Supporting Better Practice Benchmarking: A DEA-ANN Approach to Bank Branch Performance Assessment

Ioannis E. Tsolas*, **Vincent Charles**, and **Tatiana Gherman**

Abstract

The quest for best practices may lead to an increased risk of poor decision-making, especially when aiming to attain best practice levels reveals that efforts are beyond the organization's present capabilities. This situation is commonly known as the "best prac-tice trap". Motivated by such observation, the purpose of the present paper is to develop a practical methodology to support better practice benchmarking, with an application to the banking sector. In this sense, we develop a two-stage hybrid model that employs Artificial Neural Network (ANN) via integration with Data Envelopment Analysis (DEA), which is used as a preprocessor, to investigate the ability of the DEA-ANN approach to classify the sampled branches of a Greek bank into predefined efficiency classes. ANN is integrated with a family of radial and non-radial DEA models. This combined approach effectively captures the information contained in the characteristics of the sampled branches, and subsequently demonstrates a satisfactory classification ability especially for the efficient branches. Our prediction results are presented using four performance measures (hit rates): percent success rate of classifying a bank branch's performance exactly or within one class of its actual performance, as well as just one class above the actual class and just one class below the actual class. The proposed modeling approach integrates the DEA context with ANN and advances benchmarking practices to enhance the decision-making process for efficiency improvement.

Keywords Artificial Neural Network, Data Envelopment Analysis, Banking, Performance, Best Practice, Benchmarking.

*Corresponding author.

1. Introduction

In commercial banking, the operation and performance of branch networks is bound to have a significant impact on a bank's overall efficiency. In the banking industry, due to increased competition, banks monitor their branch networks in order to improve their performance. In the performance measurement literature, data envelopment analysis (DEA) has proven to be a viable technique in terms of efficiency measurement and target setting by identifying benchmarks (Charles et al., 2016; Charles et al., 2018). DEA computes the relative efficiency of an individual branch against best-performers, i.e., benchmarks, and moreover, it identifies under-performing branches so that managerial actions can be taken to improve their performance (Tsolas, 2011). DEA has been successfully used to identify bank branch benchmarks in line with three approaches (Giokas, 2008; Paradi et al., 2004), namely, production, intermediation, and profitability or profit-oriented approach (Gaganis et al., 2009). In this paper, we adopt a profit-oriented approach to model branch profitability in the DEA context.

In DEA, the units being compared are called decision-making units (DMUs), since they enjoy a certain decisional autonomy. The purpose of DEA is to identify those DMUs that are deemed to be efficient by means of assigning them an efficiency score of one, then expressing several suggestions to improve the efficiency of the inefficient units. For each inefficient unit, DEA identifies a set of best-in-class units, called a peer group, which include units that are efficient (Vercellis, 2009). DEA treats a given DMU as an entity with observed inputs and outputs, without any regard to the process by which inputs are converted into outputs; thus, DEA treats this conversion process as a "black box" common to the homogenous set of DMUs.

Neural Network (NN) modeling aims to also build a "black box" model (i.e., an Artificial Neural Network (ANN)) of the unknown relationships of the sampled data. It is a data mining method appropriate for situations in which the relationship between the input and output variables is unknown and when we are more interested in prediction rather than explanation. Although both DEA and NNs cannot provide insights regarding the mechanism of transformation of inputs into outputs, they can give information regarding the results of the transformation (Samoilenko & Osei-Bryson, 2010).

The structure of the NNs has been inspired from the human brain functions. NNs as sys-tems consist of a number of interconnected neurons. Similar to human thinking, NNs learn by example and they are trained by adjusting the weights between interconnected neurons; i.e., an input leads to a target (i.e., desired) output. Different ANN models have been proposed in the literature, such as: Multilayer perceptron (MLP), Hopfield networks, and Kohonen's self-organizing networks, among others. The MLP networks are used in most cases due to their capability to map arbitrary inputs and outputs. An MLP network includes the following components: the input (lower or information receiver) layer, the output (highest) layer, and one or more hidden (intermediate) layers. MLPs are usually trained using Back Propagation (BP) algorithms. BP networks are a class of feed-forward neural networks (i.e., the information flows from the input to the output layer) and the network's forecasts are compared with the known target or desired output and the weights are adjusted based on the forecasting error to minimize an error function (Shokrollahpour et al., 2016). For different algorithms, see Gallant & Stephen (1993) and Kheirkhah et al. (2013).

In this paper, we focus on proposing a practical methodology to support benchmarking analysis in bank branch networks by employing DEA and ANN in tandem. Branches may be motivated to learn not only from best in-class performers, but also from other better performers that lie in lower classes of performance. We experiment with a data set of bank branches by using families of radial and non-radial DEA models to extend the best in-class performers in the DEA context. The case study concerns a sample of branches of a large commercial Greek bank (henceforth, simply "The Bank"). The "black box" approach on which both methods are based fits the purpose of our study well, since we are interested in the classification ability of NNs when we have already used DEA as a preprocessor for deriving the efficiency scores of sampled DMUs. The remainder of this paper unfolds as follows: in Section 2, we provide a snapshot of the literature on DEA and ANN. Subsequently, in Section 3, we propose the DEA-ANN hybrid methodology, which is then followed by a discussion of results in Section 4. Managerial implications are presented in Section 5. Finally, Section 6 concludes the paper with limitations of current research and avenues for future research on the topic.

2. Literature review

Both DEA and ANN methods have no strict assumptions, and as a result, their applications are wide. For recent reviews on DEA and ANN, the interested reader is referred to the studies by Liu et al. (2013) and Tkáč & Verner (2016), respectively. Hybrid models that use both DEA and ANN can also be found in the literature. The possibility of using DEA and ANN in tandem was first proposed by Athanassopoulos & Curram (1996), who employed ANN as a tool for assessing the efficiency of DMUs and concluded that both DEA and ANN were comparable and complementary methods in performance evaluation. In that new strand of research, some studies considered the use of ANN as an alternative to DEA, providing varying results (Costa & Markellos, 1997; Santin et al., 2004; Wang, 2003; Azadeh et al., 2007). For a recent review, also see Samoilenko & Osei-Bryson (2010), who stated that NNs augment, rather than replace, DEA.

In line with Athanassopoulos & Curram (1996), a few researchers employed DEA as a preprocessor to produce training sets and to contribute to the computational efficiency derived by subsequent NNs (Emrouznejad & Shale, 2009; Kheirkhah et al., 2013; Pendharkar, 2011; Pendharkar & Rodger, 2003). Relevant studies in this strand focused mainly on the prediction of DEA efficiency scores (Çelebi & Bayraktar, 2008; Hsiang-Hsi et al., 2013; Kuo et al., 2010; Ozdemir & Temur, 2009; Sreekumar & Mahapatra, 2011; Wang, 2003; Wu, 2009). Wang (2003) produced more accurate predictions due to the preservation of monotonicity of the input observations. To tackle this problem of monotonicity, Misiunas et al. (2016) used the stratification DEA model proposed by Seiford & Zhu (2003).

Regarding the use of both DEA and ANN in the banking industry, apart from the work of Athanassopoulos & Curram (1996), we have the study by Wu et al. (2006), who combined the two techniques to measure the performance of a large Canadian bank and concluded that the proposed DEA-ANN method produces a more robust frontier and assists in identifying more efficient DMUs. Further, Angelidis & Lyroudi (2006) used both DEA and ANN to assess the efficiency of the Italian banking industry. Mostafa (2009) employed a probabilistic NN approach for modeling and classifying the efficiency of the Gulf Cooperation Council (GCC) banks. Finally, Shokrollahpour et al. (2016) employed an integrated DEA-ANN to find possible future benchmarks of bank branches.

The present study employs a DEA-ANN approach using a large size of branch data from a Greek commercial bank in an attempt to build a hybrid model using DEA as a preprocessor and to investigate whether benchmark analysis results can support the decision-making process to achieve efficiency improvement. In practice, branches may be motivated to learn not only from best performers (i.e., peers in the DEA context), but also from other better performers. The motivation behind the development of such a model lies in the need to avoid the so called "best practice trap", which may happen when seeking the best performers may not always yield the best results (Agarwal et al., 2013; Francis & Holloway, 2007; Kwon et al., 2016). Caution must be exercised when using the word "best'". According to the International Quality Study (1993), cited in Davies & Kochhar (2002), best practices are those practices that support lower performers to improve to medium performance, medium performers to achieve higher performance, and higher performers to stay on top and to continue to be successful. However, as Kwon et al. (2016) stated (citing Francis & Holloway, 2007), "in the midst of a benchmarking practice and literature dominated by "best practices," the conceptual development of better practice is still shallow and the lack of a proper evaluation methodology was pointed out as both a cause of this as well as a future necessity" (p. 522).

