Journal of Turkish

Operations Management

Ranking decision making units using support vector machines & ideal decision making unit

Mehmet Güray Ünsal [1*], Volkan Soner Özsoy[2], Hacı Hasan Örkcü[3]

[1] Department of Technology and Knowledge Management, Faculty of Business and Administrative Sciences, Baskent University, Ankara, Turkey.
e-mail: mgunsal@baskent.edu.tr, ORCID no: https://orcid.org/0000-0001-7081-9060

[2] Department of Management Information Systems, Faculty of Business and Administrative Sciences, Aksaray University, Aksaray, Turkey.
e-mail: volkansonerozsoy@aksaray.edu.tr, ORCID no: https://orcid.org/0000-0002-6417-8946

[3] Department of Statistics, Faculty of Sciences, Gazi University, Ankara, Turkey.
e-mail: hhorkcu@gazi.edu.tr, ORCID no: https://orcid.org/0000-0002-2888-9580

*Corresponding Author

Article Info

Article History:
Received: 13.08.2024
Revised: 15.12.2024
Accepted: 07.02.2025

Keywords

Ranking,
Decision Making Unit,
Support Vector Machines,
Ideal DMU

Abstract

Ranking of decision making units (DMUs) is an important issue in a production process. Therefore, it is one of the most frequently studied subjects in the theory and practice studies in Data Envelopment Analysis literature. Recently, machine learning-based methods have also been used for crucial problems in literature such as ranking of DMUs and determining the efficient frontier. This study proposes a new hybrid approach to rank DMUs. This approach is based on the Support Vector Machines, which is a machine learning method, and the Ideal DMU, which has an important place in the DEA literature. The theoretical details of this method are explained, and the performance of the model is demonstrated through application and simulation studies.

1. Introduction

Ranking of units in the evaluation process is a critical field of study in sectors where performance measures are involved. In this context, ranking decision-making units (DMUs) according to their performance is a popular subject among researchers in the Data Envelopment Analysis (DEA) field. DEA is commonly defined as a performance measurement and ranking tool for homogeneous DMUs that use inputs and produce outputs in a production process. The first models of DEA methodology is proposed as CCR and BCC models by Charnes et al. (1978) and Banker et al. (1983), respectively. The CCR model operates under the Constant Returns to Scale (CRS) technology, meaning that the proportion of increase in output is equal to the proportion of increase in input. CCR is generally used for benchmarking and classifying DMUs. Classical approaches, such as cross-efficiency evaluation, are commonly used to rank DMUs. The cross-efficiency method relies on a matrix of weights obtained during the efficiency score calculation process of all DMUs (Sexton et al., 1986). In addition to these classical approaches, some studies related to the ranking of units have employed multivariate statistical methods (Zhu, 1999; Premachandra, 2001; Azadeh et al., 2007; Kuosmanen and Johnson, 2010; Golçalves et al., 2013; Ünsal and Örkcü, 2017; Ünsal and Nazman, 2020), common weight sets in DEA (Jahanshahloo, 2005; Liu and Peng, 2008; Ramón et al., 2012; Saati et al., 2012; Davoodi A., Rezai H.Z., 2012; Afsharian and Ahn, 2017; Bouzidis and Karagiannis, 2022), and weight restrictions, decomposition, and determination techniques (Thompson et al., 1986; Allen et al., 1997; Podinovski, 1999; Bal et al., 2007; Bal et al., 2008; Wang et al., 2009; Lam, 2010; Mecit and Alp, 2012; Mecit and Alp, 2013; Örkcü et al., 2015; Ünsal and Örkcü, 2016; Michali et al., 2021).

