\(00\)00201-3](https://doi.org/10.1016/S0167-7012(00)00201-3)
- Bezdek, J. C. 1981. *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press.  
<https://doi.org/10.1007/978-1-4757-0450-1>
- Boyacioglu, M. A.; Kara, Y.; Baykan, O. K. 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 Systems with Applications* 36(2): 3355–3366. <https://doi.org/10.1016/j.eswa.2008.01.003>
- Cerna, S.; Donath, L.; Seulean, V.; Herbei, M.; Bărglăzan, D.; Albulescu, C.; Boldea, B. 2008. *Financial Stability*. Timișoara: West University Publishing House.
- Cimpoeru, S. 2015. Using self-organising maps for assessing systemic risk. Evidences from the Global Economic Crisis, *Economic Computation and Economic Cybernetics Studies and Research* 49(2): 71–90.
- Costea, A. 2003. Economic performance classification using neural networks, in *Proceedings of the Sixth International Conference on Economic Informatics “Digital Economy”*, 8–11 May, 2003, Bucharest, Romania: INFOREC Printing House, 390–395.
- Costea, A. 2005. *Computational intelligence methods for quantitative data mining*: PhD thesis No. 67. Åbo Akademi University and Turku Centre for Computer Science, Turku: Painosalama Oy.
- Costea, A. 2012a. Evaluating the performance of non-banking financial institutions by the means of C-Means algorithm, *Journal of US-China Public Administration* 9(4): 458–466.

- Costea, A. 2012b. Applying Fuzzy C-Means and Artificial Neural Networks for analyzing the non-banking financial institutions' sector in Romania, *Journal of Applied Quantitative Methods* 7(3): 26–32.
- Costea, A. 2013a. Performance benchmarking of non-banking financial institutions by means of Self-Organising Map algorithm, *East-West Journal of Economics and Business* 16(1): 37–58.
- Costea, A. 2013b. A decision-tree induction approach to assessing comparatively the performance of non-banking financial institutions in Romania in *Proceedings of the third international conference – International Conference on International Business (ICIB 2012)*, May 17–19, 2012, Thessaloniki, Greece, 406–416.
- Costea, A. 2014. A statistical-based approach to assessing comparatively the performance of non-banking financial institutions in Romania, Chapter 11 in A. Karasavvoglou, P. Polychronidou (Eds.). *Economic crisis in Europe and the Balkans – problems and prospects, contributions to economics*. Springer International Publishing Switzerland, 195–209.
- Costea, A.; Bleotu, V. 2012. A new fuzzy clustering algorithm for evaluating the performance of non-banking financial institutions in Romania, *Economic Computation and Economic Cybernetics Studies and Research* 46(4): 179–199.
- Costea, A.; Eklund, T. 2004. Combining clustering and classification techniques for financial performance analysis, in *Proceedings of 8th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI 2004)*, 18–21 July, 2004, Orlando, Florida, USA, 389–394.
- Costea, A.; Nastac, I. 2005. Assessing the predictive performance of ANN-based classifiers based on different data preprocessing methods, distributions and training mechanisms, *International Journal of Intelligent Systems in Accounting, Finance and Management* 13(4): 217–250.  
<https://doi.org/10.1002/isaf.269>
- Demuth, H.; Beale, M. 2001. *Neural network toolbox*. The MathWorks Inc, MA, USA: Natick Press.
- Doumpos, M.; Zopounidis, C. 2009. A multicriteria bank rating system, *European Working Group “Multiple Criteria Decision Aiding”* 3(19): 17–19.
- Drobics, M.; Winiwarter, W.; Bodenhofer, U. 2000. Interpretation of self-organizing maps with fuzzy rules, in *Proceedings of the Twelfth IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, 13–15 November, 2000, IEEE Press, Vancouver, BC, Canada, 304–311. ISBN: 0-7695-0909-6. <https://doi.org/10.1109/tai.2000.889887>
- Erdogan, B. E. 2013. Prediction of bankruptcy using support vector machines: an application to bank bankruptcy, *Journal of Statistical Computation and Simulation* 83(8): 1543–1555.  