This paper thus focuses on developing a methodology to support benchmarking analyses in bank branch networks by employing DEA and ANN in tandem. DEA as a sole method has some shortcomings. Because of the DEA optimization principles, when a new data set is fed to a DEA model, it needs to find the results from the very beginning, and this creates a high computational burden especially in the case of big data. After the development of the proposed DEA-ANN algorithm and the production of satisfactory results for a selected DEA model, the corresponding ANN model can be employed solely with a new data set by just using the pre-found ANN weights.

To the best of our knowledge, this paper is the first to adopt the profit-oriented approach to model branch profitability by employing families of radial and non-radial DEA models with ANN. Thus, the findings of our analysis may be useful for generalizing our DEA-ANN mod-eling initiative to other banks in the same or different countries. It is also expected that the addition of ANN to the performance measurement framework can enhance the decision-making process in efficiency improvement and add some value to existing literature.

We differentiate from previous studies that use DEA-ANN approaches for efficiency prediction, such as the works by Kwon et al. (2016), Shokrollahpour et al. (2016), and Wu et al. (2006), since we are only interested in the classification ability of our proposed DEA-ANN modeling. We do not only provide the steps for the employment of the DEA-ANN modeling as a classi-fication tool, but we also investigate a family of radial and non-radial DEA models within the ANN framework. By contrast, the studies by Kwon et al. (2016) and Wu et al. (2006) used only the radial CCR and BBC models, and the study by Shokrollahpour et al. (2016) used the slack-based model (SBM; Tone, 2001), which is an extension of the additive model Charnes et al. (1985), as the Russell measure (RM) that we use also is. The comparison between radial and non-radial DEA models that we perform in terms of their success as classification tools when combined with ANN is unique.

Our research can guide managers and researchers, as we investigate for first time a family of DEA models within the ANN framework. It is worth pointing out that although in conventional DEA studies the model selection with respect to returns-to-scale is done before the assessment, in our research we use a possible spectrum of returns-to-scale of some selected DEA models because we want to provide evidence of the behavior of our proposed DEA-ANN algorithm under the particular hypotheses of returns-to-scale. Moreover, in the same manner, we investigate other hypotheses such as the radial or non-radial DEA optimization principles using selected DEA models.

3. Methodology

We position the problem in view of Design Science Research Methodology (DSRM) for which we develop a simple artefact (i.e., a hybrid DEA-ANN model) that better contributes to the achievement of a goal. By ‘artefact’, we refer to a human-made object, usually developed for practical purposes Geerts (2011). Furthermore, artefacts should possess two essential characteristics: relevance and novelty (Geerts; 2011; Hevner et al., 2004) and the artefact that we propose in this paper has both. On the one hand, it is relevant, as it addresses the ongoing practical problem of the “best practice trap” in DEA, which may happen as a result of the fact that seeking the best performers may not always yield the best results. On the other hand, it is novel in the sense that, unlike previous attempts, we are interested only in the classification ability of our proposed DEA-ANN modeling and to this end we investigate a family of radial and non-radial DEA models within the ANN framework, which makes our effort unique.

In line with the above, the design problem (see, for example, Wieringa, 2014) can be formulated as follows: *Address the so called “best practice trap” in DEA by designing an approach (i.e., a hybrid DEA-ANN model) that is able to classify the sampled bank branches into predefined efficiency classes in order to support the decision-making process to achieve efficiency improvement.*

Peffers et al. (2008) introduced the Design Science Research Methodology (DSRM), consisting of a nominal sequence of activities to be followed in the process of creating an artefact; in

Table 1, we discuss the activities that are relevant in the context of the present study. The first column lists the DSRM activities, the second column describes each of these activities, and the third column pinpoints the materials from and through which the activities are executed, such as models, methods, and foundational theories, instruments and frameworks, among others (Hevner et al., 2004).

Table 1

Design Science Research Methodology (DSRM) applied to the Current Study

DSRM Activities	Activity Description	Knowledge
Problem identification and motivation	Need to address the so called "best practice trap" in DEA, which may happen as a result of the fact that seeking the best performers may not always yield the best results. Bank branches, however, may be motivated to learn not only from best performers (i.e., peers in the DEA context), but also from other better performers.	Literature review. Understanding of weaknesses of existing DEA models and benchmarking analysis. Real-world problem.
Define the objectives of a solution	Investigate whether benchmark analysis using DEA in tandem with ANN can support the decision-making process to achieve efficiency improvement.	Literature review. Knowledge of existing tools: DEA, ANN.
Design and development	Design of an approach (i.e., a hybrid DEA-ANN model, wherein DEA is used as a preprocessor) that is able to classify the sampled bank branches into predefined efficiency classes.	Radial and non-radial DEA models. Efficiency scores. ANN model.
Demonstration	Case study demonstration. The proposed hybrid DEA-ANN model is used to classify the sampled branches of a large Greek commercial bank into predefined efficiency classes.	Applying the proposed approach to a real-world case.
Evaluation	Comparative analysis.	Understanding of current solution and its advantages.

Note. The DSRM activities described here follow the arrangement proposed by Peffers et al. (2008) and Charles et al. (2019).

This study follows the above DSRM activities to develop a hybrid DEA-ANN model to classify 160 branches of a large Greek commercial bank into predefined efficiency classes. It should be noted that considering DEA in view of DSRM is a novel approach; to the best of our knowledge, only one other paper has done so Charles et al. (2019).

The proposed DEA-ANN system is developed through the following main steps: variable selection, network design, and network application process, which we proceed to explain in the following sections.

3.1. *Variable selection*

In this study, an ANN is employed to model the efficiencies of bank branches of The Bank. Data (for one year) that are likely to be considerably cleaner than standard banking data sets have been retrieved from the Management Information System (MIS) of The Bank and refer to a large sample of 160 branches which are operating all over Greece.

DEA models that have been used in bank branch performance have approached efficiency from a production, intermediation, and profitability perspectives. In this study, the profitability efficiency of the bank branches is assessed in the light of contrasting their operating cost with the monetary outcomes (i.e., incomes) that are generated by the branches. In this approach, it is examined how well different branches combine their resources (i.e., expenses) to produce revenues. In other words, profitability efficiency evaluates the ability of branches to minimize the cost of resources for the level of revenue generated from different activities. In the profitability efficiency assessment, the objective function of DEA models is the ratio of the weighted sum of revenues to the weighted sum of expenses, which is an indicator of profitability Tsolas (2010).

The input set for analysis consists of three inputs (xDEA):
xDEA1 = personnel expenses,
xDEA2 = rents and depreciation, and
xDEA3= operational expenses.
On the other hand, the output set is composed of two outputs (yDEA), namely:
yDEA1 = net interest income, and
yDEA2= non-interest income (fee and trading income).

Personnel expenses, rents and depreciation, and operational expenses are cost-related items that are used as inputs.

Net interest income (i.e., interest income from loans minus interest paid on customer's deposits) and non-interest income (i.e., fees from non-lending activities of branches, such as fee income and trading income) are revenue-related items that are used as outputs.

It should also be noted that the above inputs correspond to major cost items of bank branch operations. The output set of profitability assessment, on the other hand, includes only two items; nevertheless, these outputs account for a sufficiently large part of the total income of a bank branch, i.e., interests earned on loans (net interest income) and non-interest (gross) income (commissions and other non-interest income).

A slight deviation from the profitability efficiency model in relation to the specification of the output set can be found in the work by Oral & Yolalan (1990). Instead of using gross interest income with gross non-interest income in the output side of DEA to assess the efficiency of resource use in delivering income, origination income (net interest income) and non-interest (gross) income (commissions and other non-interest income) are considered as outputs in our profitability efficiency model.

The assessment of the profitability efficiency of the bank branches is based on their ability to generate short- and long-term profits. By short-term profitability, we refer to the income from commissions that branches generate and by long-term profitability, we indicate the income from the lending activity of the branches (see also Giokas, 2008).

For the purposes of ANN, a new variable, i.e., DEA efficiency, is obtained by means of input minimization using a family of models that includes radial models: CCR (Charnes et al., 1978), BCC (Banker et al., 1984), non-increasing returns-to-scale (NIRS) model (Byrnes et al., 1984), and non-decreasing returns-to-scale (NDRS) model (Byrnes et al., 1984); and non-radial models: free disposal hull (FDH) model (Deprins et al., 1984) and Russell measure (RM, Färe & Lovell, 1978) under constant, variable, non-increasing, and non-decreasing returns-to-scale.