Recently, methods based on machine learning (ML) have become popular for efficiency frontier estimation in the evaluation process (Esteve et al., 2020; Esteve et al., 2022; Tsionas, 2022). Support vector machines (SVM) and support vector regression (SVR) are among the ML methods commonly integrated into the performance measurement process with DEA. Yeh et al. (2010) proposed a novel model that integrates rough set theory (RST) with SVM as a tool for evaluating input/output efficiency. Kao et al. (2013) used a hybrid approach combining DEA and SVM to estimate efficiency in web security. Rezaie et al. (2013) integrated SVM and DEA in the calculation of efficiency scores. Saradhi and Girish (2015) investigated the parameter tuning issue for SVM with the help of DEA. Yang and Dimitrov (2017) applied DEA and SVM to financial data. Wanke and Barros (2017) assessed airport efficiencies using a two-stage DEA-SVM approach. Fallahpour et al. (2018) predicted efficiency scores of units in the evaluation process by using an intelligence-based model that combines SVM and DEA. Zhang and Wang (2019) proposed an efficiency measurement model combining SVM with DEA using time series data. Valero-Carreras et al. (2021) used support vector regression (SVR), a specific model within the SVM family, to estimate the efficiency frontier and introduced the Support Vector Frontier (SVF) approach. They further investigated the extension of SVFs as multi-output Support Vector Frontiers (Valero-Carreras et al., 2022). Zhu et al. (2022) compared four ML-based algorithms, including SVM-DEA and improved SVM-DEA approaches, in the efficiency measurement process of manufacturing companies in China. Zhao et al. (2022) developed a super-efficient DEA-SVM (SE-DEA-SVM) method. Petridis et al. (2022) used an SVM-based model to classify mergers and acquisitions (M&A) using financial data, including negative data. Guerrero et al. (2022) introduced the DEA-machine learning approach (DEAM), which is based on SVR and the Structural Risk Minimization (SRM) principle As seen in the literature, particularly in the last 10 years, SVM and SVR-based methods have been applied to subjects related to DEA, such as the estimation of efficiency and ranking of units.

This study proposes an alternative approach for ranking DMUs based on SVM and an ideal DMU. The method, which was developed by assigning weights to the variables in the SVM classification process and introducing new variables defined as ratios of outputs to inputs, is examined in detail. The aim is to rank DMUs using an approach that has not been suggested in the literature, thereby proposing an alternative and innovative method for the machine learning-based ranking of DMUs, a subject recently studied in the literature. The rest of the paper is structured as follows: Section 2 provides information about Support Vector Machines (SVM) and the ideal DMU, and introduces the proposed ranking method, including its algorithm and a numerical example. Section 3 presents simulation studies to demonstrate the general performance of the proposed method in terms of rankings based on cross-efficiency evaluation. A real-world data application related to this study is discussed in Section 4. Finally, Section 5 concludes the study.

2. Methodology

This section discusses the proposed approach for ranking DMUs in the evaluation process. Although the methodology to be proposed in this study is mainly based on the Support Vector Machine method, it also includes adding an ideal DMU, using newly defined proportional variables, and multiplying the weights and variable values and score calculation processes. For this reason, the methodology section is divided into sub-sections in order to present the explanation clearly.

2.1. Support Vector Machines

Support Vector Machines (SVM) one of the traditional machine learning methods is firstly proposed by Cortes and Vapnik (1995). SVM is one of a machine learning methods which is used for the classification of the units in the analysis. SVM is mathematically complex and computationally expensive although it can still assist with huge data classification issues. Instead of a regression model and an algorithm, SVM offers a classification learning model. It manipulates the straightforward mathematical equation $y = wx + b$ to enable the linear domain division. If the classes in the original domain can be separated linearly (e.g., along a straight line or hyperplane), the model is referred to as a linear SVM model. Otherwise, it is named a nonlinear SVM model. In this study, only the linear SVM model is included as the frontier of DEA is linear.
In the SVM, a straight-line equation is derived as $wx' + b = 0$. This parameterized straight line, when applied to a data domain, creates two subdomains, which we may refer to as left subdomain and right subdomain (as we do in decision tree-based models), as D_1 and,D_2 and, which we defined as $D_1 = x: wx' + b \leq 0$ and $D_2 = x: wx' + b > 0$. Labels 1 for subdivision D_1 and -1 for subdivision D_2 can be used to designate the points lying inside these subdivisions. Consequently, the SVM's parametrization goal can be described as follows $wx' + b \leq 0, x \in D_1$ and $wx' + b > 0, x \in D_2$. Furthermore, Figure 1 shows the $D_1$and D_2 hyperplanes. $(wx' + b = 1)$ vector is the positive hyperplane, $(wx' + b = -1)$ is the negative hyperplane and $(wx' + b = 0)$ vector is the optimal hyperplane (Pecha and Horak, 2018; Ramasubramanian and Singh, 2019).

