Engineering Applications of Artificial Intelligence 127 (2024) 107222

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A data-driven decision support framework for DEA target setting: an explainable AI approach

Mustafa Jahangoshai Rezaee [a], Mohsen Abbaspour Onari [b,c,*], Morteza Saberi [d]

[a] Faculty of Industrial Engineering, Urmia University of Technology, Iran
[b] Information Systems Group, Eindhoven University of Technology, the Netherlands
[c] Eindhoven Artificial Intelligence Systems Institute, the Netherlands
[d] School of Information, Systems and Modelling, University of Technology Sydney, Australia

ARTICLE INFO

Keywords:
Data envelopment analysis
Benchmarking
Target setting
Explainable artificial intelligence
LIME
Multi-objective counterfactual explanation

ABSTRACT

The intention of target setting for Decision-Making Units (DMUs) in Data Envelopment Analysis (DEA) is to perform better than their peers or reach a reference efficiency level. However, most of the time, the logic behind the target setting is based on mathematical models, which are not achievable in practice. Besides, these models are based on decreasing/increasing inputs/outputs that might not be feasible based on DMU's potential in the real world. We propose a data-driven decision support framework to set actionable and feasible targets based on vital inputs-outputs for target setting. To do so, DMUs are classified in their corresponding Efficiency Frontier (EF) levels based on multiple EFs approach and a machine learning classifier. Then, the vital inputs-outputs are determined using an Explainable Artificial Intelligence (XAI) method. Finally, a Multi-Objective Counterfactual Explanation is developed based on DEA (MOCE-DEA) to lead DMU in reaching the reference EF by adjusting actionable and feasible inputs-outputs. We studied Iranian hospitals to evaluate the proposed framework and presented two cases to demonstrate its mechanism. The results show that the performance of the DMUs is improved to reach the reference EF for studied cases. Then, a validation was conducted with the primal DEA model to show the robust improvement of DMUs after adjusting their original value based on the generated solutions by the proposed framework. It demonstrates that the adjusted values can also improve DMUs' performance in the primal DEA model.

1. Introduction

Data Envelopment Analysis (DEA) assesses the relative efficiencies of a homogeneous set of Decision-Making Units (DMUs) with multiple inputs-outputs based on the non-parametric Linear Programming (LP) technique (Lim et al., 2011). The DEA is an appropriate technique to help organizations or firms evaluate DMUs and allocate resources to utilize the organizational strategies and objectives (Lai et al., 2011). Therefore, DEA is a decision support tool that can be implemented for management monitoring, planning, and control (Ramón et al., 2018a). The DEA has been proven to be a robust approach for evaluating performance and benchmarking to ameliorate organizations' operations. Consequently, DEA is utilized as a benchmarking technique to generate a Performance Score (PS) indicating the relative distance of a unit to the best-practices to compare with its equivalent peers (Shao et al., 2018). Although DEA is primarily a diagnostic tool, it does not prescribe any reengineering strategies to turn inefficient units into efficient ones (Lim et al., 2011). Generally, DEA determines a target operating point on the Efficient Frontier (EF), which indicates the amount of input reduction and output increment required for the unit under evaluation to become efficient (Lozano et al., 2020). However, the feasibility of reducing or incrementing input and output is not considered in the real world. Besides, adjusting the selected input-output may not be possible, and this important factor should be considered. Defining actionable targets with the least adjustment can provide valuable information for managers to have policies for the future and relieve the organization's shortcomings. Since almost all benchmarking studies utilize mathematical models, they can barely consider the units' capabilities.

According to the above-mentioned information, challenges arising for an inefficient DMU in benchmarking reference targets can be summarized into three aspects. First, the reference target might be a hypothetical DMU that does not exist. It is difficult and unrealistic to learn

* Corresponding author. P.O. Box 513, NL-5600, MB, Eindhoven, the Netherlands.
E-mail addresses: m.jahangoshai@uut.ac.ir (M. Jahangoshai Rezaee), m.abbaspour.onari@tue.nl (M. Abbaspour Onari), Morteza.Saberi@uts.edu.au (M. Saberi).

https://doi.org/10.1016/j.engappai.2023.107222
Received 26 September 2022; Received in revised form 26 September 2023; Accepted 27 September 2023
Available online 10 October 2023

from a DMU that does not exist because, in practice, it is not feasible to adjust inputs-outputs to reach the hypothetical DMU (Lim et al., 2011). Second, it is not easy to benchmark multiple best-practice DMUs in the reference set simultaneously. The presence of multiple efficient DMUs in the reference set of an inefficient DMU creates a perplexing situation, as the inefficient DMU has multiple targets for benchmarking (Lim et al., 2011). Third, achieving the desired efficiency in a single step for an inefficient DMU will likely be practically infeasible. In other words, if the inefficient DMU is significantly distant from the EF, achieving the frontier in one step becomes unattainable. Instead, a more practical approach would involve making stepwise gradual improvements towards the target (Lim et al., 2011). To the best of our knowledge, none of the research in the DEA literature considered: first, setting targets based on the realistic potential of DMUs. This critical factor is missing in the literature because the reference target should be its counterpart. Without considering this factor, the DMU is compared with a reference target far away from it, and in practice, it is infeasible to reach it. Second, identifying the vital input-outputs that can positively contribute to target setting. In real-world problems, adjusting some inputs-outputs might be impossible because, potentially, it is infeasible. Hence, we need to determine inputs-outputs that are adjustable. Combining DEA and Machine Learning (ML) algorithms can break through such limitations by adjusting inputs-outputs based on actual needs. Besides, according to data selection, different levels of EFs can be defined in actual decision-making to meet the needs of multi-level development (Zhong et al., 2021a).

ML systems are implemented to identify objects in images, transcribe speech into text, match news items, posts, or products with users' interests, and select relevant search results (LeCun et al., 2015). ML models became very popular in various fields, such as forecasting weather (Chakraborty et al., 2016a), construction cost estimation (Chakraborty et al., 2016b), composite flooring systems cost estimation (Elhegazy et al., 2022), prediction of optical properties of water (El-Ssawy et al., 2022), analyzing COVID-19 cases (Kwekha-Rashid et al., 2023), radar active jamming signal classification (Zhu et al., 2023), bearing fault diagnosis (Tang et al., 2023). However, many ML models have a black-box nature, which might have too many parameters or be proprietary. Consequently, they cannot explain their predictions in an understandable human way. In such scenarios, users might struggle to understand a model's outputs enough to trust and use its predictions (Balagopalan et al., 2022). To relieve the hazard of creating and implementing AI systems, eXplainable Artificial Intelligence (XAI) creates a suite of ML techniques that (i) generate more explainable models while maintaining a high level of learning performance and (ii) enable humans to understand and trust appropriately, and effectively manage the emerging generation of artificially intelligent partners (Arrieta et al., 2020). To fill the gap of setting targets based on the realistic potential of DMUs and identifying the vital input-outputs to do so, a data-driven decision support framework is proposed in this study using XAI methods. The motivation for using XAI in this framework can be summarized as (i) a faster decision-making process, (ii) considering much more decision-making criteria (Radovanović et al., 2022), and providing transparency and realism for Decision-Makers (DMs) in the target setting.

The proposed target setting framework in this study builds on the work of Seiford and Zhu (2003) to generate multiple EF levels according to the primal DEA model introduced by Charnes et al. (1978) (CCR). The corresponding EF level is assigned as a label for each unit. A classifier with high predictive performance with corresponding labels is applied to the original dataset to identify the potential inputs-outputs (features) for target setting. Then, the intended DMU for the target setting is named as observed DMU (DMU_o). Afterward, its peer DMU on the reference EF is found based on the least Euclidean distance and named prospective DMU (DMU_p). A post-hoc explainable method is implemented to obtain each feature's contribution to the classification's outcome for DMU_o and DMU_p. Local Interpretable Model-agnostic Explanations (LIME) is a model-agnostic method that suits this approach in finding features' contributions for units. Afterward, dominant features for target setting are determined among features that positively contribute to the prediction's outcome. Then, a modified DEA model based on Multi-Objective Counterfactual Explanation (MOCE-DEA) is proposed according to the CCR dual model. This model seeks to answer the following question: "How much adjustment is needed for DMU_o to reach reference EF?" To do so, it is attempted to reduce the distance of actionable features between DMU_o and DMU_p as well as maximize the Efficiency Score (ES) simultaneously. This objective makes the problem a non-linear and an NP-hard problem. The main advantage of this non-linear model is to consider multiple features in the target setting process simultaneously, making it generalizable for different problems. The proposed model is solved based on Multi-Objective Particle Swarm Optimization (MOPSO) to obtain the Pareto solutions. Finally, to validate the performance of MOCE-DEA, the robust improvement of DMU_o is evaluated based on the primal CCR model. This approach for target setting provides four main contributions to the problem:

- First, due to the lack of knowledge and ad hoc approaches taken by many firms and managers when implementing benchmarking, there is a need for research on how analysts and managers determine reference sets for conducting competitive analyses (Baek and Lee, 2009). The proposed approach solves this problem by generating multiple EF levels and using the least Euclidean distance to find the peer DMU for benchmarking on reference EF.
- Second, considering multiple EF levels and finding the peer DMU for target setting provides a realistic procedure that relieves the problem of defining a hypothetical DMU and stepwise gradual improvements.
- Third, determining actionable inputs-outputs for target setting by LIME presents the feature importance by extracting the contribution of each feature to the ML model's prediction. This approach determines the dominant features for target setting, and there is no need to adjust all inputs-outputs to improve performance. Here, we reduce the complexity of computation by selecting only vital inputs-outputs for target setting.
- Fourth, according to counterfactual logic, only feasible inputs-outputs for adjustment will be used for target setting. Besides, MOCE-DEA will generate solutions for DMU_o to reach the reference EF by the minimum possible adjustment.

The rest of the paper has been organized as follows: In Section 2, the related works of the implemented methods are covered. In Section 3, the implemented methods in the current research are presented. The proposed target setting framework is elaborated in Section 4. In Section 5, the proposed framework is applied to a real-world case study. In Section 6, the findings of this research are presented. Finally, in Section 7, the conclusion and suggestions for future studies are discussed.

2. Related works

This section has been divided into two sub-sections covering the related works conducted for the developed target setting framework. Subsection 1 provides the application of ML methods in the DEA domain. Subsection 2 covers some state-of-the-art approaches for distance-based target setting.

2.1. Application of ML methods in DEA

ML is a form of applied statistics with increased emphasis on implementing computers to estimate complicated functions statistically (Goodfellow et al., 2016). Cielen et al. (2004) argued that DEA is a type of ML technique and can provide analytical support for decisions based on intelligent data analysis (Li et al., 2017). As a data-enabled performance evaluation technique, DEA is useful in various fields, supporting decision-making worldwide (Ebrahimi et al., 2022). So far, various ML methods have been implemented to overcome DEA shortcomings or improve its performance. A Self-Organizing Map (SOM) clustering method was implemented by Hong et al. (1999) to overcome DEA shortcomings in the lack of offering no guidelines on where relatively inefficient DMUs can improve. The efficiency prediction model proposed by Zhang (Zhang and Wang, 2019) combines information granulation and Support Vector Machine (SVM) with the DEA model to evaluate the future efficiency of decision-making over time series data. A multi-stage model-based DEA and Radom Forest (RF) was proposed by Nandy and Singh (2020) to examine and predict the impact of environmental variables on farms' performance by extracting the crucial variables in prediction with RF. To address the typical rule of thumb issue used in DEA, Lee and Cai (2020) proposed a Least Absolute Shrinkage and Selection Operator (LASSO) variable selection technique as a DEA estimator. According to the previous study, Chen et al. (2021) revisited this approach, explored a more advanced version of LASSO, the so-called Elastic Net (EN) approach, and adapted it to DEA. To tackle DEA's traditional weaknesses of being easily affected by statistical noise in data and remeasuring its performance when new evaluation units are added, Zhong et al. (2021b) applied ML algorithms. Valero-Carreras et al. (2022) showed that Free Disposal Hull (FDH) and DEA could be cases of a more general model-based SVR within ML. Also, Valero-Carreras et al. (2021) introduced denominated Support Vector Frontiers (SVF) to estimate production functions, which allows the translation of certain notions of this ML technique into the efficiency measurement world for extra returns. To determine supper-efficiency in the context of the FDH technique, Esteve et al. (2023) adapted RF to differentiate between the performance of observations. Hatamzad et al. (2022) proposed an intelligent methodology using DEA and ML prediction techniques to achieve efficient and effective winter road maintenance on the roads during winter.