<https://doi.org/10.1080/00949655.2012.666550>
- Filip, F. G. 2008. Decision support and control for large-scale complex systems, *Annual Reviews in Control* 32(1): 61–70. <https://doi.org/10.1016/j.arcontrol.2008.03.002>
- Filip, F. G.; Suduc, A. M.; Bizoi, M. 2014. DSS in numbers, *Technological and Economic Development of Economy* 20(1): 154–164. <https://doi.org/10.3846/20294913.2014.890139>
- Hagan, M. T.; Demuth, H. B.; Beale, M. 1996. *Neural network design*. Boston, USA: PWS Publishing Company.
- Hajek, P.; Olej, V.; Myskova, R. 2014. Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making, *Technological and Economic Development of Economy* 20(4): 721–738. <https://doi.org/10.3846/20294913.2014.979456>
- Herrera, F.; Herrera-Viedma, E.; Martinez, L. 2008. A fuzzy linguistic methodology to deal with unbalanced linguistic term sets, *IEEE Transactions on Fuzzy Systems* 16(2): 354–370.  
<https://doi.org/10.1109/TFUZZ.2007.896353>
- Herrera, F.; Herrera-Viedma, E.; Verdegay, J. L. 1994. Direct approach processes in group decision making using linguistic OWA operators, *Fuzzy Sets and Systems* 79: 175–190.  
[https://doi.org/10.1016/0165-0114\(95\)00162-X](https://doi.org/10.1016/0165-0114(95)00162-X)

- Hevner, A. R.; March, S. T.; Park, J.; Ram, S. 2004. Design science in information systems research, *MIS Quarterly* 28(1): 75–105.
- IMF. 2010. *Romania: financial sector stability assessment* [online], [cited 30 October 2014]. International monetary fund country report No. 10/47. Available from Internet: <http://www.imf.org/external/pubs/ft/scr/2010/cr1047.pdf>
- Ionita, R. O.; Stancu, D. 2015. Early warning models for debt crises – case study for Romania, Czech Republic and Hungary, *Economic Computation and Economic Cybernetics Studies and Research* 49(2): 107–128.
- Iturriaga, F. J. L.; Sanz, I. P. 2013. Self-organizing maps as a tool to compare financial macroeconomic imbalances: the European, Spanish and German Case, *The Spanish Review of Financial Economics* 11: 69–84. <https://doi.org/10.1016/j.srfe.2013.07.001>
- Jain, A. K.; Murty, M. N.; Flynn, P. J. 1999. Data clustering: a review, *ACM Computing Survey* 31(3): 264–323. <https://doi.org/10.1145/331499.331504>
- Kohonen, T. 1997. *Self-organising maps*. 2<sup>nd</sup> ed. Heidelberg: Springer-Verlag. <https://doi.org/10.1007/978-3-642-97966-8>
- Koskivaara, E. 2004. Artificial neural networks in analytical review procedures, *Managerial Auditing Journal* 19(2): 191–223. <https://doi.org/10.1108/02686900410517821>
- Kou, G.; Lu, Y.; Peng, Y.; Shi, Y. 2012. Evaluation of classification algorithms using MCDM and rank correlation, *International Journal of Information Technology & Decision Making* 11(1): 197–225. <https://doi.org/10.1142/S0219622012500095>
- Kou, G.; Peng, Y.; Wang, G. 2014. Evaluation of clustering algorithms for financial risk analysis using MCDM methods, *Information Sciences* 275: 1–12. <https://doi.org/10.1016/j.ins.2014.02.137>
- Kumar, P. R.; Ravi, V. 2007. Banks and firms in bankruptcy prediction via intelligent and statistical techniques – a review, *European Journal of Operational Research* 180(1): 1–28. <https://doi.org/10.1016/j.ejor.2006.08.043>
- Lachtermacher, G.; Fuller, J. D. 1995. Backpropagation in time series forecasting, *Journal of Forecasting* 14: 381–393. <https://doi.org/10.1002/for.3980140405>
- Lehtinen, J. 1996. Financial Ratios in an international comparison. Validity and reliability. Vaasa, Finland: Acta Wasaensia 49.