Since the branches typically have little or no direct control over the amount of services their customers require, input orientation was chosen.

We use the inputs and outputs presented in Figure 1 to characterize the efficiency of the branches.

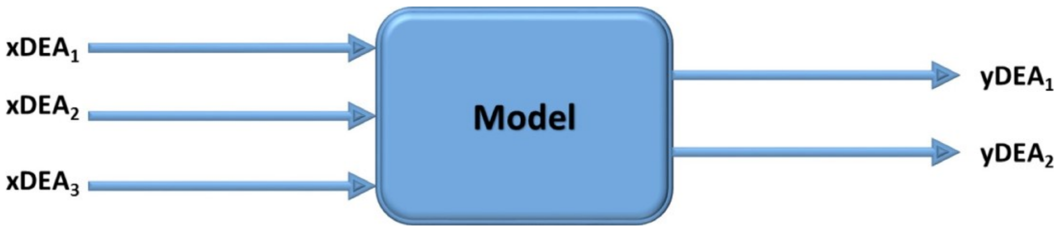

Fig. 1. DEA Model.

Given a set of n bank branches utilizing quantities of inputs $X \in R_{+}^{mxn}$ to produce quantities of outputs $Y \in R_{+}^{rxn}$, we use the radial and non-radial models presented in Table 2 to derive an efficiency score (≤ 1) based on a reference technology.

Table 2
Selected radial and non-radial DEA models.

EM	Hypothesis	Model	Reference Technology and Objective Function
Radial	CRS	CCR, Charnes et al. (1978)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, \Lambda \in R_+^n\}$ $Min\{\theta \mid (y, \theta x) \in S\}$
	VRS	BCC, Banker et al. (1984)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda = 1, \Lambda \in R_+^n\}$ $Min\{\theta \mid (y, \theta x) \in S\}$
	NIRS	NIRS, Byrnes et al. (1984)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda \leq 1, \Lambda \in R_+^n\}$ $Min\{\theta \mid (y, \theta x) \in S\}$
	NDRS	NDRS, Byrnes et al. (1984)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda \geq 1, \Lambda \in R_+^n\}$ $Min\{\theta \mid (y, \theta x) \in S\}$
Non-radial	FDH	FDH, Deprins, Simar, and Tulkens (1984)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, \Lambda = (\lambda_i \mid \forall i) = \{0, 1\}\}$ $Min\{\theta \mid (y, \theta x) \in S\}$
	CRS	Russell measure, Färe and Lovell (1978)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, \Lambda \in R_+^n\}$ $Min\{e_m' \Theta / m \mid (y, \Theta x) \in S\}$
	VRS	Russell measure, Färe and Lovell (1978)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda = 1, \Lambda \in R_+^n\}$ $Min\{e_m' \Theta / m \mid (y, \Theta x) \in S\}$
	NIRS	Russell measure, Färe and Lovell (1978)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda \leq 1, \Lambda \in R_+^n\}$ $Min\{e_m' \Theta / m \mid (y, \Theta x) \in S\}$
	NDRS	Russell measure, Färe and Lovell (1978)	$S = \{(x,y) : Y\Lambda \geq y, X\Lambda \leq x, e_n' \Lambda \geq 1, \Lambda \in R_+^n\}$ $Min\{e_m' \Theta / m \mid (y, \Theta x) \in S\}$

Note: EM - Efficiency Measure, CRS - constant returns-to-scale, VRS - variable returns-to-scale, NIRS - non-increasing returns-scale, NDRS - non-decreasing returns-to-scale, FDH - free disposal hull, θ - efficiency score, $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$, where θ_i is the score associated with i^{th} input, X_{mxn} - a matrix of m inputs, Y_{rxn} - a matrix of r outputs; x and y are vectors of inputs and outputs of the DMU of interest, respectively, $e_p' = (1, 1, \ldots, 1)$ with p elements, $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are intensity factors.

Bank branches are given an efficiency score (≤ 1) by the DEA model. They are classified into four classes, according to the efficiency score intervals in which their efficiency scores lie (Wu et al., 2006):
- Class (1): Efficiency score interval of (0.98, 1); referred to as 'strong relative efficient interval'.
- Class (2): Efficiency score interval (0.80, 0.98); referred to as 'relative efficient interval'.
- Class (3): Efficiency score interval (0.50, 0.80); referred to as 'relative inefficient interval'.
- Class (4): Efficiency score interval of (0, 0.50); referred to as 'very inefficient interval'.

It is to be noted, however, that one does not need to take a rigid approach towards the given efficiency classes, which can and should be altered in view of application and the requirements of the management teams and their targets. These classes can change even when switching from one type of bank to another (e.g., from commercial to investment banks); for an example of another practical classification, see Norman & Stoker (1991). The efficiency value intervals classes we have selected for use in the present paper are for illustrative purposes only, as our main objective is not to derive such classes, but to develop a hybrid DEA-ANN model that is able to classify the sampled bank branches into predefined efficiency classes.

3.2 Network design

The network design process deals with the determination of the network architecture, selection of the learning algorithm, and configuration of training and testing data sets.

3.2.1 Network architecture

An MLP network is preferred due to its simple architecture and the proven success of the model for solving approximation problems. An MLP network includes the following components: the input layer, one or more hidden layers, and an output layer which refers to the desired output of the system. It should be noted that the successful application of the MLP to any problem is related to the network architecture. The relevant theory states that networks with a single hidden layer only can provide more accurate results (Çelebi & Bayraktar, 2008).

Figure 2 shows a schematic representation of a back propagation of one input layer-one hidden layer-one output layer network with a 3-3-1 (3 input nodes - 3 hidden nodes - 1 output (with four classes) node structure. As shown in the figure, the whole iterative process includes real input presentation, feed (forward) information, error estimation, and back propagation of error for sequential weight adjustments (see also Kwon & Lee, 2015).

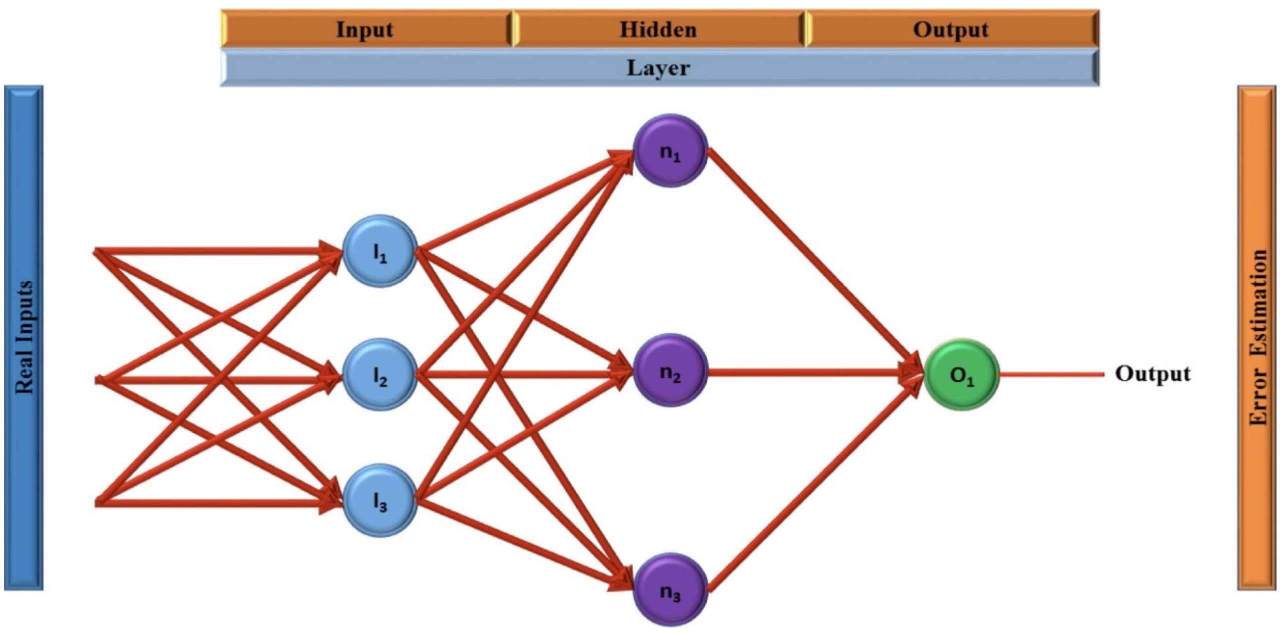

Fig. 2. Schematic representation of a 3:3:1 network.