2.2. Distance-based target setting approaches for DEA

DEA has been proven to be a practical technique for efficiency measurement and target setting by identifying benchmarks (Tsolas et al., 2020). Many attempts have been conducted in DEA literature to represent a better target setting approach that is more realistic and achievable according to the reducing distance between inefficient and efficient DMUs. Aparicio et al. (2017) claimed closer targets determine less demanding operation levels for the inefficient units' inputs/outputs to perform efficiently. Subsequently, they proposed a general approach to finding the closest targets for a given unit according to the closeness between the inputs/outputs and the proposed targets using different distance functions or efficiency measures. Baek et al. (Baek and Lee, 2009) used the least-distance measure to the strongly efficient production frontier to obtain the shortest projection from the evaluated DMU, allowing an inefficient DMU to explore the easiest way to improve its efficiency. To satisfy strong monotonicity over the strength EF, Fukuyama et al. (2014) developed and extended the least distance $\rho-$ *norm* inefficiency measures for a free disposable set and introduced a trade-off set that implements input/output substitutability. A DEA-based benchmarking approach proposed by Ruiz and Sirvent (2016) for identifying a common Best Practice Frontier (BPF) as the facet of DEA Efficient Frontier (DEA-EF) spanned by the technically efficient DMUs in a common reference group. An efficiency measure developed by Zhu et al. (2018) based on non-oriented closest targets that satisfy strong monotonicity and that is calculated by a simple Mixed-Integer Linear Programming (MILP). A two-step procedure was introduced by Ramón et al. (2018b) using models that minimize the distance to the frontier by setting more realistically achievable targets on an intermediate frontier formed by more similar units in the level of performance. An approach proposed by Ruiz and Sirvent (2019) to minimize the distance to the DEA strong EF to incorporate goal information that adjusts the DEA benchmarking to the goals to consider the improvements policy that was pursued with setting such goals. A novel procedure was designed by Ramón et al. (2020) to make reference set selection (as defined in DEA), establishing the common benchmarking framework. Then, benchmarking models are formulated to set the closest targets relative to the reference sets selected. Le et al. (2021) extended the theory of inverse frontier-based benchmarking to enhance the conventional DEA efficiency measurement approach by focusing on the derivation of meaningful benchmarks – in terms of the required level of inputs/outputs – to improve performance. The closest target model was developed by An et al. (2021) for a two-stage system by constructing the efficient frontier in that all extremely efficient stages of the DMUs are considered to form the closest target for an inefficient DMU. A minimum distance directional slack inefficiency model developed by Fukuyama et al. (2022) uses the minimum distance approach to the bank efficiency measurement and treats NPLs under the costly-disposability framework. A MILP proposed by Zhu et al. (2022) to determine the extended efficient facets based on minimizing the slacks of outputs for an inefficient unit that would guarantee the reference point lies exactly on the full-dimensional efficient facets.

Studying the research in the literature shows that our proposed framework has two main advantages compared to other works. Even though the literature stresses that finding the reference target with the least distance is an efficient way to benchmark, most do not consider the stepwise target setting to compare units with the reference target in the closest EF. Further, the inputs-outputs for target setting are not selected based on the nature of the problem in these studies. Instead, first, the mathematical model is developed, and based on its assumptions, the adjustments are proposed for inputs-outputs. Here, most likely, adjusting inputs-outputs is not possible in practice. In our developed framework, we determine the feasible inputs-outputs in the real-world application for adjustment; then, their targets are set accordingly.

3. Preliminaries

This section covers the implemented methods in the current study. In Subsection 1, the DEA model and benchmarking will be presented. Then, in Subsection 2, the concept of the LIME will be provided. Finally, in Subsection 3, the fundamentals of the MOCE will be presented.

3.1. DEA and benchmarking

For the first time, Charnes et al. (1978) introduced the DEA with multiple inputs-outputs for assessing the performance of a set of homogeneous DMUs. DEA classifies DMUs into two mutually exclusive and collectively exhaustive groups (efficient and inefficient), implementing LP and measuring the PS of each DMU (Khezrimotlagh et al., 2019). An efficient DMU means that no other DMU, by consuming fewer inputs, can produce the same outputs, known as the input-orientated approach, or by consuming the same inputs, can produce more outputs, known as the output-orientated approach (Yang et al., 2009). DEA models vary by considering constant or variable returns to scale and have been known as CRS and VRS models, respectively (Rezaee et al., 2018). The CRS model is a Non-Linear fractional Programming (NLP) model like the original DEA model. The objective function in the model for a particular DMU maximizes the single ratio of the weighted outputs over-weighted inputs, referred to as an observed DMU and denoted by DMU_o (Yang et al., 2009). Suppose there are n DMUs ($j = 1,\ldots,n$) for evaluation. Each DMU consumes varying amounts of m inputs, denoted by x_{ij} (the ith input of DMU_j for $i = 1,\ldots,m$), to produce s different outputs, denoted by y_{rj} (the rth output of DMU_j for $r = 1,\ldots,s$). The fractional formulation of the DEA model for a particular DMU is defined as follows:

$$
\begin{aligned}
&Max\ e_0 = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\
&s.t: \\
&\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \qquad \forall j = 1, \dots, n \\
&u_r, v_i > 0 \qquad \forall r = 1, \dots, s;\ i = 1, \dots, m
\end{aligned} \tag{1}
$$

where it should be noted that u_r and v_i are the weights for output r and input i, respectively. The optimal ES is denoted by e_0 (a possible range of $0 < e_0 \leq 1$). The complete ES for e_0 equals 1 and $0 < e_0 < 1$ indicates the presence of inefficiency (Yang et al., 2009). The virtual input-output should represent all DMUs in a bi-dimensional plot. In the standard CCR (or BCC) model, a constraint is added to the problem of virtual input equals one to linearize the formulation states. Consequently, all DMUs could be located on the same vertical straight line in a virtual input versus virtual output plot, which causes a meaningless graphical representation. A different constraint is necessary to avoid the multiple optimal solutions for Model (1). Therefore, another constraint states the total sum of the input weights equal to one is proposed instead of the usual normalization (e Costa et al., 2016). It turns the primal CCR model in Model (1) into the following model:

$$
\begin{aligned}
&Maxz = \sum_{r=1}^{s} u_r y_{ro} \\
&s.t: \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \qquad \forall j = 1, \dots, n \\
&\sum_{i=1}^{m} v_i x_{io} = 1 \\
&u_r, v_i > 0 \qquad \forall r = 1, \dots, s; i = 1, \dots, m
\end{aligned} \tag{2}
$$

The dual model of the above-mentioned model is as follows:

$$
\begin{aligned}
&\theta^* = Min\ \theta_o \\
&s.t: \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_o x_{io} \qquad i = 1, \dots, m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \qquad r = 1, \dots, s \\
&\lambda_j \geq 0 \qquad j = 1, \dots, n
\end{aligned} \tag{3}
$$

For each DMU_o in the CCR dual Model (3), an imaginary composite unit is constructed that outperforms DMU_o. λ_j indicates the proportion to which DMU_j is utilized to construct the composite unit for DMU_o ($j = 1, \dots, n$), and it will be efficient if the $\theta^* = 1$ (Yang et al., 2009).

Benchmarking is originally defined and designed as a comprehensive quality tool to improve business operations and organizational performance. Benchmarking is an effort to attain superior performance by searching for the best practices of others and endeavoring to adopt these practices to suit the conditions of the organization (Rostamzadeh et al., 2021). The DEA has proven to be a strong tool for performance evaluation and benchmarking to improve organizations or companies' operations (Lai et al., 2011). The models that set the closest targets have significantly contributed to DEA as a benchmarking tool. Methodologically, the DEA models that minimize the distance to EF are extended to incorporate information on goals. The closest targets attempt to minimize the gap between actual performances and best practices. Therefore, they can show a way for DMUs to improve their performance with as little effort as possible (Ruiz and Sirvent, 2019). For this purpose, it is possible to define the common BPF as the facet of the DEA-EF spanned by a set of technically efficient DMUs, which can be seen as a common reference group. Targets will be chosen among projections of the DMUs on to this common BPF, which identifies the closest targets—minimizing the gap between actual inputs-outputs and targets guarantees identifying the best peers that are the most similar in global to the real performances of the DMUs. Therefore, they may be a benchmark for DMUs to find the easiest way to improve (Ruiz and Sirvent, 2016).

3.2. LIME

The rules guiding a classifier's output can be difficult to interpret in terms of content, even for transparent classifiers. The classifier's interpretability can become increasingly arduous when the model becomes complex or feature extraction is replaced by feature learning (Mishra et al., 2017). Hence, Ribeiro et al. (2016) attempted to answer the question, "Why should DMs trust the results of a classifier?" They stressed that trust is vital for effective human interaction with ML systems and that explaining individual predictions is indispensable in assessing trust. Then, they proposed an extensible instance-based model-agnostic algorithm named LIME to present a locally faithful explanation of the predictions of any classifier in an interpretable manner (Mishra et al., 2017).

The mechanism of LIME is based on illuminating reasons for a given prediction. For example, LIME for a system S applies label j to the instance x_i with probability y_{ij}. Then, to explain the prediction lists three reasons for the input x_i: R_1, R_2, and R_3. R_1 and R_2 are correlated positively with the decision, and R_3 is correlated negatively (Mishra et al., 2017). Specifically, suppose that an explanation is defined as a model $g \in G$, where G is a class of potentially interpretable models such as SVM, Decision Tree (DT), or artificial neural network. The domain of g is $\{0,1\}^d$ might then be a vector of binary values representing the absence or presence of the interpretable components. Then, let $\Omega(g)$ be a measure of complexity or complexity penalty of $g \in G$ which is opposed to interpretability. Let denote the explained model by $f: R^d \to R$. In classification, $f(x)$ is the probability (or a binary indicator) that indicates x belongs to a certain class. $\pi_x(z)$ is further utilized as a proximity measure between an instance z to x define locality around x. Finally, let's consider $\rho(f, g, \pi_x)$ as a measure of how unfaithful g is in approximating f in the locality defined by π_x. $\rho(f, g, \pi_x)$ must be minimized to ensure both interpretability and local fidelity while $\Omega(g)$ being low enough to be interpretable by humans. LIME learns a model g over the interpretable space by the minimization (Ribeiro et al., 2016):

$$\xi(x) = \underset{g \in G}{arg\ min} \rho(f, g, \pi_x) + \Omega(g) \tag{4}$$

In practice, G is considered as the set of linear regression models, with Ω restricting that only some explanatory features can have non-zero regression weights (although other types of explanation models could be implemented). The loss function is calculated through the weighted Euclidean distance (Peltola, 2018):

$$\rho(f, g, \pi_x) = \sum_i \pi_x(z_i) \left(f(z_i) - g\left(z_i'\right)\right)^2 \tag{5}$$

Where z_i is a perturbed data point in the original data space, z_i' is the corresponding interpretable representation, and the sum goes over a set of sampled perturbed points around x, $\{(z_i, z_i'), i = 1, \dots, m\}$. $\pi_x(z_i)$ weighs the samples according to their similarity to x, the point where the classification result is being explained.