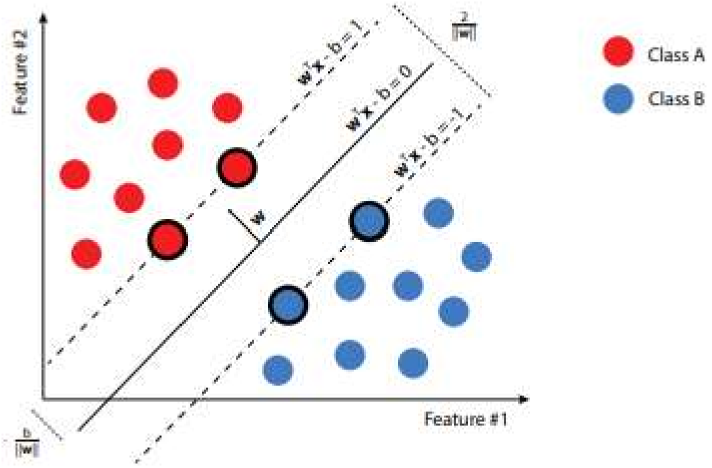

Figure 1. SVM Clustering Graph (Pecha and Horak, 2018)

To help in describing boundaries between the clusters, two straight lines (or hyperplanes) were created in the parametrization goals. It is aimed to search for parameter values that are the optimum values of the objective function that maximizes the distance between straight lines. The distance between straight lines is also called a margin. The standard distance equation between two parallel lines $y = mx + b_1$ and $y = mx + b_2$ can merely be used because these lines are parallel to each other (Boser et al., 1992; Hearts et al., 1998; Franc and Hlavac, 2002; Huang et al., 2011; Suthaharan, 2016).

$$d = \frac{(b_2 - b_1)}{\sqrt{m^2 + 1}} \tag{1}$$

where the slopes of the straight lines are,$m = w$ and their intercepts are $b_1 = \gamma + 1$ and $b_2 = \gamma - 1$. As this distance formula will be the measure for the optimization problem, we can establish the following by substituting these variables without loss of generality.

$$d = \frac{\mp 2}{\sqrt{ww'}} = \frac{\mp 2}{||w||^2} \tag{2}$$

To maximize the margin (the distance between the lines), it is necessary to simplify the equation. Therefore, by squaring both sides of the equation, and then dividing both sides of the equation by the value of 2, a simple mathematical relationship can be obtained.

$$\frac{d^2}{2} = \frac{1}{\frac{||w||^2}{2}} \tag{3}$$

The same goal is achieved by minimizing the denominator instead of maximizing the distance. Furthermore, the error function can be defined for the prediction error between $x \in D$ and y: $e = 1 - y(wx' + \gamma)$. By combining the aim and constraint, the following optimization problem can be created for the two-class SVM (Boser et al., 1992; Hearts et al., 1998; Franc and Hlavac, 2002; Huang et al., 2011; Suthaharan, 2016) :

$$\begin{gathered} min\ \frac{||w||^2}{2} \\ \text{Subject to : } e \leq 0\ or\ y(wx' + \gamma) \geq 1 \end{gathered} \tag{4}$$

2.2. Ideal and Anti-Ideal DMU

The production process is based on producing different amounts of output using different amounts of input. In this production process, the values of inputs and outputs of $DMU_j (j = 1, \ldots, n)$ are $x_{ij} (i = 1, \ldots, m)$ and,$y_{rj} (r = 1, \ldots, s)$ during evaluated of each DMU with m inputs and s outputs. If a virtual DMU uses the fewest inputs while producing the greatest amount of outputs, it is considered to be an ideal DMU (IDMU). A virtual DMU is considered be an anti-ideal DMU (ADMU) if it uses the most inputs to generate the fewest outputs. The inputs of IDMU can be determined as $x_i^{min} = \min_j x_{ij}, i = 1, \ldots, n$ while the outputs of IDMU can be determined as $y_r^{max} =$

$\max_j y_{rj}, r = 1, \dots, s$. Additionally, the inputs of ADMU can be determined as $x_i^{max} = \max_j x_{ij}, i = 1, \dots, n$ whereas the outputs of ADMU can be determined $y_r^{min} = \min_j y_{rj}, r = 1, \dots, s$ (Jahanshahloo et al., 2010; Wang et al., 2011; Lozano et al., 2020).