In our case, as a base scenario, the MLP structure to construct the classification model and its input layer, hidden layer, and output layer are as follows:

- Input layer: it includes (a) six input values, which are the three inputs (xDEA1, xDEA2, xDEA3) and the two outputs (yDEA1, yDEA2) of the DEA model and (b) the DEA-derived efficiency score.
- Hidden layer: one hidden layer.
- Output layer: it includes one output value, which classifies the DEA scores according to the predefined efficiency intervals.

3.2.2 Learning algorithm

Having selected the architecture of the MLP, the connection weights of the network are estimated by means of a training procedure based on the selected training data set. Levenberg–Marquardt (LM), a well-known BP algorithm, has been selected as the learning algorithm for the training of the MLP due to its high level of accuracy and low level of complexity.

The process of the BP algorithm is as follows (Ahmad et al., 2017):

1. Present a training sample and propagate it through the neural network to derive the desired output.

2. Use small random and threshold values to initialize all weights.

3. Calculate the input to the j^{th} node in the hidden layer in line with Eq.(1)

$$zNET_j = \sum_{i=1}^{N} w_{ij}z_i - \tau_j, \tag{1}$$

where $zNET_j$ is the input to the j^{th} node in the hidden layer, w_{ij} is the weight from the i^{th} input node to the j^{th} hidden layer node, τ_j is the threshold value between the input and hidden layers, and z_i is the input, and N is the number of inputs.

4. Calculate the output from the j^{th} node in the hidden layer using Eqs. (2) and (3):

$$h_j = f_h\left(\sum_{i=1}^{N} w_{ij}z_i - \tau_j\right), \tag{2}$$

$$f_h(z) = \frac{1}{1 + e^{-\lambda_h z}}, \tag{3}$$

where h_j is the vector of hidden-layer neurons, $f_h(z)$ is the logistic sigmoid activation function, λ_h is the slope control variable of the sigmoid function and w_{ij}, z_i, and τ_j as above.

5. Calculate the input to the k^{th} node in the hidden layer using Eq.(4):

$$zNET_k = \sum_{j} w_{kj}z_j - \tau_k, \tag{4}$$

where $zNET_k$ is the input to the k^{th} node in the hidden layer, w_{kj} is the weight from the j^{th} hidden layer node to the k^{th} output layer node, τ_k is the threshold value between the hidden and output layers, and z_j is the input from the j^{th} hidden layer.

6. Calculate the output of the k^{th} node of the output layer using Eqs.(3) and (5):

$$y_k = f_k(\sum_j w_{kj} z_j - \tau_k), \tag{5}$$

where y_k is the output of the k^{th} node of the output layer, and $f_k(z)$, w_{kj}, z_j, and τ_k as above.

7. Use Eqs. (6) and (7) to calculate the errors from the output layer:

$$\delta_k = -(d_k - y_k) f_k, \tag{6}$$

$$f_k = y_k(1 - y_k), \tag{7}$$

where δ_k is the errors vector for each output neuron, d_k is the target activation of the output layer, f_k is the local slope of the node activation function for the output nodes, and y_k as above.

8. Use Eqs.(8) and (9) to calculate the errors from the hidden layers:

$$\delta_j = f_k \sum_{k=1}^{N} w_{kj} \delta_k, \tag{8}$$

$$f_h = h_j(1 - h_j), \tag{9}$$

where δ_j is the errors vector for each hidden layer's neuron, and f_k, w_{kj} and , δ_k as above.

9. Adjust the weights and thresholds in the output layer.

3.2.3 Configuration of the training and testing data set

In the network design, the data of 160 bank branches are used. The complete data set is divided into two subsets: the training and the testing data set. Hundred twenty-eight branches (i.e., 80% of all data) are randomly chosen for the training set and thirty two (i.e., 20% of all data) for the testing process. The DEA-ANN system design utilised in this study is shown in Figure 3.

3.2.4 Bingo and 1-Away Metrics

In this section, we define the percent hit rates, which we will use to measure the predictive performance of our neural network approach. The percent success rate is arguably the most intuitive measure of discrimination for predictive accuracy of classification problems (Sharda & Delen, 2006), with bigger values indicating better classification performance.

Let **C** be the confusion matrix (also known as classification matrix) representing q classes, which has the form:

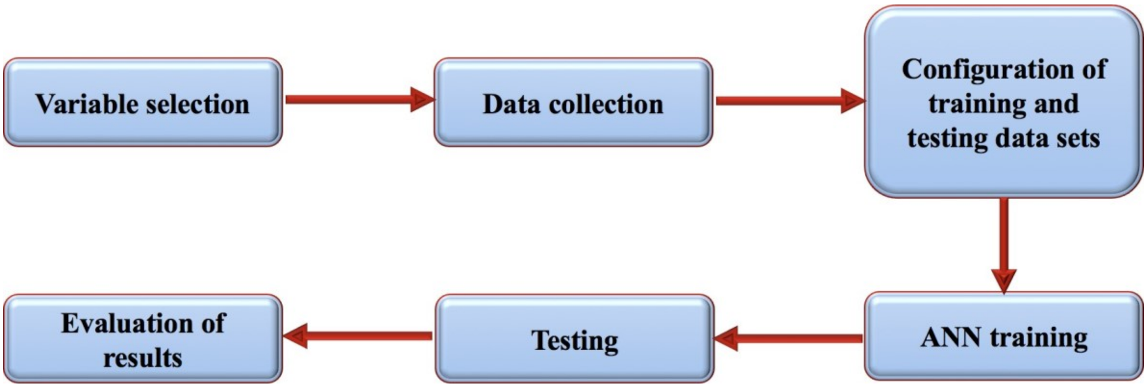

Fig. 3. The DEA-ANN system.

$$\mathbf{C} = \text{Actual} \begin{array}{c} \text{Predicted} \\ \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \vdots & & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qq} \end{bmatrix} \end{array} \tag{10}$$

where c_{lk} (wherein l is the row identifier and k is the column identifier) indicates the cases belonging to the l^{th} actual class that have been classified as the k^{th} predicted class. In this sense, the elements in the diagonal (c_{ll}) are the cases correctly classified, while the elements off the diagonal are the cases misclassified. In line with the above, we define four different hit rates: (a) two existent hit rates, which are the exact (dubbed "*Bingo*") hit rate and the "*OneAway* " class hit rate (see, for example, Zhang et al., 2009; Sharda & Delen, 2006); and (b) two newly-derived hit rates, the "*OneAway(above)*" hit rate; and the "*OneAway(below)*" hit rate, which as we shall see, can provide a more granular view of the results.

Bingo represents exact accuracy and only accounts for the correct classifications. It is defined as the ratio of sum of correct classifications to the total number of samples:

$$Bingo = \frac{\sum_{l=1}^{q} c_{ll}}{\sum_{l=1}^{q} \sum_{k=1}^{q} c_{lk}} \tag{11}$$

OneAway represents near-exact accuracy and is the within one class hit rate; it accounts for the cases correctly classified and the cases predicted into the adjacent classes (one class away from the diagonal). It is calculated as the ratio of the sum of total correct classifications and classifications predicted within one class to the total number of samples:

$$OneAway = \frac{\sum_{l=1}^{q-1}(c_{l+1l} + c_{ll+1}) + \sum_{l=1}^{q} c_{ll}}{\sum_{l=1}^{q}\sum_{k=1}^{q} c_{lk}} \tag{12}$$

OneAway(above) accounts only for the cases that have been predicted one class above the actual class; it is calculated as the ratio of the sum of classifications predicted just one class above the actual class to the total number of samples predicted above the actual class:

$$OneAway(above) = \frac{\sum_{l=1}^{q-1} c_{l+1l}}{\sum_{\substack{l=1 \\ l>k}}^{q}\sum_{k=1}^{q} c_{lk}} \tag{13}$$

OneAway(below) accounts only for the cases that have been predicted one class below the actual class; it is calculated as the ratio of the sum of classifications predicted just one class below the actual class to the total number of samples predicted below the actual class:

$$OneAway(below) = \frac{\sum_{l=1}^{q-1} c_{ll+1}}{\sum_{\substack{l=1 \\ l<k}}^{q}\sum_{k=1}^{q} c_{lk}} \tag{14}$$

4. Results

4.1. Descriptive statistics

In this section, we proceed to present and discuss the descriptive statistics. In this sense, Figure 4 and Figure 5 visually show the efficiency distribution of radial and non-radial models and the details of quartiles for radial and non-radial models, respectively.