3.3. MOCE

Counterfactual Explanation (CE) is a post-hoc method that has attracted much attention recently. Most existing CE methods are gradient-based or heuristic searches (Kanamori et al., 2020). By definition, when the desired prediction has not been obtained for supervised ML setups, CEs are applicable (Verma et al., 2020). More precisely, CEs provide information to users on what they might do to change the outcome of an automated decision (Keane et al., 2021). To re-run the classic example, suppose a customer seeks a home mortgage loan in a bank. The decision is largely impacted by an ML classifier that considers the customer's feature vector of {Income, Credit, Sex, Age, Marital, Education}. Unfortunately, the customer is denied the loan he/she seeks, and the following questions arise: (i) why was the loan denied? and (ii) what actions he/she can take differently in the future to approve the loan? The first question might be answered with explanations like: "Income was not satisfying." The latter question forms the basis of a CE: what small changes could be feasible for the customer to acquire validation to obtain the loan? For example, he/she can increase his/her credit (Verma et al., 2020).

CE was introduced by Wachter et al. (2017) as an optimization problem for the first time. The formalized objective is optimized by minimizing the distance between the counterfactual (x') and the original data point (x), subject to the constraint that the output of the classifier on the counterfactual is the desired label ($y' \in \mathrm{Y}$) (see Eq. (6)).

$$\begin{aligned} &\arg\min_{x'} d(x, x') \\ &\quad s.t: \\ &f(x') = y' \end{aligned} \tag{6}$$

It is possible that converting the objective into a differentiable, unconstrained form yields two terms. The first term inclines the classifier's output on the counterfactual to the desired class. On the other hand, the second term forces the counterfactual to be close to the original data point (see Eq. (7)).

$$\arg\min_{x'}\max_{\lambda} \lambda (f(x') - y')^2 + d(x, x') \tag{7}$$

A metric d measures the distance between two data points $x, x' \in \mathrm{X}$, which can be the L1/L2 distance, quadratic distance, or distance functions that take the cumulative distribution functions of the features (Mishra et al., 2017).

Major contributions of research in CE have sought to incorporate increasingly complex constraints on counterfactuals to ensure a truly actionable and useful resulting explanation. There are perhaps four big ideas that attempt to generate CEs with the following characteristics: (i) guided by proximity, (ii) feature-focused, (iii) distributionally faithful, and possibly, (iv) instance-based (Keane et al., 2021). Nevertheless, a single counterfactual might propose an interpretable but not actionable or counterproductive strategy in more general contexts (Dandl et al., 2020). Based on the best of our knowledge, Dandl et al. (2020) proposed the concept of MOCE to formalize the counterfactual search as a multi-objective optimization problem. MOCE returns a Pareto set of counterfactuals representing different trade-offs between proposed objectives, which are diverse in feature space. Changing to different features can lead to a desired counterfactual prediction, which seems preferable, and it is more likely that some counterfactuals meet a user's (hidden) preferences. Moreover, if multiple otherwise quite different counterfactuals propose changes to the same feature, the user rests surely that the feature is a significant lever to attain the desired outcome (Dandl et al., 2020).

4. Proposed approach

The proposed framework can be implemented in six phases, which are elaborated in this section.

Step 1: Obtaining multiple DEA-EFs

For providing an actionable and realistic target setting for DMUs, the concept of multiple EF proposed by Seiford and Zhu (2003) has been utilized in this study. The terms actionable and feasible refer to the fact that each unit should have achievable targets based on its policies, potentials, facilities, limitations, technologies, etc.; otherwise, targets would be invalid, and the peer DMU would be hypothetical. This approach proposes reaching the performance of the peer DMU in the reference set and is considered a benchmark. For example, consider a domestic online store that provides a specific product in a small country region. It is unachievable and unrealistic to set a target to have the same performance and efficiency as Amazon, which supplies various products worldwide. However, it is realistic to compare it with some of the equivalent online stores in the region, then, after growing, compete in the continent. Finally, if possible and competitive, in the worldwide market with Amazon. To do so, first, the primal CCR model is applied to the data set. After determining the efficient DMUs, they are considered the first group of efficient DMUs labeled EF_0 and then removed from the original dataset. In fact, EF_0 is located at the superior EF, emphasizing that collected DMUs are globally efficient. Then, another CCR model is applied to the rest of the data, and again the efficient DMUs are determined and removed from the data set, and the remains are moved for the next step of determining the efficient DMUs. Needless to say, these DMUs are not globally efficient since they are considered efficient in the absence of globally efficient DMUs (EF_0) and are located at the lower priority of efficiency. Finally, based on the experts' opinion, this process is terminated, and every group of DMUs is labeled according to their corresponding reference set. The last group of DMUs is considered inefficient DMUs since they could not be efficient even after removing various DMUs from the dataset. This technique causes classifying DMUs in their corresponding homogeneous common practice frontier that DMUs are counterparts based on their potential and performance. Fig. (1) illustrates the concept of multiple EFs, which consists of two EFs with nine DMUs. In this figure, EF_0 is globally efficient and DMU_F on EF_1 will try to get closer to DMU_D on EF_0 to turn into an efficient DMU. Among three efficient DMUs on EF_0, DMU_D has the least distance from it and has been considered as DMU_p accordingly.

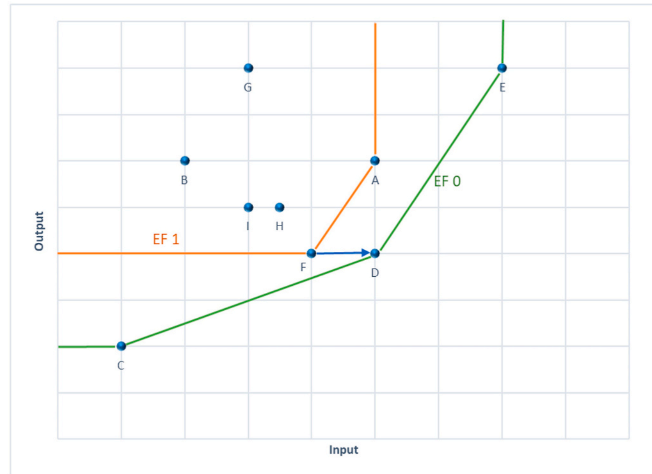

Fig. 1. DEA-EFs in two levels.

Step 2: Classifying DMUs

This approach implements the supervised classification model to classify DMUs in their corresponding reference sets. The main logic behind this idea is extracting vital features that positively contributed to the classification's outcome. In the real world, vital inputs-outputs for each DMU might vary. So, it does not make sense to adjust a specific input-output globally for the target setting of all DMUs. Hence, it is attempted to find vital inputs-outputs for each DMU specifically. To do so, developing a classification model with high predictive performance among different existing methods can be a way forward. From the previous step, we have labels of DMUs, which are their corresponding EF showing their efficiency level. Accordingly, after exploiting them, the best model with the highest predictive performance is considered the main classification model.

Step 3: Distance measure to find peer DMU

The current target setting framework is a distance-based approach; hence, Euclidean distance is used to find the distance between efficient and inefficient DMUs. First, among all non-global efficient DMUs, the one that we are interested in studying is selected as DMU_o. According to the provided information, a DMU's strategy should be to improve its performance to reach the reference EF, and it is supposed to set targets according to the peer DMU. For example, suppose a DMU is selected from non-global efficient DMUs and considered the DMU_o. If it locates in EF_3, it could improve its performance to reach the EF_2 so it should set a target based on its equivalent peer in EF_2. Needless to say, the strategy of reaching the EF_0 at the beginning could not be actionable and feasible. The Euclidean distance is calculated between DMU_o and all DMUs in the reference EF and DMU with the least distance are considered DMU_p. For calculating the Euclidean distance, DMUs' features (inputs-outputs) are considered the elements of the Euclidean distance equation.

Step 4: Feature importance based on post-hoc LIME method

LIME explains the classifier's predictions, which can be understandable for DMs. LIME determines the contributions of each feature to the classification's outcome. Hence, it is a practical tool to determine vital features that caused DMUs to be classified in their equivalent BPF. By analyzing the vital features, it is possible to set targets for DMUs to improve their performance to reach the reference EF. To do so, the dominant features that positively contribute to the prediction's output of both DMU_o and DMU_p are investigated. Then, the mutual and actionable features with positive correlations are selected for the target setting. The main motivation for this issue is selecting features that can positively and directly contribute to the target setting. Features with negative correlation represent the classification of the DMU in another class, and adjusting them cannot be effective for target setting.

Step 5: Developing MOCE-DEA mathematical programming model

After extracting the dominant and mutual features of DMU_o and DMU_p, it is possible to develop the MOCE-DEA for target setting. It is worth recalling that based on the CE logic, it is not possible to adjust every feature because, in some cases, it is impossible to adjust it, or it does not make sense to do so. Hence, among the mutual features between DMU_o and DMU_p those that can be actionable according to DEA logic are collected, and the remaining are ignored. Here, the idea for developing MOCE-DEA is to consider DMU_o as a new member of the reference EF. Given this hypothesis, it is possible to obtain DMU_o efficiency by solving the modified LP model of Model (3) by adding it to the reference EF. For solving the MOCE-DEA, there are two objectives: (i) the distance between DMU_o and DMU_p should be minimized by adjusting the actionable and feasible features, (ii) this minimization must maximize ES of DMU_o. Based on the provided information, it is possible to take advantage of CF by adjusting the value of actionable and feasible features to achieve the highest ES. Although this is a non-linear and NP-hard problem, it considers multiple features simultaneously, making it generalizable for different problems. In what follows, the CF model is developed as a multi-objective optimization problem according to Model (3):

$$
\begin{aligned}
&\text{Min}\left(\|\theta_o-1\|,\left\|x_{io}-\Delta x_{io}-x^*_{ip}\right\|,\left\|y_{ro}+\Delta y_{ro}-y^*_{rp}\right\|\right) \quad \forall i\in CI,\forall r\in CO\\
&s.t:\\
&\sum_{j=1}^{n}\lambda_j x_{ij}\le\theta_o x_{io} \quad \forall i\in NCI\\
&\sum_{\substack{j=1\\ j\neq o}}^{n}\lambda_j x_{ij}+\lambda_o(x_{io}-\Delta x_{io})\le\theta_o(x_{io}-\Delta x_{io}) \quad \forall i\in CI\\
&\sum_{j=1}^{n}\lambda_j y_{rj}\ge y_{ro} \quad \forall r\in NCO\\
&\sum_{\substack{j=1\\ j\neq o}}^{n}\lambda_j y_{rj}+\lambda_o(y_{ro}+\Delta y_{ro})\ge y_{ro}+\Delta y_{ro} \quad \forall r\in CO\\
&x_{io}-\Delta x_{io}\ge x^*_{ip} \quad \forall i\in CI\\
&y_{ro}+\Delta y_{ro}\le y^*_{rp} \quad \forall r\in CO\\
&\lambda_j\ge 0,\quad j=1,\dots,n\\
&\Delta x_{io},\Delta y_{ro}\ge 0,\ \theta\ \textit{is free}
\end{aligned}
\tag{8}
$$

where CI and CO are sets of inputs and outputs that contribute to target setting. In the same way, NCI and the NCO are sets of inputs and outputs that do not contribute to the target setting. In this model, objective functions are modeled as follows.