2.3. Ranking DMUs Using a Support Vector Machine & Ideal DMU

In this sub-section, SVM method and ideal DMU (IDMU) approach are linked in order to rank DMUs in the evaluation process. The process will be discussed by giving an algorithm and solving a basic example to indicate the details of the proposed method.

The steps of the proposed method can be defined as follows:
Step 1. Define an IDMU and add it into original dataset.
Step 2. Calculate new variables as output/input ratios.
Step 3. Define two classes such that one class consists of the original units and the other class has just IDMU.
Step 4. Run SVM to obtain weights for each variable based on output/input ratio.
Step 5. Use the absolute values of the weights are to focus on the impact of weights and multiply these absolute weights by the variable values of each unit to calculate a score value for each unit.
Step 6. Rank DMUs were performed according to the calculated score values.

To illustrate a simple numerical application, suppose there is a production process with 9 DMUs, 1 input (x), and 2 outputs (y1, y2), as shown in Table 1. Zhu (1999) defined output/input ratios by using output variables as numerators and input variables as denominators in fractional structures. If *m* is the number of inputs and *s* is the number of outputs, $m \times s$ new rational variables can be defined, such as y_1/x and y_2/x. At this point, following the steps, an IDMU can be added to the analysis in addition to the 9 original DMUs, as seen in the last row of Table 1. In the "Class" column of Table 1, the DMUs are divided into 2 groups: 0 and 1. Then, the classification-based method SVM is applied to these 10 units to calculate the weights for each variable.

Table 1. Numerical illustration of the proposed method

DMU	x	y_1	y_2	y_1/x	y_2/x	Class
1	11	25	35	2.2727	3.1818	0
2	8	15	32	1.875	4	0
3	16	38	42	2.375	2.625	0
4	21	35	45	1.6667	2.1428	0
5	5	10	14	2	2.8	0
6	4	8	10	2	2.5	0
7	25	40	38	1.6	1.52	0
8	20	25	48	1.25	2.4	0
9	6	11	15	1.8333	2.5	0
Ideal DMU	4	40	48	10	12	1