- Lindholm, C. K.; Liu, S. 2003. *Fuzzy clustering analysis of the early warning signs of financial crisis* [online], [cited 25 October 2014]. Working Papers No. 472, IAMSR, Åbo Akademi University. Available from Internet: [http://iamsr.abo.fi/publications/openFile.php?pub\\_id=472](http://iamsr.abo.fi/publications/openFile.php?pub_id=472)
- Massanet, S.; Riera, J. V.; Torrens, J.; Herrera-Viedma, E. 2014. A new linguistic computational model based on discrete fuzzy numbers for computing with words, *Information Sciences* 258: 277–290. <https://doi.org/10.1016/j.ins.2013.06.055>
- Masters, T. 1994. *Practical neural network recipes in C++*. Boston, Ma: Academic Press.
- Moller, M. F. 1993. A scaled conjugate gradient algorithm for fast supervised learning, *Neural Networks* 6: 525–533. [https://doi.org/10.1016/S0893-6080\(05\)80056-5](https://doi.org/10.1016/S0893-6080(05)80056-5)
- Nastac, I.; Koskivaara, E. 2003. *A neural network model for prediction: architecture and training analysis* [online], [cited 10 January 2015]. TUCS Technical Report No. 521. Available from Internet: <http://www.tucs.fi/publications/insight.php?id=tNaKo03a&table=techreport>
- Pendharkar, P. C. 2002. A computational study on the performance of artificial neural networks under changing structural design and data distribution, *European Journal of Operational Research* 138: 155–177. [https://doi.org/10.1016/S0377-2217\(01\)00085-6](https://doi.org/10.1016/S0377-2217(01)00085-6)
- Pendharkar, P. C.; Rodger, J. A. 2004. An empirical study of impact of crossover operators on the performance of non-binary genetic algorithm based neural approaches for classification, *Computers and Operations Research* 31(4): 481–498. [https://doi.org/10.1016/S0305-0548\(02\)00229-0](https://doi.org/10.1016/S0305-0548(02)00229-0)

- Peng, Y.; Kou, G.; Shi, Y.; Chen, Z. 2008. A descriptive framework for the field of Data Mining and Knowledge Discovery, *International Journal of Information Technology and Decision Making* 7(4): 639–682. <https://doi.org/10.1142/S0219622008003204>
- Peng, Y.; Kou, G.; Wang, G.; Shi, Y. 2011a. FAMCDM: A fusion approach of MCDM methods to rank multiclass classification algorithms, *Omega* 39(6): 677–689. <https://doi.org/10.1016/j.omega.2011.01.009>
- Peng, Y.; Wang, G.; Kou, G.; Shi, Y. 2011b. An empirical study of classification algorithm evaluation for financial risk prediction, *Applied Soft Computing* 11(2): 2906–2915. <https://doi.org/10.1016/j.asoc.2010.11.028>
- Pivk, A.; Vasilecas, O.; Kalibatiene, D.; Rupnik, R. 2013. On approach for the implementation of data mining to business process optimisation in commercial companies, *Technological and Economic Development of Economy* 19(2): 237–256. <https://doi.org/10.3846/20294913.2013.796501>
- Rodriquez, R.; Martinez, L.; Herrera, F. 2012. Hesitant fuzzy linguistic term sets for decision making, *IEEE Transactions on Fuzzy Systems* 20(1): 109–119. <https://doi.org/10.1109/TFUZZ.2011.2170076>
- Schaffer, J. D. 1994. Combinations of genetic algorithms with neural networks or fuzzy systems, in J. M. Zurada, R. J. Marks, C. J. Robinson (Eds). *Computational Intelligence: Imitating Life*. New York: IEEE Press, 371–382.