1. Minimizing the distance between θ_o and 1, which is the ES. The higher value θ_o inclines it to the reference EF.
2. Minimizing the distance of actionable features of DMU_o and DMU_p. x_{io} and y_{ro} represent the value of actionable features of DMU_o referring to input and output, respectively. Also, x^*_{ip} and y^*_{rp} have the same role for DMU_p. Now, it is clear that for minimizing the actionable features' distance of DMU_o and DMU_p, the values of Δx_{io} and Δy_{ro} should be maximized, which is the subtraction of corresponding features. Needless to say, there can be several objectives based on the number of actionable features extracted from LIME, which is the main advantage of this model.

In the constraints, it is critical to consider DMU_o as a new member of the reference set. The main logic behind this is determining the amount of adjustment that can lead DMU_o to the reference EF. The concept of x_{ij}, y_{rj}, x_{io}, y_{ro} have the same definition as the primal CCR model. However, the values of λ_j should be calculated by a multi-objective optimization model. Note that because of the nature of the objective function, the fifth and sixth constraints are redundant and can be discarded from the model.

It should be taken into consideration that one of the objectives tries to minimize the distance of the DMUs, and the other maximizes its θ_o. On the other side, according to Model (8), the value of λ_j has an important role in finding the optimal values of objective functions because a metaheuristic algorithm generates their values, affecting θ_o. Therefore, the value θ_o cannot properly reflect the robust θ_o in practice. For obtaining a vivid insight into the counterfactual change of the DMU_o, it is verified by the primal CCR model by calculating the robust PS.

Step 6: Algorithm selection and solving the mathematical program

To solve MOCE-DEA, Multi-Objective Particle Swarm Optimization (MOPSO) introduced by Coello et al. (2004) is implemented. MOPSO is the extension of the PSO algorithm, a population-based metaheuristic algorithm introduced by Kennedy and Eberhart (1995). The experimental results show that the PSO can converge fast because it does not involve selection operation or mutation calculation, so the search can be performed by repeatedly varying the particle's speed. Also, the performance of PSO is not susceptible to population size, and PSO scales well (Abbaspour Onari et al., 2021). PSO generates new solutions based on

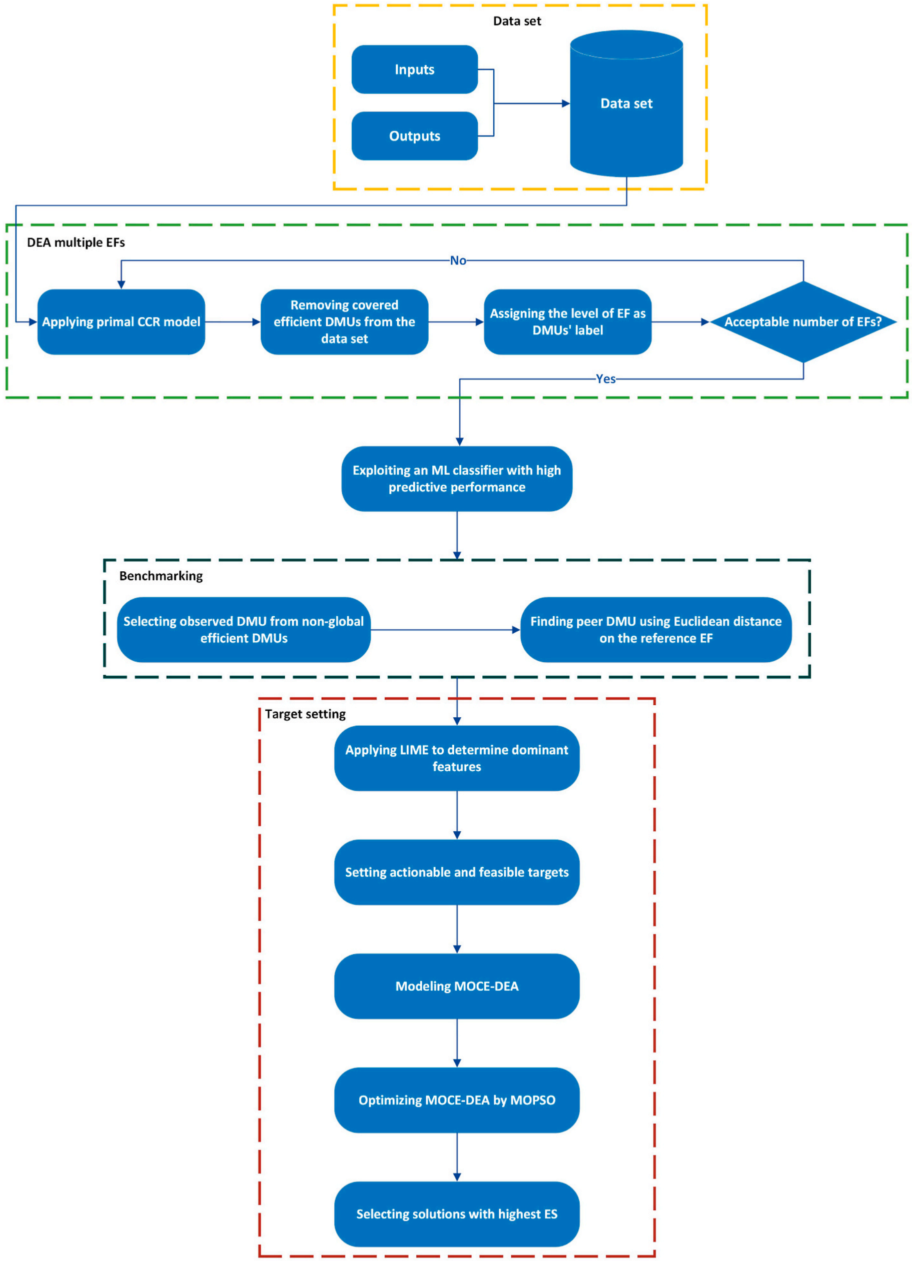

Fig. 2. Diagram of the proposed framework.

two equations:

$$v_i(t+1) = w * v_i(t) + c_1 * rnd() * (pbest_i(t) - x_i(t)) + c_2 * rnd() * (gbest(t) - x_i(t)) \tag{9}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{10}$$

where c_1 and c_2 denote the acceleration constant for weighting the stochastic acceleration terms that pull a single particle toward personal best (*pbest*) and global best (*gbest*) positions. $rnd()$ indicates a random variable that is generated by uniform distribution between [0, 1]. w, x, v refer to inertia weight, the position vector, and velocity vector, respectively (Abbaspour Onari and Jahangoshai Rezaee, 2022, 2023).

The multi-objective procedures should supply two main properties. First, generating high-quality non-dominated solutions on the Pareto frontier of the Multi-Objective Decision-Making (MODM) problem (Khalili-Damghani et al., 2013). To do so, the best non-dominated positions (leaders) are implemented to guide particles. An additional set called an external archive is used for storing the non-dominated solutions for Pareto optimal solutions. Besides, the external archive is limited; accordingly, it is necessary to use methods to replace existing solutions with new ones (Meza et al., 2017). Second, concerning a proper diversity for the generated solutions on the Pareto frontier of MODM problem (Khalili-Damghani et al., 2013). For selecting the individual best of each particle, a single position *pbest* is maintained and is only replaced if x_i is better than p_i. In the meantime, selecting the best position group can be performed randomly (Khalili-Damghani et al., 2013). The MOPSO algorithm has been selected for this study because the proposed model is non-linear, and also multiple solutions are required. Besides, it is a popular algorithm for non-linear programming problems (Jabbari et al., 2022). Finally, among multiple Pareto optimal solutions, only one solution with the highest θ_o is selected as the best solution. The infographic of the proposed framework has been provided in Fig. (2).

Table 1
Hospitals' statistical information.

Type	Variable	Mean	Standard Deviation	Min	Max	Description
Inputs	Labor	468.78	353.853	21	2792	Number of personnel, permanent personnel, contract workers, and other personnel.
	Facility	325.5	330.99	10	2676	Reliable and favorable performance, unreliable and unfavorable performance, salvage equipment, and useless equipment.
	Active-bed	167.402	130.65	14	806	Number of available and active cots.
Outputs	Inpatients	107,596.05	118,490.18	365	818,257	Patients that are referred directly to the hospital and patients that transferred from other hospitals.
	Outpatients	187,805	204,731.95	48	1,914,155	Patients who are not hospitalized for 24 h.
	Starred	37,908	49,240.72	6	489,031	Having special diseases and needing special care.
	Bed-day	44,824.8	40,467.93	189	230,058	Number of operational beds in a month.
	Occupancy-rate	67.79	16.73	1.95	116.41	The number of occupancy bed-day divided by total bed-days × 100. Occupancy bed-day is equal to the operational beds × number of occupancy days.

Table 2
DMUs and their corresponding EF level with PS before reaching the corresponding EF.