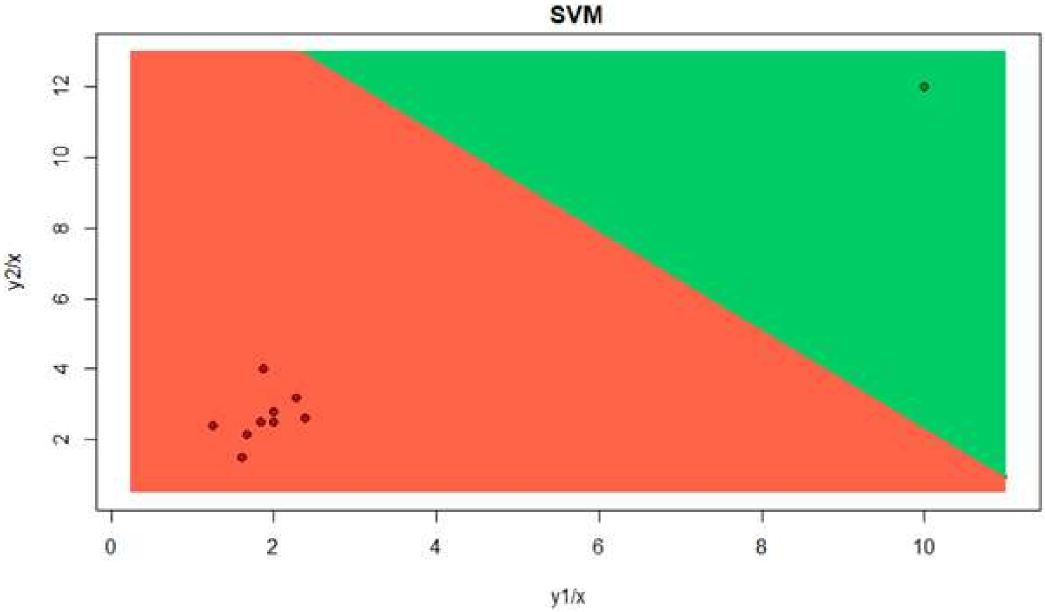

Figure 2. SVM margins for original DMUs and Ideal DMU

As shown in Figure 2, the SVM line passes through a straight line between red coloured original DMUs and green coloured IDMU. Since the margin length will be maximum due to the nature of SVM, the original DMUs and the IDMU are separated as much as possible. This is an advantage of our proposed method. As a result, a line-oriented weights that separates the IDMU and the original DMUs are obtained. Podinovski (1999), Gonçalves et al. (2013), Ünsal and Örkcü (2016) emphasize that it is possible and important to use absolute values in the variable weighting process. DEA models also have a restriction that the weight must be greater than zero. Thus, the scores of the original units are obtained by multiplication of absolute weights and new variables. The results of the proposed method can be seen in the last two columns of Table 2.

Table 2. Scores and Rankings of the Original DMUs for different methods

DMU	CCR		Cross Efficiency		The Proposed Approach	
	Scores	Rank	Scores	Rank	Scores	Rank
1	0.6562	3	0.9677	1	1.8479	2
2	1.0000	1	0.9051	3	1.9564	1
3	0.5181	8	0.9569	2	1.7113	3
4	0.6027	6	0.6949	7	1.2954	7
5	0.6562	4	0.8516	4	1.6262	4
6	0.5859	7	0.8283	5	1.5320	5
7	0.4453	9	0.6254	8	1.0749	9
8	0.9000	2	0.5827	9	1.2205	8
9	0.6392	5	0.7754	6	1.4698	6

3. Simulation Studies

The simulation study is conducted for different combinations of N (number of original DMUs), m (number of inputs) and s (number of outputs) sets. For each combination set, thousand trials are done, a spearman rank correlation coefficient between rankings of the cross efficiency which is an alternative method to rank DMUs as mentioned in Section 1 and the proposed method is obtained in each run and the average Spearman rank correlation value is obtained for each N, m, and s combination set. Furthermore, the data production process is managed according to the production function and methodology in Giraleas et al. (2012). The values obtained from the simulation study can be seen in Table 3 and Figure 3.

Table 3. Spearman correlations between cross efficiency and the proposed method

N	m	s	Spearman Corr.	N	m	s	Spearman Corr.
10	1	2	0.9796	10	2	1	0.8256
20			0.9859	20			0.8647
30			0.9872	30			0.8790
40			0.9895	40			0.8816
50			0.9905	50			0.8891
60			0.9905	60			0.8873
70			0.9919	70			0.8898
80			0.9924	80			0.8912
90			0.9927	90			0.8925
100			0.9930	100			0.8876
N	**m**	**s**	**Spearman Corr.**	**N**	**m**	**s**	**Spearman Corr.**
10	2	2	0.8085	10	2	3	0.7997
20			0.8517	20			0.8497
30			0.8649	30			0.8548
40			0.8720	40			0.8652
50			0.8784	50			0.8644
60			0.8791	60			0.8721
70			0.8811	70			0.8722
80			0.8819	80			0.8717
90			0.8803	90			0.8753
100			0.8837	100			0.8763
N	**m**	**s**	**Spearman Corr.**	**N**	**m**	**s**	**Spearman Corr.**
10	3	2	0.7043	10	3	3	0.6500
20			0.7817	20			0.7303
30			0.8025	30			0.7629
40			0.8150	40			0.7779
50			0.8259	50			0.7919
60			0.8323	60			0.7930
70			0.8363	70			0.7965
80			0.8403	80			0.8025
90			0.8418	90			0.8091
100			0.8459	100			0.8093