- Stanujkic, D. 2013. An extension of the MOORA method for solving fuzzy decision making problems, *Technological and Economic Development of Economy* 19(S1): S228–S255. <https://doi.org/10.3846/20294913.2013.880083>
- Swicegood, P.; Clark, J. A. 2001. Off-site monitoring systems for predicting bank underperformance: a comparison of neural networks, discriminant analysis, and human Professional Judgement, *Intelligent Systems in Accounting, Finance and Management* 10(3): 169–186. <https://doi.org/10.1002/isaf.201>
- Șerban, F.; Ștefănescu, M. V.; Dedu, S. 2012. Algorithm in three stages for portfolio optimization – application in case of Italian Stock market, in *Proceedings of the 19<sup>th</sup> International Economic Conference (IECS)*, 15–16 June 2012, Sibiu, Romania.
- Ștefănescu, M. V.; Ferrara, M.; Dedu, S. 2008. Algorithms for hierarchical classification with applications in portfolio management, *Economic Computation and Economic Cybernetics Studies and Research* 42(3–4): 109–122.
- Toma, A.; Dedu, S. 2014. Quantitative techniques for financial risk assesment: a comparative approach using different risk measures and estimation methods, *Procedia Economics and Finance* 8: 712–719. [https://doi.org/10.1016/S2212-5671\(14\)00149-X](https://doi.org/10.1016/S2212-5671(14)00149-X)
- Tudor, M.; Dedu, S. 2012. A new risk management model using quantile-based risk measure, with applications to non-normal distributions, *Recent Researches in Applied Information Science, Recent Advances in Computer Engineering Series 2*: 183–186.
- Upadhyaya, B. R.; Eryurek, E. 1992. Application of neural network for sensory validation and plant monitoring, *Neural Technology* 97: 170–176.
- Wang, J. W.; Cheng, C. H.; Huang, K. C. 2009. Fuzzy hierarchical TOPSIS for supplier selection, *Applied Soft Computing* 9(1): 377–386. <https://doi.org/10.1016/j.asoc.2008.04.014>
- Yao, X. 1999. Evolving artificial neural networks, *Proceedings of the IEEE* 87(9): 1423–1447. <https://doi.org/10.1109/5.784219>

**Adrian COSTEA.** He is Professor of Data Mining and Econometrics at Bucharest University of Economic Studies, Department of Statistics and Econometrics. He received his doctorate in Information Systems from Åbo Akademi University in 2005. His recent publications include *Towards an early-warning system of distressed non-banking financial institutions* (ECECSR 48(2): 75–90, 2014) and *A statistical-based approach to assessing comparatively the performance of non-banking financial institutions in Romania* (Book Chapter 11, pp. 195–209 in A. Karasavvoglu and P. Polychronidou (Eds.), *Economic Crisis in Europe and the Balkans, Contributions to Economics*, Springer International Publishing Switzerland, 240 pg., 2014). His research interests include the use of Data Mining methods for assessing comparatively the performance of financial entities, and he is currently pursuing the CFA program (he passed Level II Exam in June 2016).

**Massimiliano FERRARA.** He is Full Professor of Mathematical Economics at “Mediterranea” University of Reggio Calabria where he is Vice Rector, actually. He was the Founder and Director of MEDAlics and Vice Rector at “Dante Alighieri” University of Reggio Calabria. Research Affiliate at CRIOS – Center for Research on Innovation, Organization and Strategy, University Bocconi of Milan. He was also Visiting Professor at Harvard University, Cambridge (USA), Morgan State University in Baltimore (USA), Western Michigan University (USA), New Jersey Institute of Technology in Newark (NJ) (USA). He is editor and co-editor of several international journals and Member of Mathematical Association of America and of the Indian Academy of Mathematics, Indore. Author and co-author of 160 papers on peer-review and ISI journals. His main research interests are: dynamical systems, patterns of growth and sustainable development, mathematical economics, game theory, optimization theory, applied Economics.

**Florentin ȘERBAN.** He is Lecturer, Department of Applied Mathematics, Bucharest University of Economic Studies. He is PhD in Economics, Bucharest University of Economics and PhD candidate in Mathematics, Doctoral School of Mathematics, University of Bucharest. His doctoral theses are about Portfolio Optimization and his research interests include: financial mathematics, risk theory, portfolio optimization, data analysis. He published 8 books, over 20 papers and has had over 15 talks to National and International Conferences.