Available online at www.sciencedirect.com

ScienceDirect

Procedia Economics and Finance 10 (2014) 4 – 9

Procedia Economics and Finance

www.elsevier.com/locate/procedia

7th International Conference on Applied Statistics

Applying fuzzy logic and machine learning techniques in financial performance predictions

Adrian Costea[a,b,*]

[a]University of Economic Studies, Statistics and Econometrics Department, Virgil Madgearu Building, Calea Dorobantilor No.15-17, 6th floor, Sector 1, Bucharest 010552, Romania

[b]National Bank of Romania, Bucharest Regional Branch, Statistical Reporting and Business Surveys Unit, Lipscani st. No. 16, 5th floor, room 5090, Sector 3, Bucharest 030035, Romania

Abstract

In this article we apply a fuzzy logic technique, namely Fuzzy C-Means clustering, and artificial intelligence algorithms for evaluating comparatively the financial performance of non-banking financial institutions (NFIs) in Romania. The NFIs' performance dataset consists of indicators that define the capital adequacy, assets' quality and profitability performance dimensions. The class performance variable is obtained by applying on the performance dataset the Fuzzy C-Means algorithm and obtaining clusters with similar performance. We attach to each input dataset observation a performance class depending on which cluster contains the observation given the characterization and hierarchy of the clusters in "good", "medium" and "poor" performance clusters. Finally, we apply artificial neural networks (ANNs) trained with genetic algorithms in order to find a function that maps the input performance space on the newly constructed performance class variable. The classification model obtained can be used by different beneficiaries (e.g.: the Supervision Department of National Bank of Romania) to classify new NFIs as having a "good" or "poor" performance so that the limited resources of the supervision authority to be better allocated.

Selection and peer-review under responsibility of the Department of Statistics and Econometrics, Bucharest University of Economic Studies.

Keywords: classification models; artificial neural networks; genetic algorithms; non-banking financial institutions

* Corresponding author. Tel.: +4-031-132-6207; fax: +4-021-319-1793.
E-mail address: adrian.costea@csie.ase.ro

Selection and peer-review under responsibility of the Department of Statistics and Econometrics, Bucharest University of Economic Studies.
doi:10.1016/S2212-5671(14)00271-8

1. Introduction

The non-banking financial institutions' (NFIs') sector has been recently regulated in Romania. The National Bank of Romania (NBR) proposed and the Parliament passed a series of provisions regarding the activities carried out by these entities in order to strengthen the stability of the financial sector as a whole. According to these provisions, the NBR is obliged to prudentially supervise the most important NFIs: those register in the NBR' Special register. The other NFIs are monitored and only in special cases are further scrutinized. However, all NFIs have to report on their activity to the Supervision Department. One such reporting consists of periodic financial statements (PFSs) that a NFI has to sent to NBR quarterly. Currently, these PFSs are analysed manually by the inspectors from the department. Based on their assessment, NFIs that present difficulties are further scrutinized and, eventually, an on-site inspection is organized. The scarce time and personnel resources of the Supervision Department and the need to balance the subjective interpretation of the inspectors motivate the use of some sorts of techniques that would classify the NFIs as having a "good" or "poor" performance. Based on these techniques we can build so-called classification models that might provide additional help in taking the decision for an on-site inspection.

Data Mining techniques (Han & Kamber, 2006) can help with building the clustering/classification models. Clustering techniques can be used to find performance clusters within the NFIs' performance dataset and classification techniques can be used to place a new NFI into a predefined cluster as data become available. In other words, clustering techniques have descriptive properties and classification techniques have prescriptive ones.

In this article we apply the Fuzzy C-Means algorithm (Bezdek, 1981) and obtaining clusters with similar performance. We attach to each input dataset observation a performance class depending on which cluster contains the observation given the characterization and hierarchy of the clusters in "good", "medium" and "poor" performance clusters. Finally, we apply artificial neural networks (ANNs) trained with genetic algorithms in order to find a function that maps the input performance space on the newly constructed performance class variable. In this way we overcome the problem of accommodating new NFIs on the performance map as data become available.

Next, we present in more detail our methodology. Then, we present the NFIs' performance dataset and our experiment on applying the Data Mining techniques on these data. Finally, we draw our conclusions.

2. Methodological framework

We use Fuzzy C-Means (FCM) algorithm (Bezdek, 1981) to group the NFIs with similar characteristics. The FCM algorithm minimizes the following objective function, $J_m(U, v)$:

$$J_m(U,v) = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m (d_{ik})^2 \tag{1}$$

where c is the number of clusters, n is the number of observations, $U \in M_{fc}$ is a fuzzy c-partition of the data set X, $u_{ik} \in [0,1]$ is the membership degree of observation x_k in cluster i,

$$d_{ik} = \|x_k - v_i\| = \left[\sum_{j=1}^{p}(x_{kj} - v_{ij})^2\right]^{1/2} \tag{2}$$

is the Euclidean distance between the cluster center v_i and observation x_k for p attributes (financial ratios in our case), $m \in [1,\infty)$ is the weighting exponent, and the following constraint holds

$$\sum_{i=1}^{c} u_{ik} = 1. \tag{3}$$

The solution for the FCM clustering consist of the final cluster centers and membership degrees of the observations. Once we obtain the final cluster centers and membership degrees we construct the performance class