Note: The correlations are significant for α=0.05

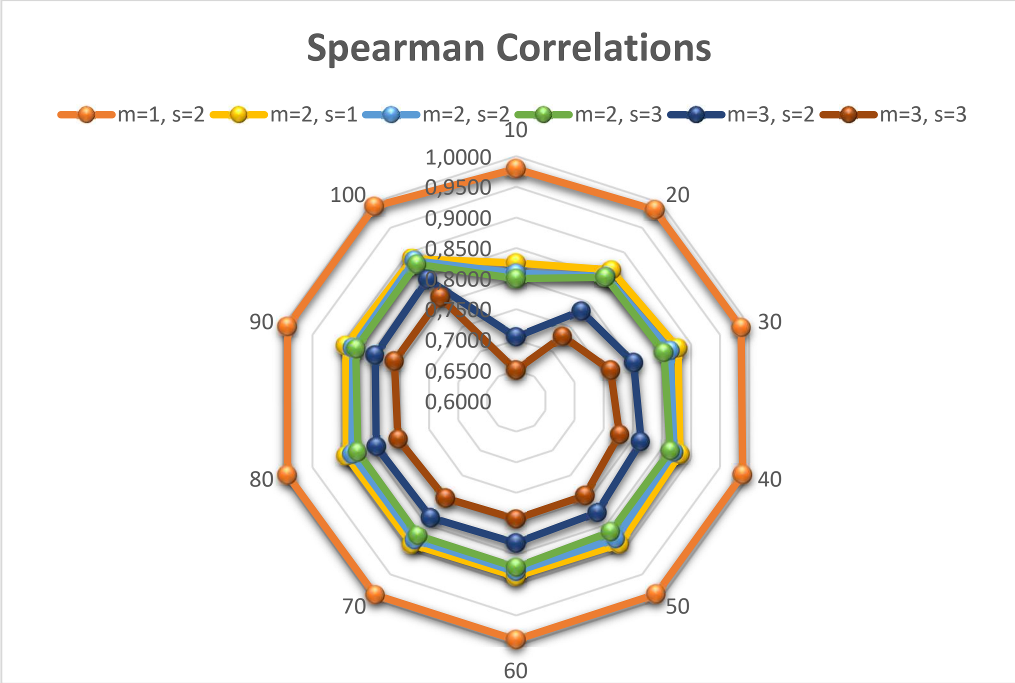

Figure 3. Radar Graph for Spearman correlations

As shown in Table 3 and Figure 3, in each combination set, rankings of the cross efficiency and the proposed method give high, positive and statistically significant correlations. As the number of units increases, the correlation between the cross efficiency and the proposed method increases in all combination sets. According to the simulation results, it can be generalized that the proposed method can be suggested as an alternative ranking method for decision-making units in the evaluation process.

4. A Real World Data Application

In this section, a real world data application study based on 21 national airports in Turkey is discussed to investigate an additional numerical example and compare the methods in this study. There are two inputs as $x_{1:}$ terminal domain (m^2), $x_{2:}$ expenses (1000 Turkish Liras) and two outputs as $y_{1:}$ income (1000 Turkish Liras), y_2: plain traffic$_{:}$ in the application. The data belongs to 2009 and is obtained from official website of General Directorate of State Airports Authority of Turkey. The calculated scores and rankings of the methods can be seen in Table 4.

Spearman rank correlation coefficients are calculated to see the relationship in the rankings of the methods. Accordingly, it is observed from Table 5 that there is a statistically significant, positive, and linear relationship between the rankings of the methods. It can also be seen that the highest correlation coefficient is calculated between the proposed method and the cross efficiency.