DMU	EF	PS	DMU	EF	PS	DMU	EF	PS	DMU	EF	PS	DMU	EF	PS	DMU	EF	PS
0	1	0.8418	48	2	0.7534	96	1	0.8877	144	5	0.9978	192	5	0.9043	240	2	0.9222
1	2	0.9236	49	5	0.9980	97	2	0.9759	145	4	0.9662	193	3	0.9713	241	1	0.5502
2	1	0.7980	50	3	0.9258	98	2	0.8682	146	2	0.6060	194	2	0.8337	242	1	0.8583
3	0	–	51	3	0.9316	99	1	0.8196	147	2	0.8849	195	2	0.9149	243	1	0.8591
4	0	–	52	2	0.9979	100	2	0.9599	148	1	0.9775	196	5	0.9178	244	3	0.9911
5	0	–	53	0	–	101	3	0.9348	149	3	0.8825	197	4	0.9415	245	6	0.8958
6	5	0.9098	54	2	0.9483	102	3	0.9851	150	5	0.8457	198	3	0.9940	246	5	0.7712
7	1	0.8755	55	0	–	103	4	0.9257	151	4	0.8972	199	4	0.9963	247	3	0.9303
8	2	0.7757	56	0	–	104	6	0.9542	152	0	–	200	4	0.8567	248	1	0.9430
9	4	0.9762	57	2	0.9149	105	6	0.7731	153	6	0.9765	201	0	–	249	4	0.9977
10	4	0.9413	58	3	0.9481	106	3	0.9524	154	6	0.9762	202	3	0.9292	250	0	–
11	4	0.8536	59	3	0.9214	107	1	0.6521	155	2	0.9035	203	1	0.7874	251	1	0.9079
12	2	0.9009	60	3	0.7865	108	1	0.7610	156	1	0.9401	204	6	0.8109	252	5	0.8796
13	0	–	61	3	0.8775	109	2	0.8993	157	2	0.9168	205	2	0.7356	253	2	0.9131
14	2	0.9550	62	6	0.7351	110	0	–	158	6	0.8393	206	3	0.8834	254	3	0.893
15	2	0.8979	63	6	0.7922	111	3	0.8587	159	4	0.9395	207	4	0.8515	255	2	0.9272
16	5	0.9081	64	4	0.9350	112	0	–	160	3	0.9319	208	6	0.8130	256	3	0.9157
17	4	0.9278	65	4	0.9301	113	3	0.8845	161	3	0.9727	209	5	0.9990	257	2	0.8864
18	5	0.8333	66	1	0.5933	114	1	0.9118	162	2	0.9603	210	6	0.7653	258	2	0.7722
19	5	0.9887	67	5	0.9152	115	1	0.9732	163	1	0.8897	211	3	0.8762	259	4	0.8554
20	2	0.6910	68	5	0.8150	116	1	0.8138	164	3	0.9961	212	1	0.9036	260	3	0.8763
21	4	0.9306	69	3	0.9411	117	0	–	165	3	0.8855	213	5	0.9225	261	3	0.9289
22	2	0.9874	70	4	0.8283	118	0	–	166	1	0.8555	214	1	0.8408	262	4	0.8412
23	3	0.9129	71	0	–	119	6	0.8538	167	4	0.7472	215	3	0.9786	263	3	0.9829
24	1	0.8866	72	3	0.9205	120	3	0.8845	168	5	0.9540	216	4	0.9166	264	4	0.8921
25	2	0.9917	73	3	0.9393	121	1	0.6846	169	2	0.8432	217	0	–	265	3	0.8984
26	5	0.9307	74	2	0.7878	122	5	0.9711	170	1	0.8552	218	3	0.8822	266	4	0.9109
27	4	0.9555	75	6	0.9393	123	3	0.9180	171	4	0.9569	219	3	0.9238	267	0	–
28	3	0.9087	76	2	0.7098	124	4	0.8187	172	1	0.9148	220	2	0.992	268	2	0.8800
29	1	0.8333	77	6	0.8746	125	0	–	173	4	0.9755	221	4	0.9487	269	5	0.8580
30	6	0.7848	78	3	0.9173	126	1	0.9190	174	3	0.9676	222	5	0.8745	270	5	0.9216
31	6	0.6243	79	2	0.8584	127	0	–	175	1	0.7486	223	1	0.8676	271	2	0.9293
32	2	0.8289	80	0	–	128	2	0.9178	176	4	0.9909	224	2	0.8653	272	4	0.9764
33	2	0.8706	81	1	0.9123	129	5	0.9388	177	1	0.9749	225	4	0.8382	273	1	0.9137
34	3	0.8524	82	5	0.9975	130	6	0.9909	178	1	0.7887	226	2	0.8238	274	3	0.7622
35	3	0.8909	83	1	0.6582	131	3	0.8287	179	1	0.9386	227	4	0.9649	275	1	0.6941
36	1	0.7092	84	1	0.6714	132	3	0.9585	180	3	0.9881	228	5	0.9348	276	2	0.9523
37	2	0.7747	85	2	0.7489	133	4	0.8852	181	4	0.9588	229	3	0.8356	277	3	0.9489
38	2	0.9531	86	6	0.8258	134	5	0.9897	182	3	0.7700	230	4	0.9115	278	3	0.9032
39	6	0.7947	87	4	0.7951	135	2	0.8905	183	5	0.9968	231	3	0.97	279	6	0.5982
40	6	0.9481	88	6	0.9806	136	2	0.9300	184	2	0.9014	232	3	0.9754	280	1	0.7722
41	6	0.9552	89	1	0.8733	137	5	0.9280	185	4	0.9658	233	6	0.9803	281	4	0.8772
42	2	0.9795	90	4	0.9863	138	3	0.7662	186	4	0.9840	234	2	0.7535	282	4	0.8295
43	6	0.9251	91	3	0.8967	139	2	0.7379	187	5	0.8167	235	1	0.6547	283	5	0.8301
44	2	0.8265	92	1	0.7892	140	0	–	188	1	0.5387	236	2	0.9084	284	4	0.9481
45	4	0.9633	93	5	0.9759	141	2	0.7773	189	0	–	237	1	0.844	285	0	–
46	2	0.6158	94	3	0.9716	142	2	0.9299	190	0	–	238	3	0.6435	286	0	–
47	6	0.9050	95	5	0.9294	143	1	0.7989	191	3	0.9563	239	0	–	287	0	–

5. Case study and analysis of the results

This section implements the proposed framework for benchmarking a group of hospitals' data set. In Subsection 1, the data preprocessing is presented, and the DEA multiple EFs are determined. In Subsection 2, the most accurate classification model is selected among some classification models. Eventually, in Subsection 3, the target setting process, which applies LIME and MOCE-DEA, is covered.

Table 3
Predictive performance of different classifiers.

Classification method	Accuracy	Precision	Recall	F1 score
LR	0.3611	0.3999	0.3611	0.3366
SVM	0.6007	06777	0.6007	0.6151
DT	0.7083	0.7204	0.7083	0.7083
MLP	0.8403	0.8409	0.8403	0.8402
RF	**0.9722**	**0.9736**	**0.9722**	**0.9725**

5.1. Data preprocessing and DEA multiple EFs

For evaluating the proposed approach, a group of hospitals has been selected to evaluate the efficient DMUs and set targets to improve the performance of the inefficient DMUs or non-global efficient DMUs. This data set consists of 288 hospitals from all 31 provinces in Iran located in the capitals of the provinces. Like all DEA models, each DMU has some inputs-outputs, which are considered the features of the classification model. In this case, DMUs have three inputs: the number of hospital personnel, medical equipment in the hospital, and available hospital cots. On the one side, there are five outputs: the number of inpatients, outpatients, special patients, operational beds, and bed occupancy rate. The hospitals' information has been shown in Table 1 (Rezaee and Karimdadi, 2015).

The first primal CCR model is applied according to the main approach, and the first EF is determined. The efficient DMUs are labeled according to their corresponding EF (EF_0). Then, they are dropped from the data set, and the primal CCR model is applied again, and the efficient DMUs are obtained and labeled according to their corresponding frontier (EF_1). This process is continued until the acceptable number of EF makes sense according to the problem and experts' opinions. Table 2 demonstrates the results of the primal CCR model and corresponding labels with DMUs' PS before reaching the corresponding EF for ease of comparison.

5.2. Applying different classification models

Applying an accurate classification model for the labeled dataset is important in this step to obtain an appropriate explanation from the LIME. Among various classification methods, five classifiers are collected: Logistic Regression (LR), DT, SVM, Multi-Layer Perceptron (MLP), and RF. Since the dataset does not have any outliers or missing values, only data normalization is applied for LR, SVM, and MLP. For DT and RF, data normalization is not mandatory. In order to avoid overfitting, 5-fold cross-validation is used to validate models. Also, parameter optimization is applied to train classifiers. Finally, the classifiers' prediction results are represented in Table 3.

RF has obtained the best predictive performance among other classifiers so far. This performance is considered outstanding because the data is small, and RF with such a small amount of data can obtain this performance. Other classification methods have a poor performance, which cannot be reliable for the current study.

5.3. MOCE-DEA simulation and results analysis

To clarify the proposed approach, two cases are studied in this subsection. First, a random DMU_o is selected among non-global efficient DMUs for target setting. Then, the DMU_p is found based on the Euclidean distance among the reference EF. Subsequently, the LIME method is applied to DMU_o and DMU_p to acquire the contribution of features on the obtained classification outcome. In this study, the algorithm developed by Ribeiro et al. (2016) has been implemented for applying LIME. Then, according to the MOPSO, MOCE-DEA is simulated, and the optimal solutions are obtained. In this study, the maximum iteration is 500, and the population and repository size is 100. Only one solution is collected with the highest θ_o among Pareto solutions for each simulation batch.

Case 1. $DMU_o = 99$ (Underevaluation), $DMU_p = 3$ (Identified by Euclidean distance)

The $DMU_o = 99$ locates in EF_1 and substantially is an inefficient DMU. For setting an actionable and feasible target for it, the dominant features should be extracted by LIME. The priority of the contributions of each feature for the classification's outcome has been demonstrated in Fig. (3).

According to the results of LIME in Fig. (3), Facility, Labor, Outpatient, Active-bed, and Bed-day are dominant features in classifying $DMU_o = 99$ in EF_1, respectively. In the same way, the results in Fig. (4) reflect the contribution of features for $DMU_p = 3$, which locates in EF_0 as a global efficiency.

For $DMU_p = 3$, Facility, Labor, Active-bed, and Outpatient are dominant features to classify it in EF_0, respectively. Among all dominant features for DMUs, Facility, Labor, Outpatient, and Active-bed are considered mutual features because they all positively contribute to both DMUs' classification outcomes. However, it is critical to confirm them logically for target setting. Facility, Labor, and Active-bed are inputs of the model, and their value should be reduced based on DEA logic. LIME for all of them suggests the same strategy. However, the correlation of Active-bed is zero according to LIME, meaning it is not a vital feature (see Fig. (3)). Hence, it is ignored from target setting. On the other side, Outpatient is the model's outcome, and its value should be increased. However, LIME suggests an opposite strategy which is against DEA logic. Consequently, it is being ignored for target setting. Finally, only Facility and Labor are actionable and feasible features for target setting. Table 4 represents the target setting strategy of $DMU_o = 99$.

MOCE-DEA has three objective functions: two functions decrease the distance between DMU_o and DMU_p, and one strives to maximize θ_o. It is worth recalling that it is not expected that DMU_o turning an efficient DMU because it is not a realistic target based on the various factors that a DMU should be acquired. It is important to improve its performance according to its equivalent DMU_p. After 30 independent simulations of MOCE-DEA, solutions with the highest θ_o in the repository of Pareto solutions are collected in Table 5. The highest θ_o value has been obtained

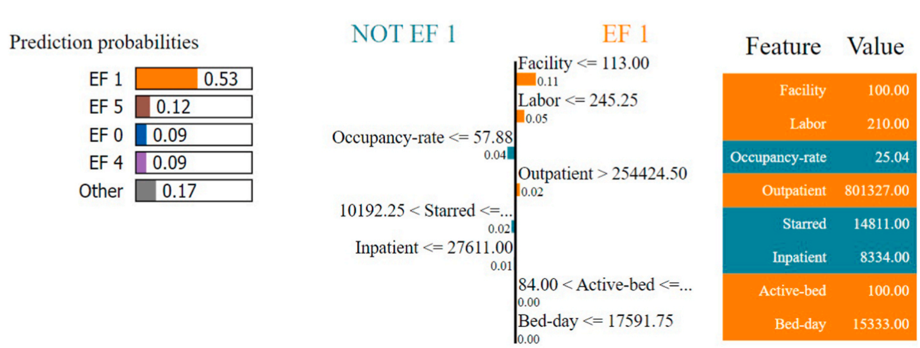

Fig. 3. Contribution of each feature to the prediction outcome of $DMU_o = 99$.

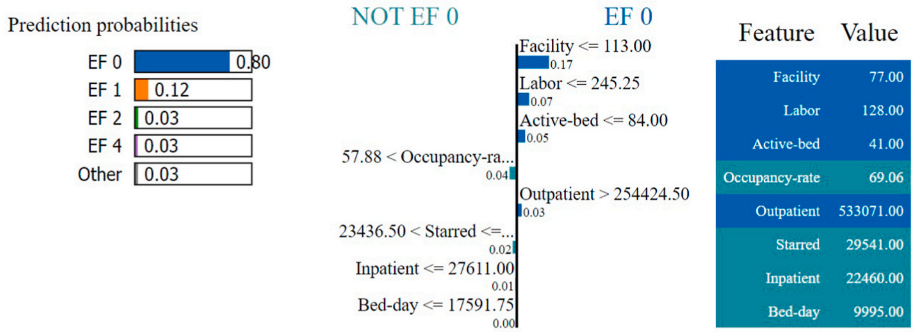

Fig. 4. Contribution of each feature on the prediction outcome of $DMU_p = 3$.