Based on Figure 5, we can note that the mean efficiency of the models ranges from as low as 0.56 (CRS) to as high as 0.88 (FDH). The medians of the models range from 0.40 (CRS) to 0.82 (FDH). The standard deviation of models ranges from 0.13 (RM-VRS, RM-CRS) to 0.23 (NIRS). The efficiency scores of the non-radial models are distributed around the 0.37 to 1.00 (max) efficiency range; this is the case of the FDH model. Greater minimum values have been found: 0.43 for both the RM-CRS and RM-NIRS models, and 0.52 for both the RM-VRS and RM-NDRS models.

With regard to the efficiency scores distributions, as they stem from the density plot histogram (see Figure 4), we can observe the following: The FDH efficiency scores take the highest peak at 1.00 and the rest of the DMUs are distributed around the 0.37 to 0.99 efficiency range, with most of them getting scores higher than 0.6; also, it is clear from the box plot of FDH that the average efficiency is 0.88 ± 0.17, as well as that 50% of the units are falling within the 0.82 to 1.00 efficiency range. Regarding the peak of the distribution at 1.00, after FDH the RM-VRS follows and the rest of the DMUs are distributed around the 0.52 to 0.99 efficiency range. The average efficiency is 0.79 ± 0.13 and 50% of the units are falling within the 0.69 to

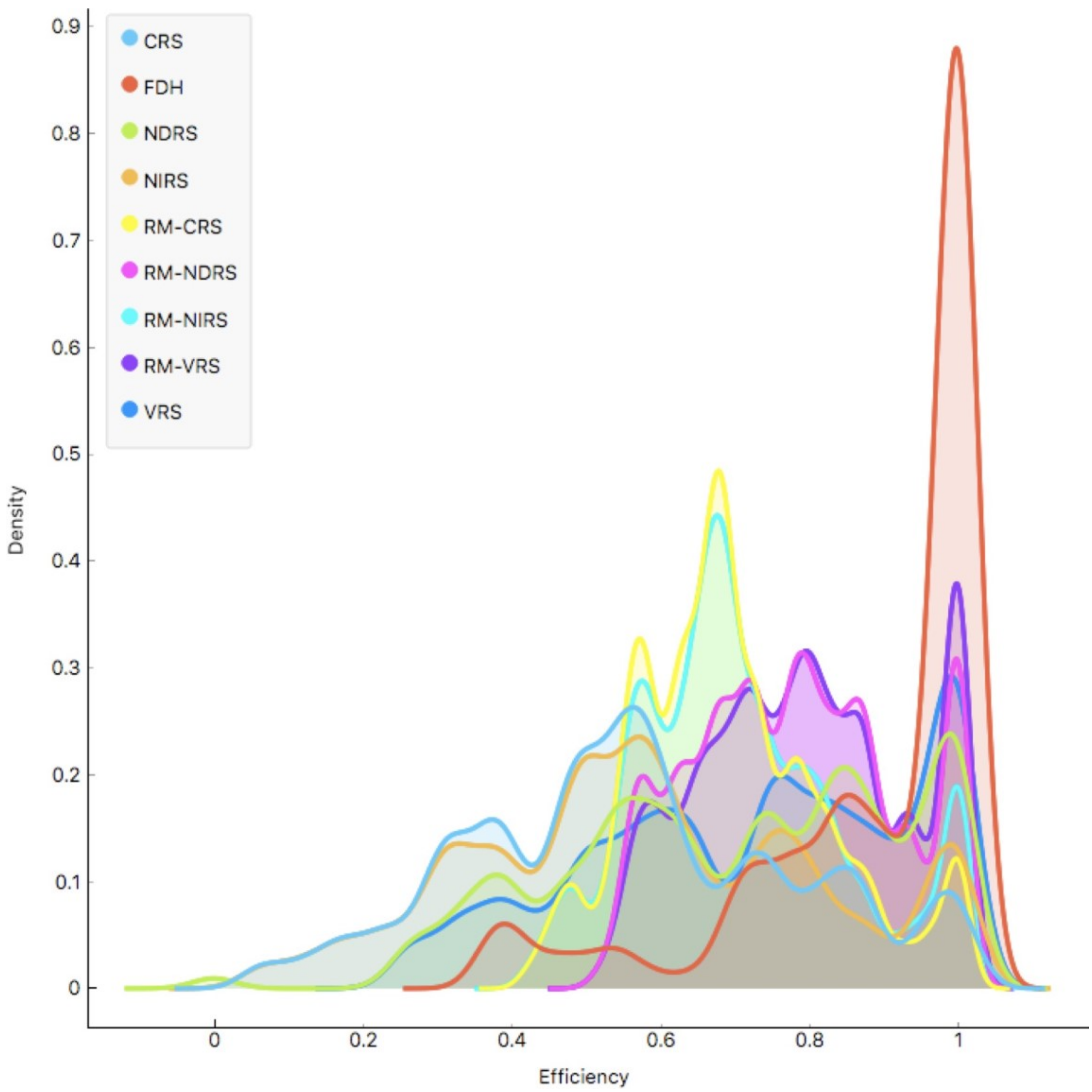

Fig. 4. Efficiency distribution of radial and non-radial models.

0.89 efficiency range.

The next model with the highest peak at 1.00 is the RM-NDRS model and the rest of the DMUs are distributed around the 0.52 to 0.99 efficiency range. The average efficiency is 0.78 ± 0.13 and 50% of the units are falling within the 0.68 to 0.88 efficiency range. The next model with the highest peak at 1.00 is the VRS conventional BCC model and the rest of the DMUs are distributed around the 0.25 to 0.99 efficiency range; this wider range (when compared to the previous models) may be due to the DMUs being of different sizes. The average efficiency is 0.72 ± 0.21 and 50% of the units are falling within the 0.56 to 0.91 efficiency range. Next in the rank with the highest peak at 1.00 is the RM-NIRS model and the rest of the DMUs are distributed around the 0.43 to 0.99 efficiency range. The average efficiency is 0.71 ± 0.14 and 50% of the units are falling within the 0.61 to 0.79 efficiency range.

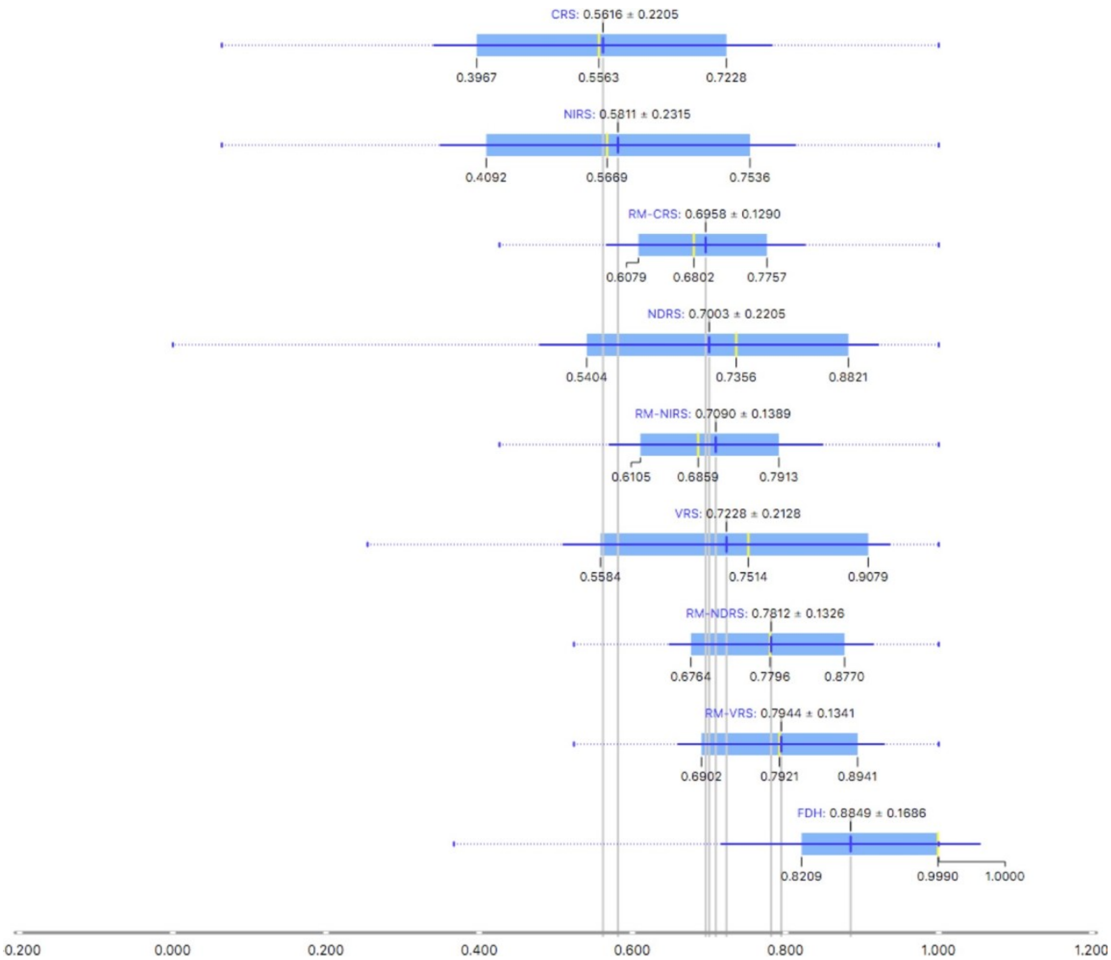

Fig. 5. Details of quartiles for radial and non-radial models.