Table 4. Application data and Scores and Rankings of the methods

DMU (Airports)	x_1	x_2	y_1	y_2	CCR	Cross Efficiency	The Proposed Approach
Adana	10365	20610	20164	26242	1 [1]	0.9974 [1]	0.8066 [1]
Trabzon	23745	19446	13139	14892	0.6906 [6]	0.4843 [6]	0.3121 [8]
S.Demirel	5400	5654	419	2105	0.1822 [18]	0.1468 [18]	0.1093 [18]
Nevşehir	3500	8571	354	1524	0.1646 [19]	0.1128 [19]	0.0910 [19]
Erzurum	12950	10781	3507	5230	0.3563 [14]	0.2665 [13]	0.1769 [15]
Gaziantep	22790	18608	6533	8161	0.3589 [13]	0.2627 [14]	0.1708 [16]
Bursa	12716	11609	499	2228	0.0964 [21]	0.0734 [21]	0.0537 [21]
Çanakkale	1200	3029	236	1326	0.4159 [12]	0.273 [12]	0.2190 [12]
Çardak	16890	9492	1005	1774	0.1272 [20]	0.0865 [20]	0.0543 [20]
Diyarbakır	8085	7725	5150	8897	0.7910 [3]	0.6138 [3]	0.4242 [3]
Elazığ	1400	6014	826	2544	0.6923 [5]	0.4122 [8]	0.3583 [6]
Erzincan	1242	6105	541	1667	0.5114 [10]	0.2936 [11]	0.2585 [11]
Kars	2860	6791	953	2276	0.3056 [15]	0.2349 [15]	0.1929 [13]
Kayseri	11000	11922	8314	7281	0.7128 [4]	0.4780 [7]	0.3249 [7]
Malatya	3585	5220	2189	4566	0.5556 [8]	0.5031 [5]	0.3806 [5]
Mardin	1500	4496	497	2098	0.5290 [9]	0.3384 [10]	0.2782 [9]
Muş	1503	3333	515	1111	0.2849 [16]	0.2313 [16]	0.1886 [14]
Samsun	11500	12735	5916	7856	0.4796 [11]	0.3879 [9]	0.2720 [10]
Sivas	2217	3390	503	1232	0.2174 [17]	0.2028 [17]	0.1551 [17]
Tekirdağ	6521	7399	1256	17481	1 [1]	0.8426 [2]	0.6479 [2]
Ferit Melen	4410	7421	2613	6720	0.5836 [7]	0.5208 [4]	0.4047 [4]

Table 5. Spearman Correlations between the methods

	CCR- Cross	CCR-Proposed	Cross-Proposed
Spearman Correlations	0,9711	0,9659	0,9831

Note: The correlations are significant for α=0.05

5. Conclusion

The ranking of units in the evaluation process is a critical area of study in sectors where performance measures play a significant role. Consequently, it is one of the most frequently studied topics in both theoretical and applied research within the DEA literature. Recently, methods based on machine learning have gained popularity for efficiency frontier estimation in the evaluation process. This study introduces an alternative approach for ranking DMUs, utilizing SVM and an IDMU. The rankings derived from cross-efficiency and the proposed method exhibit high, positive, and statistically significant correlations. Moreover, as the number of units increases, the correlation between cross-efficiency and the proposed method improves across all combination sets in both simulation and application studies. These findings demonstrate that the proposed method can serve as a viable alternative ranking approach based on machine learning techniques in the existing literature. The alternative ranking method presented in this paper can be applied to various sectors (e.g., energy, education, healthcare) where DEA is used. The limitations of the study include the consideration of simulation studies for inputs and outputs with a specific structure. The possibility of expanding these simulation studies can be explored by repeating the variable data production process using different distribution structures and value ranges. Furthermore, the application study can be extended to real-world datasets containing a greater or smaller number of DMUs.

Contribution of Authors

In this research, Mehmet Güray Ünsal, contributed to the definition of the problem, literature research, determining and applying the method, analysis of the results, writing the article; Volkan Soner Özsoy, contributed to the definition of the problem, literature research, writing of the article and Hacı Hasan Örkcü, contributed to the definition of the problem, literature research, determination of the method.

Conflict of Interest
All authors have no conflicts of interest.