Table 4
Target setting for $DMU_o = 99$ according to actionable features.

Current Value	Target Value
Facility = 100	Facility = 77
Labor = 210	Labor = 128

in Run 23. The Pareto frontier for this simulation batch has been demonstrated in Fig. (5). In this figure, red stars demonstrate the optimal solutions in the Pareto frontier. According to the best solution, it can improve the value of Facility by 81.3% with the optimized value of 81.3025 for it. The improved value for Labor is 61.1%, with the optimized value of 128.3043. These adjustments lead DMU_o to a very close distance from DMU_o with $\theta_o = 0.9920$, which is very close to one.

In order to validate the robustness of the generated solution by MOCE-DEA, the optimized values for actionable and feasible features are replaced with their original value for $DMU_o = 99$. Then, the conventional CCR model based on Model (3) is rebuilt. Table 6 represents the robust improvement of PS. $DMU_o = 99$ turns into an efficient DMU on EF_0 representing adjusted targets can lead it to be an efficient global DMU. These adjustments have improved the PS of $DMU_o = 99$ by 18.04% is sufficient to make it a globally efficient DMU.

Case 2. $DMU_o = 209$ (Underevaluation), $DMU_p = 87$ (Identified by Euclidean distance)

In this case, $DMU_o = 209$ locates in EF_5 and by target setting, it will be endeavored to reach EF_4. For $DMU_o = 209$, Occupancy-rate, Facility, Starred, and Outpatient are dominant features according to LIME's results in Fig. (6).

$DMU_p = 87$ locates in EF_4, and except Starred and Bed-day, other features are dominant according to LIME's result (see Fig. (7)). Among dominant features for DMUs, Occupancy-rate, Facility, and Outpatient are considered mutual features. According to DEA logic, the value of Occupancy-rate and Outpatient should be increased because they are outputs that coincide with LIME's suggestion. Also, the value of Facility should be reduced because it is input for the model. However, according to LIME's result for both DMUs, Outpatient has not been considered a vital feature due to zero correlation value, so it is ignored for target setting. Hence, only Occupancy-rate and Facility are actionable and

Table 5
The best Pareto solutions with the highest ES for each simulation batch ($DMU_o = 99$).

	$\Delta y_{Facility}$	Optimized value	Improvement Percentage	Δy_{Labor}	Optimized value	Improvement Percentage	Cumulative Improvement	Rank	θ_o
Run01	1.8527	98.1473	8.1	79.4515	130.5485	62.2	70.2%	24	0.9298
Run02	0.1520	99.8480	0.7	67.5307	142.4693	67.8	68.5%	25	0.8591
Run03	3.1241	96.8759	13.6	78.1817	131.8183	62.8	76.4%	16	0.8252
Run04	5.9329	94.0671	25.8	79.5753	130.4247	62.1	87.9%	10	0.9408
Run05	5.6547	94.3453	24.6	80.0062	129.9938	61.9	86.5%	11	0.8189
Run06	17.3606	82.6394	75.5	78.0135	131.9865	62.9	138.3%	3	0.8863
Run07	10.9032	89.0968	47.4	81.8068	128.1932	61.0	108.4%	5	0.8470
Run08	2.9642	97.0358	12.9	81.3128	128.6872	61.3	74.2%	19	0.9365
Run09	3.4158	96.5842	14.9	81.3436	128.6564	61.3	76.1%	17	0.9660
Run10	0.2599	99.7401	1.1	81.0115	128.9885	61.4	62.6%	29	0.6010
Run11	17.7969	82.2031	77.4	77.1939	132.8061	63.2	140.6%	2	0.9313
Run12	5.3275	94.6725	23.2	73.4762	136.5238	65.0	88.2%	9	0.9300
Run13	0.4306	99.5694	1.9	15.3009	194.6991	92.7	94.6%	6	0.9307
Run14	2.1987	97.8013	9.6	77.6202	132.3798	63.0	72.6%	22	0.7275
Run15	3.7643	96.2357	16.4	79.0778	130.9222	62.3	78.7%	15	0.9151
Run16	0.7718	99.2282	3.4	47.3806	162.6194	77.4	80.8%	13	0.7173
Run17	0.3591	99.6409	1.6	70.3498	139.6502	66.5	68.1%	26	0.9777
Run18	1.0124	98.9876	4.4	50.1745	159.8255	76.1	80.5%	14	0.9251
Run19	0.6351	99.3649	2.8	73.1642	136.8358	65.2	67.9%	27	0.9759
Run20	2.4779	97.5221	10.8	77.8504	132.1496	62.9	73.7%	20	0.8846
Run21	0.6435	99.3565	2.8	20.5355	189.4645	90.2	93.0%	7	0.8510
Run22	3.2283	96.7717	14.0	80.0628	129.9372	61.9	75.9%	18	0.7430
Run23	18.6975	81.3025	81.3	81.6957	128.3043	61.1	142.4%	1	0.9920
Run24	5.4834	94.5166	23.8	81.5574	128.4426	61.2	85.0%	12	0.8947
Run25	1.3829	98.6171	6.0	69.7702	140.2298	66.8	72.8%	21	0.9352
Run26	0.9015	99.0985	3.9	67.1075	142.8925	68.0	72.0%	23	0.6989
Run27	0.5908	99.4092	2.6	81.8519	128.1481	61.0	63.6%	28	0.6036
Run28	5.8401	94.1599	25.4	76.4843	133.5157	63.6	89.0%	8	0.6794
Run29	17.2078	82.7922	74.8	81.0495	128.9505	61.4	136.2%	4	0.8195
Run30	0.0187	99.9813	0.1	80.2312	129.7688	61.8	61.9%	30	0.9677

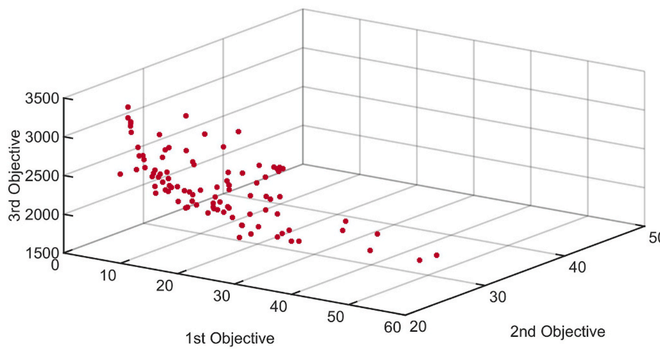

Fig. 5. Pareto frontier for efficient solution in Run 23.

Table 6
Robust improvement of $DMU_o = 99$ according to Model (3).

	Genuine PS	Improved PS	Improvement Percentage
EF 0	0.8196	1.000	18.04

feasible features for target setting. Table 7 represents the target setting strategy for $DMU_o = 209$ according to the actionable features.

MOCE-DEA for this case, has three objective functions to obtain the optimal solution as well. Accordingly, two functions decrease the distance between DMU_o and DMU_p, and one strives to maximize θ_o. The batch of solutions after 30 independent simulations of MOCE-DEA has been provided in Table 8 for the highest value of θ_o. The best θ_o belongs to Run 9. The Pareto frontier for this solution has been presented in Fig. (8).

The value of $\theta_o = 0.9996$ after applying the optimal solution indicates that $DMU_o = 209$ can reach the reference EF_4. The value of Occupancy-rate is improved by 19.9% with the optimized value of 50.2808. Also, the improved value for Facility is 99.1%, with the optimized value of 297.5774. This is an ideal outcome because the DMU, by adjusting the optimized solution of MOCE-DEA, can improve its performance and reach reference EF. Table 9 represents the robust improvement of DMU_o based on Model (3). The improved PS for $DMU_o = 209$ from EF_0 to EF_4 indicates clear improvement after adjustment. From EF_0 to EF_3, $DMU_o = 209$ can not reach corresponding EFs, which is expectable because the target has been set to improve its performance to the closest reference EF. Finally, in EF_4 according to the expectation, it can reach the reference EF.

6. Managerial insight and discussion

In the competitive market, setting goals and planning to achieve them is the main step for the growth of an organization. Furthermore, the performance of healthcare systems has a significant social impact. As important as setting goals is for an organization, achieving them is even more important. Therefore, determining achievable goals is one of the main tasks of managers. One of the challenges of managers in this field is determining the pattern, specifying the target value, or considering the combined goals to improve the organization. In previous studies in the DEA target setting field, it was impossible to determine an appropriate benchmark. Additionally, the importance of inputs-outputs could not be easily tracked, which XAI has made possible. Finally, in conventional methods, the target values were determined separately for input-outputs, while the MOCE model allows for combined settings for input-output values to be possible for the DMU to achieve maximum efficiency at its own level. By exploiting XAI methods, we attempted to address two research gaps in DEA literature. First, setting targets based on the realistic potential of DMUs. The solution consists of two steps: generating multiple EFs by primal CCR model, which determines the

Table 7
Target setting for $DMU_o = 209$ according to actionable features.

Genuine Value	Target Value
Occupancy-rate = 41.95 **Facility = 433**	Occupancy-rate = 50.36 Facility = 278

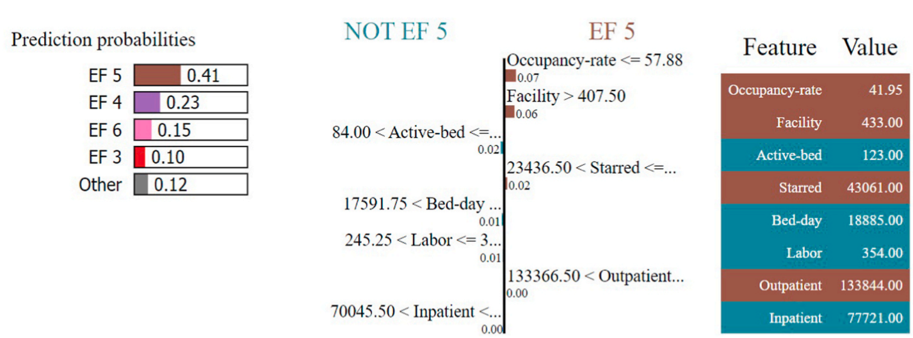

Fig. 6. Contribution of each feature on the prediction outcome of $DMU_o = 209$.

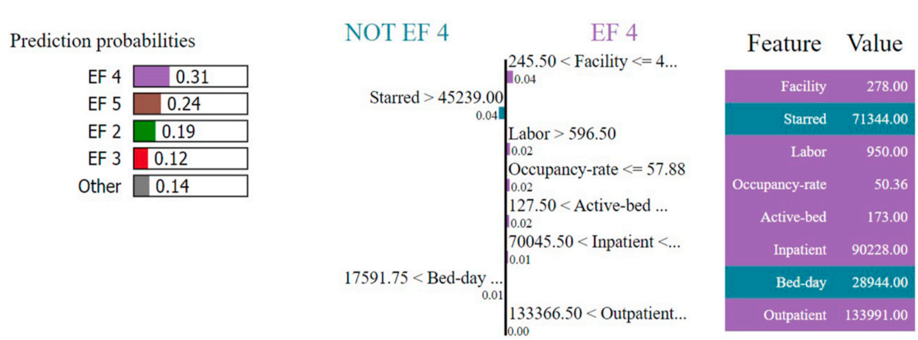

Fig. 7. Contribution of each feature on the prediction outcome of $DMU_p = 87$.

Table 8
The best Pareto solutions with the highest ES for each simulation batch ($DMU_o = 209$).