The next model with the highest peak at 1.00 is the NDRS conventional model and the rest of the DMUs are distributed around the full spectrum of the efficiency range; this wider range (when compared to the previous models) may be due once again to the DMUs being of different sizes. The average efficiency is 0.70 ± 0.22 and 50% of the units are falling within the 0.54 to 0.88 efficiency range. Next in the rank with the highest peak at about 0.68 is the RM-CRS model and all the DMUs are distributed around the 0.43 to 1.00 efficiency range. The average efficiency is 0.70 ± 0.13 and 50% of the units are falling within the 0.61 to 0.78 efficiency range. Next in the rank with the highest peak at 1.00 is the NIRS conventional model and the rest of the DMUs are distributed almost entirely around the full spectrum of efficiency range; this wider range when compared to the previous models may be due, yet again, to the DMUs being of different sizes. The average efficiency is 0.58 ± 0.23 and 50% of the units are falling within the 0.41 to 0.75 efficiency range. Finally, the next model with the highest peaks at 0.46 and 0.55 is the CRS conventional CCR model, with all DMUs being distributed almost entirely around the full spectrum of the efficiency range; same as before, this wider range when compared to the previous models may be due to the DMUs being of different sizes. The average efficiency is 0.56 ± 0.22 and 50% of the units are falling within the 0.40 to 0.72 efficiency range.

4.2. *Classification Analysis*

The accuracy or diagnostic performance of the neural network systems is evaluated considering the following two situations: (1) without 'feeding' the DEA efficiency scores into the ANN, in other words, by running the ANN only based on inputs and outputs; (2) by 'feeding' the DEA efficiency scores into the ANN, along with the inputs and outputs.

Classification accuracy (CA) is the most commonly used index when evaluating the classification performance; it shows the proportion of data that were correctly classified out of all data. We furthermore report the following quantitative indicators: precision and recall (Fawcett, 2006). Precision can be thought of as a measure of a classifier's exactness, with a low precision indicating a large number of False Positives. On the other hand, Recall (also called Sensitivity) can be thought of as a measure of a classifier's completeness, with a low recall indicating many False Negatives. As sometimes precision and recall may be contradictory, we also employ the F1 Score (van Rijsbergen, 1979), calculated as the harmonic mean between precision and recall, for comprehensive consideration. Furthermore, we use the area under the receiver ROC curve (AUC) to evaluate the diagnostic performance. To assess the robustness of our approach, we will look at all these performance measures. The testing results of the ANN system based on inputs and outputs only are given in Table 3; and the testing results of the ANN system based on inputs, outputs, and DEA efficiency scores are reported in Table 4.

Table 3
Testing results (%) of the ANN system without DEA Scores.

EM	Technology	AUC	CA	F1	Precision	Recall
Radial	CRS	96.1	78.1	80.3	82.9	78.1
	VRS	92.3	78.1	77.8	81.4	78.1
	NIRS	98.4	90.6	89.8	93.3	90.6
	NDRS	95.9	78.1	78.3	80.7	78.1
Non-Radial	FDH	82.9	62.5	60.5	59.8	62.5
	RMCRS	94.9	93.8	94.7	97.0	93.8
	RMVRS	93.3	81.2	80.4	80.0	81.2
	RMNIRS	87.4	84.4	83.8	83.3	84.4
	RMNDRS	95.6	81.2	80.9	83.3	81.2

Note: EM - Efficiency Measure.

Table 4
Testing results (%) of the ANN system with DEA Scores.

EM	Technology	AUC	CA	F1	Precision	Recall
Radial	CRS	93.3	84.4	84.5	88.7	84.4
	VRS	96.1	90.6	89.8	91.4	90.6
	NIRS	94.8	81.2	83.7	86.5	81.2
	NDRS	99.3	90.6	90.9	93.2	90.6
Non-Radial	FDH	99.7	93.8	93.6	94.5	93.8
	RMCRS	100	93.8	94.3	97.9	93.8
	RMVRS	100	93.8	93.8	94.8	93.8
	RMNIRS	99.6	96.9	97.0	97.9	96.9
	RMNDRS	98.5	90.6	90.7	91.0	90.6

Note: EM - Efficiency Measure.

Overall, we can conclude that the proposed DEA-ANN approach is much better concerning the AUC, accuracy, F1-scores, precision, and recall, when compared to the ANN without DEA scores approach. As it can be observed from the performance measure values in Tables 3 and 4, values are generally higher when DEA scores are fed into the ANN system. There are two additional observations to make at this point. First, the performance measure values when using non-radial DEA models are generally higher than the performance measure values when using radial DEA models, independent of whether DEA efficiency scores are used as an input for the ANN system or not. Noteworthy, nonetheless, is that when DEA scores are used as an input, these measures when using non-radial DEA are higher in all the cases.

A second observation to be made is that when we 'feed' the DEA scores into the ANN system, the accuracy level improves significantly for all DEA technologies, with two exceptions: CRS and NIRS technologies. For CRS, the change is minimal, with only the AUC value decreasing from 96.1% to 93.3%. But in the case of NIRS, the values of all performance metrics decrease after inputting the DEA scores into the ANN. This points to the fact that in the context of our application, NIRS is the only DEA technology that when integrated with ANN, it does not improve the system's ability to classify the sampled branches into predefined efficiency classes.

This can further be appreciated from Figure 6, which depicts the type II errors for ANN with and without DEA scores. Lower type II errors ensures a higher level of correct classifications. Here, it can be noticed that 'feeding' the DEA scores to the ANN system results in lower type II errors; the only exception being posed by the NIRS technology.

We further use the percent hit rates defined in Section 3.2.4 to measure the predictive performance of our neural network approach. The classification (available and defined) results are presented in Table 5. In summary, the results show that 81.25% to 90.63% of the test sample for radial models and 90.63% to 96.88% of the test sample for non-radial models can be successfully classified according to their success performance; furthermore, most values are

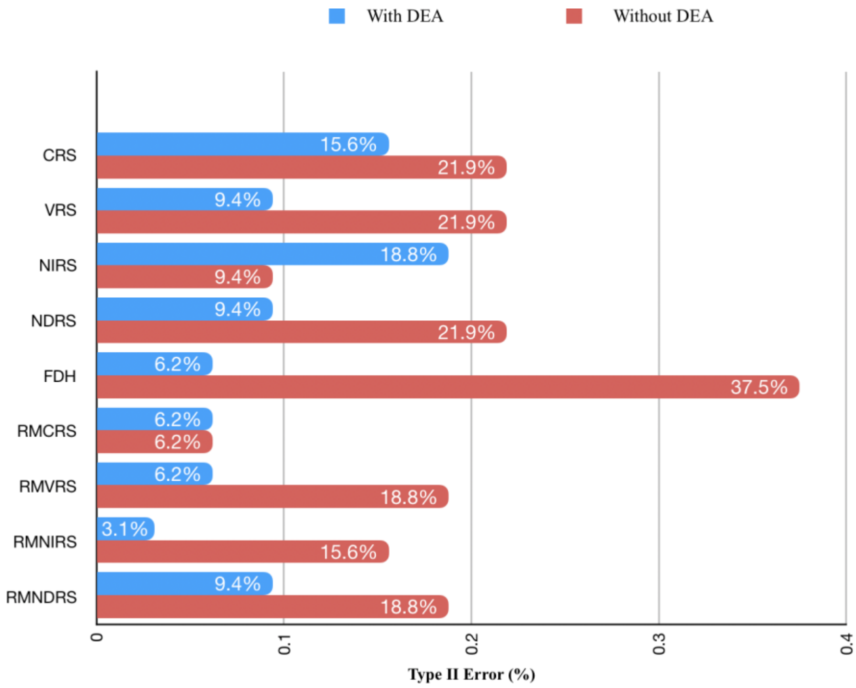

Fig. 6. Type II Errors for ANN with/without DEA Scores.

above 90%, which indicates a good classification performance. Overall, the best performers are: VRS and NDRS followed by the CCR model and the NIRS model for the radial models; and RMNIRS, followed by FDH, RMCRS, RMVRS and RMNDRS, respectively, for the non-radial models.

A well-designed ANN system can provide information about the efficiency category to which branches most probably belong if their success criterion is known. If a decision-maker has the data about a new candidate branch, it should be evaluated by using this structure and then the system results can be used for strategic decisions on bank branches. However, it is expected that by using this chosen network structure, 'efficient' branches of class (1) will be classified and predicted more successfully than 'inefficient' branches, as it can be observed from Table 5. This is also the case for branches of class (4) by employing non-radial models; this is because of their better discriminating power. In general, it can be noticed that ANN performs better when non-radial efficiency scores are 'fed' into the model, as opposed to radial efficiency scores.

Table 5
Classification results (%) of Bingo and OneAway.

		Classes					*OneAway*		
EM	Technology	(1)	(2)	(3)	(4)	Bingo	(below)	(above)	OneAway
Radial	CRS	*	75.00	73.33	100.00	84.38	80	–	96.88
	VRS	100	100	93.33	50.00	90.63	–	100	100
	NIRS	*	33.33	87.50	91.67	81.25	100	100	100
	NDRS	100	77.78	92.86	100	90.63	100	100	100
Non-Radial	FDH	100	77.78	100	100	93.75	–	100	100
	RMCRS	100	50.00	100	100	93.75	–	100	100
	RMVRS	100	100	88.24	*	93.75	–	100	100
	RMNIRS	100	75.00	100	100	96.88	–	100	100
	RMNDRS	100	90.00	88.89	*	90.63	100	100	100

Note: EM - Efficiency Measure. The * indicates that there is no bank (in the 20% test data) falling under the particular class, for the given technology. The - indicates that there is no misclassification.

Without a doubt, the correct classification of samples is an essential element of any study. Misclassification occurs when units of analysis are assigned to a different class than the one they should be in. This can lead to incorrect associations being observed between the assigned classes and the outcomes of interest. There are fields (such as healthcare) where misclassification poses a significant problem, as it can cause one to under- or over-estimate health risks, which can make a difference between life and death. But in the banking field, stakeholders might be glad to predict within one (or maybe even two classes) on either side. This observation is important as we can notice that our prediction results in view of *OneAway* indicate perfect (100%) classification accuracy hit rate, except for the case of CRS, which achieves only 96.88% correct classifications. Nevertheless, if we consider the accuracy within two classes, then even CRS achieves 100% hit rate. A further look at the *OneAway(below)* and *OneAway(above)* metrics highlights the fact that in the case of CRS, 80% of all the misclassified cases are misclassified in one class below the actual class.

5. Discussion and managerial implications

The managerial implications are quite straightforward. DEA determines the current profitability efficiency level of each bank branch of the sample and moreover, the best-in-class branches that attain the maximum performance. Nevertheless, the quest for best practices may lead to an increased risk of poor decision-making, since seeking the best performers may not always yield the best results, a situation that is commonly known as the "best practice trap". As such, in practice, branches may be motivated to learn not only from the best-in-class performers, but also from other better branch performers that lie in lower classes of performance.

Since the DEA as a sole method has some shortcomings, a hybrid DEA-ANN algorithm was suggested. After the development of this algorithm and the derivation of results for a selected

DEA model, the corresponding ANN model can be employed solely with a new data set by just using the pre-found ANN weights. The new data set will contain the initial data on branch inputs and outputs in the DEA context as well as the DEA-produced (ex post) efficiency scores and the new branch data on inputs and outputs and the predefined (ex ante) efficiency scores, e.g., equal to unity if we want the new banks to be ex ante fully efficient. The employment of the ANN model using a new data set provides consistent results. This is due to the fact that the ANN model will use the pre-found ANN weights to classify the new data. In contrast, the results produced by a DEA model are sample specific and thus, if new data are added to previous data to form an extended data set, this action is expected to cause inconsistencies with the previous results. These inconsistencies may occur when a new DEA assessment with an extended data set is performed, as DMUs previously found as efficient can be found as non-efficient and vice versa.

Managers can immediately benefit from the proposed approach. In practice, they should usually set achievable small targets and implement initiatives for performance improvement rather than look for big goals that may be impractical. The synthesized baseline establishment of branch classification through the combined use of DEA and ANN can aid managers find a better practice benchmarking configuration in the branch banking and beyond.

It is worth pointing out that although in conventional DEA studies the model selection with respect to returns-to-scale is done before the assessment, in our research we used a possible spectrum of returns-to-scale of some selected DEA models because we want to provide evidence on the behavior of our proposed DEA-ANN algorithm under the particular hypotheses of returns-to-scale. Moreover, in the same manner, we investigated other hypotheses such as the radial or non-radial DEA optimisation principles using selected DEA models.

Our rationale behind showing the base models with different technologies was to demonstrate the applicability of obtaining the knowledge (efficiency scores) from various DEA models, which could be fed into the ANN. Of course, when it comes to consulting projects, it is the duty of the consultant to select a model in conjunction with the decision-makers. As the technology is changing fast (i.e., computing technology is available and has become cost-effective, as well as computational power is not a problem anymore), one does not need to choose a model from the start, but rather one can use many models at the same time and afterwards compromise on a specific one in concern with the users of the results.

6. Conclusions

The primary purpose of this study was to investigate the potential capabilities of ANN in bank branch benchmarking, and for this task, a hybrid DEA-ANN model was presented with encouraging results. Taking advantage of the nature of these two methods, this paper presented an approach wherein DEA and ANN methods are used in tandem. In light of the results of this study, it can be concluded that overall the non-radial models seem to perform more successfully than the radial ones, although both radial and non-radial models classify more successfully the ’efficient’ branches of class (1) (i.e., the best performers).

To the best of the authors’ knowledge, this study is the first to adopt the profit-oriented approach to model branch profitability by incorporating families of radial and non-radial DEA models into the ANN framework and, moreover, to analyse the behavior of MLP with more than one hidden layer. This research explores the best practice benchmarking paradigm in the DEA context through the integration of DEA bank branch profit-efficiency measurement with ANN prediction models.

Moreover, this study may steer the theoretical development of DEA-ANN models by both academicians and practitioners towards new possibilities. Although there are a myriad of studies in the DEA literature, few researchers have examined this area by means of combining DEA with ANN, especially in the banking industry. The findings of our analysis are based on real data gathered by the MIS of a Greek bank, and as such may be useful for generalizing the DEA-ANN modeling to other banks in the same or different countries. As previously mentioned, the addition of ANN to the performance measurement framework may enhance the decision-making process for bank branch efficiency improvement. Avenues for future research include the deployment of the proposed model into the cloud-based service, in an attempt to assist the decision-makers in making informed or guided decisions.

Acknowledgments

The authors would like to thank the Editor and the three anonymous reviewers for their valuable feedback on the previous version of this manuscript.

References

Agarwal, R., Green, R., Brown, P. J., Tan, H., & Randhawa, K. (2013). Determinants of quality management practices: an empirical study of New Zealand manufacturing firms. International Journal of Production Economics, 142(1), 130–145.

Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy and Buildings, 147, 77–89.

Angelidis, D., & Lyroudi, K. (2006). Efficiency in the Italian banking industry: Data envelopment analysis and neural networks. International Research Journal of Finance and Economics, 1(5), 155–165.

Athanassopoulos, A. D., & Curram, S. P. (1996). A comparison of data envelopment analysis and artificial neural networks as tools for assessing. Journal of the Operational Research Society, 47(8), 1000–1016.

Azadeh, A., Ghaderi, S. F., Tarverdian, S., & Saberi, M. (2007). Forecasting electrical consumption by integration of neural network, time series and ANOVA. Applied Mathematics and Computation, 186(2), 1753–1761.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Models for estimating technical and scale efficiencies in Data Envelopment Analysis. Management Science, 30(9), 1078–1092.