	$\Delta y_{Occupancy-rate}$	Optimized value	Improvement Percentage	$\Delta y_{Facility}$	Optimized value	Improvement Percentage	Cumulative Improvement	Rank	θ_o
Run01	7.7716	49.7216	18.5	27.3168	405.6832	92.4	110.0	10	0.7778
Run02	8.3684	50.3184	19.9	56.2161	376.7839	99.5	135.8	7	0.5255
Run03	8.3943	50.3443	20.0	5.8817	427.1183	99.8	103.6	14	0.9647
Run04	6.5665	48.5165	15.7	8.2835	424.7165	78.1	83.4	18	0.7794
Run05	8.1635	50.1135	19.5	9.0555	423.9445	97.1	102.9	13	0.8007
Run06	7.6301	49.5801	18.2	14.7339	418.2661	90.7	100.2	12	0.9503
Run07	7.9676	49.9176	19.0	41.1633	391.8367	94.7	121.3	8	0.8151
Run08	2.3355	44.2855	5.6	0.8270	432.1730	27.8	28.3	26	0.9951
Run09	8.3308	50.2808	19.9	135.4226	297.5774	99.1	186.4	1	0.9996
Run10	8.0936	50.0436	19.3	91.5815	341.4185	96.2	155.3	4	0.9607
Run11	4.6896	46.6396	11.2	2.1744	430.8256	55.8	57.2	22	0.8527
Run12	8.1524	50.1024	19.4	29.4316	403.5684	96.9	115.9	9	0.9745
Run13	8.2386	50.1886	19.6	97.3800	335.6200	98.0	160.8	3	0.7509
Run14	0.5230	42.4730	1.2	0.1438	432.8562	6.2	6.3	30	0.7819
Run15	4.8147	46.7647	11.5	1.5554	431.4446	57.2	58.3	23	0.7950
Run16	7.8576	49.8076	18.7	25.4838	407.5162	93.4	109.9	11	0.9794
Run17	8.2569	50.2069	19.7	2.3075	430.6925	98.2	99.7	17	0.4739
Run18	5.3864	47.3364	12.8	0.5695	432.4305	64.0	64.4	21	0.7831
Run19	8.3007	50.2507	19.8	122.7899	310.2101	98.7	177.9	2	0.8341
Run20	0.8077	42.7577	1.9	2.0561	430.9439	9.6	10.9	27	0.1810
Run21	8.3297	50.2797	19.9	68.9031	364.0969	99.0	143.5	6	0.9143
Run22	7.8230	49.7730	18.6	4.5297	428.4703	93.0	95.9	15	0.6060
Run23	8.3661	50.3161	19.9	79.7849	353.2151	99.5	151.0	5	0.9090
Run24	6.8153	48.7653	16.2	4.2789	428.7211	81.0	83.8	19	0.8392
Run25	6.2573	48.2073	14.9	0.3588	432.6412	74.4	74.6	20	0.6261
Run26	1.0318	42.9818	2.5	0.4784	432.5216	12.3	12.6	28	0.9396
Run27	2.3428	44.2928	5.6	1.3027	431.6973	27.9	28.7	25	0.5136
Run28	0.8195	42.7695	2.0	0.9778	432.0222	9.7	10.4	29	0.8524
Run29	1.9219	43.8719	4.6	7.3916	425.6084	22.9	27.6	24	0.8594
Run30	6.4619	48.4119	15.4	9.0315	423.9685	76.8	82.7	16	0.8743

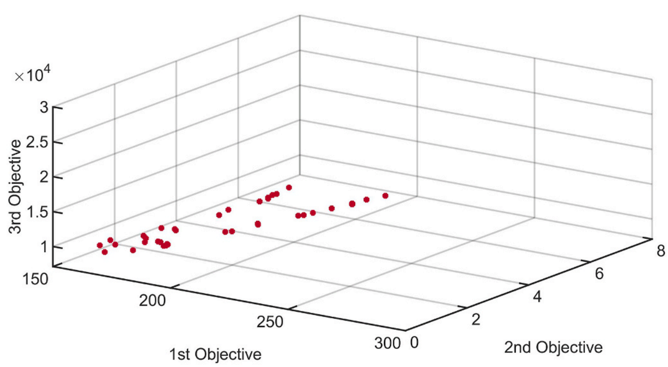

Fig. 8. Pareto frontier for efficient solution in Run 9.

Table 9
Robust improvement of $DMU_o = 209$ according to Model (3).

	Genuine PS	Improved PS	Improvement Percentage
EF 0	0.3819	0.4047	5.97
EF 1	0.4589	0.4758	3.68
EF 2	0.6081	0.7153	17.63
EF 3	0.7781	0.8813	13.27
EF 4	0.9990	1.0000	0.1

global efficient DMUs, then in each step, covered DMUs are removed from the original data set, and new EFs are generated. The main motivation to implement this method is following a stepwise gradual target setting strategy based on units' real potential. In our studied cases, in the first case, we analyzed $DMU_o = 99$, which locates in EF_1 and its target is reaching EF_0 to turn into a globally efficient DMU. In the second case, we studied $DMU_o = 209$, which locates in EF_5 and its target is reaching EF_4. In the next step, by implementing benchmarking, we determined DMU_p in the reference EF for each DMU_o to set realistic targets that need the least adjustment to reach the reference EF. In our studied cases, $DMU_p = 3$ and $DMU_p = 87$ are peer DMUs for $DMU_o = 99$ and $DMU_o = 209$ in the EF_0 and EF_4, respectively.

Second, identifying the vital input-outputs that can positively contribute to target setting implementing XAI methods. This approach can help us to adjust only inputs-outputs that are vital and ignore others that might not be feasible or important in practice. Besides, every DMU_o will have a unique strategy to adjust inputs-outputs based on their potential. After classifying DMUs in their corresponding EFs, LIME extracts local feature importance for DMUs. According to LIME explanations, for $DMU_o = 99$, facility and labor, which are inputs of the model, are dominant features and must be reduced to reach the reference EF. In the same way, for $DMU_o = 209$, occupancy-rate and facility, which the first one is output and the second one is input, must be increased and decreased, respectively. In the last step, MOCE-DEA is developed and optimized to find the least adjustment to lead DMUs to reach reference EF. For $DMU_o = 99$, after adjusting inputs, an 18.04% improvement was obtained, which could successfully lead to EF_0. For $DMU_o = 209$, after adjusting input-output, it can reach EF_4, by improving its PS by 5.97%, 3.68%, 17.63%, 13.27%, and 0.1% in EF_0, EF_1, EF_2, EF_3 and EF_4, respectively. It shows the practicality of the stepwise gradual improvement of DMU's performance.

Regarding the proposed target setting framework is a local one that should be applied for DMUs separately, a question that might be raised is the efficiency of the whole framework. We want to declare that it is possible that for some DMUs, the framework cannot lead them to reference EF, which logically does make sense. The reason is existing some fundamental reasons for the DMUs. In this case, this framework shows that further investigations are necessary to discover the management and fundamental problems with the organization's monitoring, planning, and control.

7. Conclusion and future research

The main aim of the current research is to propose a data-driven decision support framework according to the XAI approach for DEA benchmarking. The proposed framework first creates multiple EFs by the primal CCR model with their corresponding labels to create multiple EFs to propose a hierarchical strategy for target setting. According to the RF classifier, the labeled dataset is classified, and the prediction's outcome is utilized to obtain the dominant features of DMUs. Then, it is possible to set the actionable and feasible target for a DMU by finding its equivalent peer in the reference EF by Euclidean distance. LIME can extract dominant features for them that positively contribute to the prediction's outcome. Afterward, MOCE-DEA is implemented for target setting by helping DMU_o to improve its PS to reach the ES of DMU_p by adjusting actionable and feasible features. For this purpose, DMU_o is considered a new member of the reference EF. Two objectives should be met: i) the distance of actionable and feasible features of DMU_o with DMU_p should be minimized; ii) the ES of DMU_o should be maximized. MOPSO solves this model, and the best solution with the highest ES among Pareto optimal solutions for each simulation batch is collected. For evaluating the model's performance and obtaining a vivid insight into the target setting, applied adjustments based on MOCE-DEA on DMU_o are verified by the primal CCR model. The robust evaluation shows that based on optimized values DMU_o can reach the reference EF in the primal CCR model as well. The successful performance of this approach makes it possible to set realistic targets. Besides, it can support DMs in finding feasible targets, and rest assured that target setting is based on vital inputs-outputs that are adjustable.

Determining targets for DMUs based on their inputs-outputs, which are used to evaluate them, and the possibility of combining these inputs-outputs is one of the strengths of this research. This approach can be used for all DMUs whose efficiency is evaluated relatively and at different levels. This approach allows DMUs and their managers to set target values based solely on the DMU's capabilities and achieve the benchmark unit with minimal changes. For future studies, it is suggested to use global explanation methods instead of LIME, a local explanation method. This can help to have a robust explanation of the problem. Also, it is possible to have a better insight into setting targets. Moreover, the strong mathematical background of SHapley Additive exPlanations (SHAP) can be implemented to find vital inputs-outputs to compare with LIME. One of the effective ways to apply CE is finding causality between features. It would be very illustrative to acquire causality between inputs-outputs in decision-making to improve the method for leading DMUs to be efficient. This perspective might help to adjust them even less than the expected amount because it may impact other DMUs by manipulating only the fewer inputs-outputs and, finally, the whole model. One of the challenges in implementing AI is the concept of "trust," which has attracted a lot of attention in recent years. It would be very interesting to develop the proposed framework in a trustable manner.

CRediT authorship contribution statement

Mustafa Jahangoshai Rezaee: Formal analysis, Conceptualization, Methodology, Validation, Writing – review & editing. **Mohsen Abbaspour Onari:** Formal analysis, Methodology, Project administration, Software, Validation, Writing – original draft, Writing – review & editing. **Morteza Saberi:** Formal analysis, Writing – review and editing