Byrnes, P., Färe, R., & Grosskopf, S. (1984). Measuring Productive Efficiency: An Application to Illinois Strip Mines. Management Science, 30(6), 671–681.

Ç elebi, D., & Bayraktar, D. (2008). An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. Expert Systems with Applications, 35(4), 1698–1710.

Charles, V., Aparicio, J., & Zhu, J. (2019). The curse of dimensionality: A simple approach to increase the discriminatory power of data envelopment analysis. European Journal of Operational Research, 279(3), 929-940.

Charles, V., Peretto, C., & Gherman, T. (2016). Technical Efficiency Analytics for the Public Banks in Argentina. ICPE Public Enterprise Half-Yearly Journal, 22(1), 118-139.

Charles, V., Tsolas, I. E., & Gherman, T. (2018). Satisficing data envelopment analysis: A Bayesian approach for peer mining in the banking sector. Annals of Operations Research, 269(1-2), 81-102.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. Journal of Econometrics, 30, 91–l07.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429–444.

Costa, A., & Markellos, R. N. (1997). Evaluating public transport efficiency with neural network models. Transportation Research, 5(5), 301–312.

Davies, A. J., & Kochhar, A. K. (2002). Manufacturing best practice and performance studies: a critique. International Journal of Operations Production Management, 22(3), 289-305.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring Labour Efficiency in Post Offices. In M. Marchand, P. Petieau, H. Tulkens (eds), The Performance of Public Enterprises: Concepts and Measurement. Amsterdam: North Holland.

Emrouznejad, A., & Shale, E. (2009). A combined neural network and DEA for measuring efficiency of large scale data sets. Computers Industrial Engineering, 56(1), 249–254.

Färe, R., & Lovell, C. A. K. (1978). Measuring the technical efficiency of production. Journal of Economic Theory, 19, 150–162.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27, 861–874.

Francis, G., & Holloway, J. (2007). What have we learned? Themes from the literature on best-practice benchmarking. International Journal of Management Reviews, 9(3), 171–189.

Gaganis, C., Liadaki, A., Doumpos, M., & Zopounidis, C. (2009). Estimating and analyzing the efficiency and productivity of bank branches. Evidence from Greece. Managerial Finance, 35(2), 202–218.

Gallant, A., & Stephen, I. (1993). Neural network learning and expert systems. MIT Press.

Geerts, G. L. (2011). A design science research methodology and its application to accounting information systems research. International Journal of Accounting Information Systems, 12, 142-151.

Giokas, D. I. (2008). Assessing the efficiency in operations of a large Greek bank branch network adopting different economic behaviors. Economic Modelling, 25(3), 559–574.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75-105.

Hsiang-Hsi, L., Tser-Yieth, C., Yung-Ho, C., & Fu-Hsiang, K. (2013). A comparison of three-stage DEA and artificial neural network on the operational efficiency of semi-conductor firms in Taiwan. Modern Economy, 4(1), 20–31.

International Quality Study. (1993). Best Practices Report: An Analysis of Management Practices that Impact Performance. New York, NY: American Quality Foundation and Ernst Young.

Kheirkhah, A. A., Azadeh, A. A., Saberi, M. M., Azaron, A. A., & Shakouri, H. H. (2013). Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis. Computers Industrial Engineering, 64(1), 425–441.

Kuo, R. J., Wang, Y. C., & Tien, F. C. (2010). Integration of artificial neural network and MADA methods for green supplier selection. Journal of Cleaner Production, 18(12), 1161–1170.

Kwon, H.-B., & Lee, J. (2015). Two-stage production modeling of large U.S. banks: A DEA-neural network approach. Expert Systems with Applications, 42, 6758–6766.

Kwon, H.-B., Lee, J., & Roh, J. J. (2016). Best performance modeling using complementary DEA-ANN approach. Application to Japanese electronics manufacturing firms. Benchmarking: An International Journal, 23(3), 704–721.

Kwon, H.-B., Roh, J. J., & Miceli, N. (2016). Better practice prediction using neural networks: an application to the smartphone industry. Benchmarking: An International Journal, 23(3), 519–539.

Liu, J. S., Lu, L. Y., Lu, W., & Lin, B. Y. (2013). A survey of DEA applications. Omega, 41(5), 893–902.

Misiunas, N., Oztekin, A., Chen, Y., & Chandra, K. (2016). DEANN: A healthcare analytic methodology of data envelopment analysis and artificial neural networks for the prediction of organ recipient functional status. Omega, 58, 46–54.

Mostafa, M. (2009). A probabilistic neural network approach for modeling and classifying efficiency of GCC banks. International Journal of Business Performance Management, 11(3), 236–258.

Norman, M., & Stoker, B. (1991). Data Envelopment Analysis: The Assessment of Performance. Chichester, NY: John Wiley Sons.

Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. European Journal of Operational Research, 46, 282–294.

Ozdemir, D., & Temur, G. T. (2009). DEAANN approach in supplier evaluation system. World Academy of Science, Engineering and Technology, 54, 343–348.

Paradi, C. J., & Schaffnit, C. (2004). Commercial branch performance evaluation and results communication in a Canadian bank - a DEA application. European Journal of Operational Research, 156(3), 719–735.

Paradi, J. C., Vela, S., & Yang, Z. (2004). Assessing bank and bank branch performance - Modeling considerations and approaches. In W. W. Cooper, L. M. Seiford, & J. Zhu (eds.), Handbook on Data Envelopment Analysis (pp. 349-400). New York, NY: Kluwer Academic Publishers.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77.

Pendharkar, P. C. (2011). A hybrid radial basis function and data envelopment analysis neural network for classification. Computers Operations Research, 38(1), 256–266.

Pendharkar, P. C., & Rodger, J. A. (2003). Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption. Decision Support Systems, 36(1), 117–136.

Samoilenko, S., & Osei-Bryson, K. M. (2010). Determining sources of relative inefficiency in heterogeneous samples: Methodology using Cluster Analysis, DEA and Neural Networks. European Journal of Operational Research, 206(2), 479–487.

Santin, D., Delgado, F. J., & Valino, A. (2004). The measurement of technical efficiency: a neural network approach. Applied Economics, 36(6), 627–635.

Seiford, L. M., & Zhu, J. (2003). Context-dependent data envelopment analysis – measuring attractiveness and progress. Omega, 31(5), 397–480.

Sharda, R., & Delen, D. (2006). Predicting box-office success of motion pictures with neural networks. Expert Systems with Applications, 30, 243–254.

Shokrollahpour, E., Hosseinzadeh Lotfi, F., & Zandieh, M. (2016). An integrated data envelopment analysis artificial neural network approach for benchmarking of bank branches. Journal of Industrial Engineering International, 12(2), 137–143.

Sreekumar, S., & Mahapatra, S. S. (2011). Performance modeling of Indian business schools: A DEA-neural network approach. Benchmarking: An International Journal, 18(2), 221–239.

Tkáč, M., & Verner, R. (2016). Artificial neural networks in business: Two decades of research. Applied Soft Computing, 38, 788–804.

Tone, K. (2001). A slacks based measure of efficiency in Data Envelopment Analysis. European Journal of Operational Research, 130, 498–509.

Tsolas, I. E. (2010). Modeling bank branch profitability and effectiveness by means of DEA. International Journal of Productivity and Performance Management, 59(5), 432–451.

Tsolas, I. E. (2011). Bank branch-level DEA to assess overall efficiency. EuroMed Journal of Business, 6(3), 359–377.

Van Rijsbergen, C. J. (1979). Information retrieval (2nd edition). London: Butterworths.

Vercellis, C. (2009). Business Intelligence: Data Mining and Optimization for Decision Making. Chichester, UK: John Wiley Sons.

Wang, S. (2003). Adaptive non-parametric efficiency frontier analysis: a neural-network-based model. Computers & Operations Research, 30(2), 279–296.

Wieringa, R. J. (2014). Design Science Methodology for Information Systems and Software Engineering. New York, NY: Springer-Verlag.

Wu, D. (2009). Supplier selection: A hybrid model using DEA, decision tree and neural network. Expert Systems with Applications, 36(5), 9105–9112.

Wu, D., Yang Z., & Liang, L. (2006). Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. Expert Systems with Applications, 31(1), 108–115.

Zhang, L., Luo, J., & Yang, S. (2009). Forecasting box office revenue of movies with BP neural network. Expert Systems with Applications, 36, 6580-6587.