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

References

Abbaspour Onari, M., Jahangoshai Rezaee, M., 2022. A fuzzy cognitive map based on Nash bargaining game for supplier selection problem: a case study on auto parts industry. Oper. Res. 22 (3), 2133–2171.
Abbaspour Onari, M., Jahangoshai Rezaee, M., 2023. Implementing bargaining game-based fuzzy cognitive map and mixed-motive games for group decisions in the healthcare supplier selection. Artif. Intell. Rev. 1–34.
Abbaspour Onari, M., Yousefi, S., Jahangoshai Rezaee, M., 2021. Risk assessment in discrete production processes considering uncertainty and reliability: Z-number multi-stage fuzzy cognitive map with fuzzy learning algorithm. Artif. Intell. Rev. 54 (2), 1349–1383.
An, Q., Wu, Q., Zhou, X., Chen, X., 2021. Closest target setting for two-stage network system: an application to the commercial banks in China. Expert Syst. Appl. 175, 114799.
Aparicio, J., Cordero, J.M., Pastor, J.T., 2017. The determination of the least distance to the strongly efficient frontier in data envelopment analysis oriented models: modelling and computational aspects. Omega 71, 1–10.
Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al., 2020. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115.
Baek, C., Lee, J.D., 2009. The relevance of DEA benchmarking information and the least-distance measure. Math. Comput. Model. 49 (1–2), 265–275.
Balagopalan, A., Zhang, H., Hamidieh, K., Hartvigsen, T., Rudzicz, F., Ghassemi, M., 2022. The Road to Explainability Is Paved with Bias: Measuring the Fairness of Explanations. https://doi.org/10.1145/3531146.3533179.
Chakraborty, D., Elzarka, H., Bhatnagar, R., 2016a. Generation of accurate weather files using a hybrid machine learning methodology for design and analysis of sustainable and resilient buildings. Sustain. Cities Soc. 24, 33–41.
Chakraborty, D., Elzarka, H., Bhatnagar, R., 2016b. Generation of accurate weather files using a hybrid machine learning methodology for design and analysis of sustainable and resilient buildings. Sustain. Cities Soc. 24, 33–41.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. Eur. J. Oper. Res. 2 (6), 429–444.
Chen, Y., Tsionas, M.G., Zelenyuk, V., 2021. LASSO+ DEA for small and big wide data. Omega 102, 102419.
Cielen, A., Peeters, L., Vanhoof, K., 2004. Bankruptcy prediction using a data envelopment analysis. Eur. J. Oper. Res. 154 (2), 526–532.
Coello, C.A.C., Pulido, G.T., Lechuga, M.S., 2004. Handling multiple objectives with particle swarm optimization. IEEE Trans. Evol. Comput. 8 (3), 256–279.
Dandl, S., Molnar, C., Binder, M., Bischl, B., 2020. Multi-objective counterfactual explanations. In: International Conference on Parallel Problem Solving from Nature. Springer, Cham, pp. 448–469.
e Costa, C.A.B., de Mello, J.C.C.S., Meza, L.A., 2016. A new approach to the bi-dimensional representation of the DEA efficient frontier with multiple inputs and outputs. Eur. J. Oper. Res. 255 (1), 175–186.
Ebrahimi, B., Dhamotharan, L., Ghasemi, M.R., Charles, V., 2022. A cross-inefficiency approach based on the deviation variables framework. Omega 111, 102668.
El-Ssawy, W., Elhegazy, H., Abd-Elrahman, H., et al., 2022. Identification of the best model to predict optical properties of water. Environ. Dev. Sustain. https://doi.org/10.1007/s10668-022-02331-5.
Elhegazy, H., Chakraborty, D., Elzarka, H., Ebid, A.M., Mahdi, I.M., Aboul Haggag, S.Y., Abdel Rashid, I., 2022. Artificial intelligence for developing accurate preliminary cost estimates for composite flooring systems of multi-storey buildings. J. Asian Architect. Build Eng. 21 (1), 120–132. https://doi.org/10.1080/13467581.2020.1838288.
Esteve, M., Aparicio, J., Rodriguez-Sala, J.J., Zhu, J., 2023. Random forests and the measurement of super-efficiency in the context of Free Disposal Hull. Eur. J. Oper. Res. 304 (2), 729–744.
Fukuyama, H., Maeda, Y., Sekitani, K., Shi, J., 2014. Input–output substitutability and strongly monotonic p-norm least distance DEA measures. Eur. J. Oper. Res. 237 (3), 997–1007.
Fukuyama, H., Matousek, R., Tzeremes, N.G., 2022. Bank Production with Nonperforming Loans: A Minimum Distance Directional Slack Inefficiency Approach. Omega, 102706.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.
Hatamzad, M., Pinerez, G.C.P., Casselgren, J., 2022. Intelligent cost-effective winter road maintenance by predicting road surface temperature using machine learning techniques. Knowl. Base Syst. 247, 108682.
Hong, H.K., Ha, S.H., Shin, C.K., Park, S.C., Kim, S.H., 1999. Evaluating the efficiency of system integration projects using data envelopment analysis (DEA) and machine learning. Expert Syst. Appl. 16 (3), 283–296.
Jabbari, M., Sheikh, S., Rabiee, M., Oztekin, A., 2022. A collaborative decision support system for multi-criteria automatic clustering. Decis. Support Syst. 153, 113671.
Kanamori, K., Takagi, T., Kobayashi, K., Arimura, H., 2020. DACE: distribution-aware counterfactual explanation by mixed-integer linear optimization. In: IJCAI, pp. 2855–2862.

Keane, M.T., Kenny, E.M., Delaney, E., Smyth, B., 2021. If Only We Had Better Counterfactual Explanations: Five Key Deficits to Rectify in the Evaluation of Counterfactual Xai Techniques. *arXiv preprint arXiv:2103.01035*.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of ICNN'95-international Conference on Neural Networks, vol. 4. IEEE, pp. 1942–1948.

Khalili-Damghani, K., Abtahi, A.R., Tavana, M., 2013. A new multi-objective particle swarm optimization method for solving reliability redundancy allocation problems. Reliab. Eng. Syst. Saf. 111, 58–75.

Khezrimotlagh, D., Zhu, J., Cook, W.D., Toloo, M., 2019. Data envelopment analysis and big data. Eur. J. Oper. Res. 274 (3), 1047–1054.

Kwekha-Rashid, A.S., Abduljabbar, H.N., Alhayani, B., 2023. Coronavirus disease (COVID-19) cases analysis using machine-learning applications. Appl. Nanosci. 13 (3), 2013–2025.

Lai, M.-C., Huang, H.-C., Wang, W.-K., 2011. Designing a knowledge-based system for benchmarking: a DEA approach. Knowl. Base Syst. 24 (5), 662–671. https://doi.org/10.1016/j.knosys.2011.02.006.

Le, M.H., Afsharian, M., Ahn, H., 2021. Inverse frontier-based benchmarking for investigating the efficiency and achieving the targets in the Vietnamese education system. Omega 103, 102427.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Lee, C.Y., Cai, J.Y., 2020. LASSO variable selection in data envelopment analysis with small datasets. Omega 91, 102019.

Li, Z., Crook, J., Andreeva, G., 2017. Dynamic prediction of financial distress using Malmquist DEA. Expert Syst. Appl. 80, 94–106.

Lim, S., Bae, H., Lee, L.H., 2011. A study on the selection of benchmarking paths in DEA. Expert Syst. Appl. 38 (6), 7665–7673. https://doi.org/10.1016/j.eswa.2010.12.148.

Lozano, S., Soltani, N., Dehnokhalaji, A., 2020. A compromise programming approach for target setting in DEA. Ann. Oper. Res. 288 (1), 363–390. https://doi.org/10.1007/s10479-019-03486-7.

Meza, J., Espitia, H., Montenegro, C., Giménez, E., González-Crespo, R., 2017. MOVPSO: vortex multi-objective particle swarm optimization. Appl. Soft Comput. 52, 1042–1057.

Mishra, S., Sturm, B.L., Dixon, S., 2017. Local interpretable model-agnostic explanations for music content analysis. ISMIR 53, 537–543.

Nandy, A., Singh, P.K., 2020. Farm efficiency estimation using a hybrid approach of machine-learning and data envelopment analysis: evidence from rural eastern India. J. Clean. Prod. 267, 122106.

Peltola, T., 2018. Local Interpretable Model-Agnostic Explanations of Bayesian Predictive Models via Kullback-Leibler Projections. *arXiv preprint arXiv:1810.02678*.

Radovanović, S., Savić, G., Delibašić, B., Suknović, M., 2022. FairDEA—removing disparate impact from efficiency scores. Eur. J. Oper. Res. 301 (3), 1088–1098.

Ramón, N., Ruiz, J.L., Sirvent, I., 2018a. Two-step benchmarking: setting more realistically achievable targets in DEA. Expert Syst. Appl. 92, 124–131. https://doi.org/10.1016/j.eswa.2017.09.044.

Ramón, N., Ruiz, J.L., Sirvent, I., 2018b. Two-step benchmarking: setting more realistically achievable targets in DEA. Expert Syst. Appl. 92, 124–131.

Ramón, N., Ruiz, J.L., Sirvent, I., 2020. Cross-benchmarking for performance evaluation: looking across best practices of different peer groups using DEA. Omega 92, 102169.

Rezaee, M.J., Karimdadi, A., 2015. Do geographical locations affect in hospitals performance? A multi-group data envelopment analysis. J. Med. Syst. 39 (9), 1–11.

Rezaee, M.J., Jozmaleki, M., Valipour, M., 2018. Integrating dynamic fuzzy C-means, data envelopment analysis and artificial neural network to online prediction performance of companies in stock exchange. Phys. Stat. Mech. Appl. 489, 78–93.

Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why should i trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144.

Rostamzadeh, R., Akbarian, O., Banaitis, A., Soltani, Z., 2021. Application of DEA in benchmarking: a systematic literature review from 2003–2020. Technol. Econ. Dev. Econ. 27 (1), 175–222.

Ruiz, J.L., Sirvent, I., 2016. Common benchmarking and ranking of units with DEA. Omega 65, 1–9.

Ruiz, J.L., Sirvent, I., 2019. Performance evaluation through DEA benchmarking adjusted to goals. Omega 87, 150–157.

Seiford, L.M., Zhu, J., 2003. Context-dependent data envelopment analysis—measuring attractiveness and progress. Omega 31 (5), 397–408.

Shao, B.B., Shi, Z.M., Choi, T.Y., Chae, S., 2018. A data-analytics approach to identifying hidden critical suppliers in supply networks: development of nexus supplier index. Decis. Support Syst. 114, 37–48.

Tang, G., Yi, C., Liu, L., Yang, X., Xu, D., Zhou, Q., Lin, J., 2023. A novel transfer learning network with adaptive input length selection and lightweight structure for bearing fault diagnosis. Eng. Appl. Artif. Intell. 123, 106395.

Tsolas, I.E., Charles, V., Gherman, T., 2020. Supporting better practice benchmarking: a DEA-ANN approach to bank branch performance assessment. Expert Syst. Appl. 160, 113599.

Valero-Carreras, D., Aparicio, J., Guerrero, N.M., 2021. Support vector frontiers: a new approach for estimating production functions through support vector machines. Omega 104, 102490.

Valero-Carreras, D., Aparicio, J., Guerrero, N.M., 2022. Multi-output support vector frontiers. Comput. Oper. Res. 143, 105765.

Verma, S., Dickerson, J., Hines, K., 2020. Counterfactual Explanations for Machine Learning: A Review. *arXiv preprint arXiv:2010.10596*.

Wachter, S., Mittelstadt, B., Russell, C., 2017. Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harv. JL & Tech. 31, 841.

Yang, J.B., Wong, B.Y., Xu, D.L., Stewart, T.J., 2009. Integrating DEA-oriented performance assessment and target setting using interactive MOLP methods. Eur. J. Oper. Res. 195 (1), 205–222.

Zhang, Q., Wang, C., 2019. DEA efficiency prediction based on IG–SVM. Neural Comput. Appl. 31 (12), 8369–8378.

Zhong, K., Wang, Y., Pei, J., Tang, S., Han, Z., 2021a. Super efficiency SBM-DEA and neural network for performance evaluation. Inf. Process. Manag. 58 (6), 102728 https://doi.org/10.1016/j.ipm.2021.102728.

Zhong, K., Wang, Y., Pei, J., Tang, S., Han, Z., 2021b. Super efficiency SBM-DEA and neural network for performance evaluation. Inf. Process. Manag. 58 (6), 102728.

Zhu, Q., Wu, J., Ji, X., Li, F., 2018. A simple MILP to determine closest targets in non-oriented DEA model satisfying strong monotonicity. Omega 79, 1–8.

Zhu, Q., Aparicio, J., Li, F., Wu, J., Kou, G., 2022. Determining closest targets on the extended facet production possibility set in data envelopment analysis: modeling and computational aspects. Eur. J. Oper. Res. 296 (3), 927–939.

Zhu, H., Guo, S., Sheng, W., 2023. RDJCNN: a micro-convolutional neural network for radar active jamming signal classification. Eng. Appl. Artif. Intell. 123, 106417.