Data Availability
The data used in this study as a real world application are available to the public in official website of General Directorate of State Airports Authority of Turkey: https://dhmi.gov.tr/Sayfalar/AnnualReports.aspx

References

Afsharian, M., & Ahn, H. (2017). Multi-period productivity measurement under centralized management with an empirical illustration to German saving banks. *OR Spectrum*, *39*(3), 881-911. https://doi.org/10.1007/s00291-016-0465-8

Allen, R., Athanassopoulos, A., Dyson, R. G., & Thanassoulis, E. (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Annals of Operations Research*, *73*, 13-34. https://doi.org/10.1023/A:1018968909638

Azadeh, A., Amalnick, M. S., Ghaderi, S. F., & Asadzadeh, S. M. (2007). An integrated DEA PCA numerical taxonomy approach for energy efficiency assessment and consumption optimization in energy intensive manufacturing sectors. *Energy Policy*, *35*(7), 3792-3806. https://doi.org/10.1016/j.enpol.2007.01.018

Bal, H., & Örkcü, H. H. (2007). A goal programming approach to weight dispersion in data envelopment analysis. *Gazi University Journal of Science*, *20*(4), 117-125. Retrieved from https://dergipark.org.tr/en/pub/gujs/issue/7399/96859

Bal, H., Örkcü, H. H., & Çelebioğlu, S. (2008). A new method based on the dispersion of weights in data envelopment analysis. *Computers & Industrial Engineering*, *54*(3), 502-512. https://doi.org/10.1016/j.cie.2007.09.001

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational Learning Theory* (pp. 144-152). https://doi.org/10.1145/130385.130401

Bouzidis, T., & Karagiannis, G. (2022). An alternative ranking of DMUs performance for the ZSG-DEA model. *Socio-Economic Planning Sciences*, *81*, 101179. https://doi.org/10.1016/j.seps.2021.101179

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297. https://doi.org/10.1007/BF00994018

Davoodi, A., & Rezai, H. Z. (2012). Common set of weights in data envelopment analysis: a linear programming problem. *Central European Journal of Operations Research*, *20*(2), 355-365. https://doi.org/10.1007/s10100-011-0195-6

Esteve, M., Aparicio, J., Rodriguez-Sala, J. J., & Zhu, J. (2022). Random Forests and the measurement of super-efficiency in the context of Free Disposal Hull. *European Journal of Operational Research*. 304(2), 729-744. https://doi.org/10.1016/j.ejor.2022.04.024

Fallahpour, A., Kazemi, N., Molani, M., Nayyeri,S., Ehsani, M. (2018) An Intelligence-Based Model for Supplier Selection Integrating Data Envelopment Analysis and Support Vector Machine. *Interdiciplinary Journal of Management Studies*, 11 (2), 209-241. https://doi.org/10.22059/ijms.2018.237965.672750

Franc, V., & Hlavác, V. (2002). Multi-class support vector machine. In *2002 International Conference on Pattern Recognition,* 2, 236-239. IEEE. https://doi.org/10.1109/ICPR.2002.1048282

Giraleas, D., Emrouznejad, A., & Thanassoulis, E. (2012). Productivity change using growth accounting and frontier-based approaches: Evidence from a Monte Carlo analysis. *European Journal of Operational Research* 222 (3):673–683. https://doi.org/10.1016/j.ejor.2012.05.015

Gonçalves, A. C., Almeida, R. M., Lins, M. P. E., & Samanez, C. P. (2013). Canonical correlation analysis in the definition of weight restrictions for data envelopment analysis. *Journal of Applied Statistics*, *40*(5), 1032-1043. https://doi.org/10.1080/02664763.2013.772571

Guerrero, N. M., Aparicio, J., & Valero-Carreras, D. (2022). Combining Data Envelopment Analysis and Machine Learning. *Mathematics*, *10*(6), 909. https://doi.org/10.3390/math10060909

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, *13*(4), 18-28. https://doi.org/10.1109/5254.708428

Huang, G., Chen, H., Zhou, Z., Yin, F., & Guo, K. (2011). Two-class support vector data description. *Pattern Recognition*, *44*(2), 320-329. https://doi.org/10.1016/j.patcog.2010.08.025

Jahanshahloo, G. R., Hosseinzadeh Lotfi, F., Khanmohammadi, M., Kazemimanesh, M., & Rezaie, V. (2010). Ranking of units by positive ideal DMU with common weights. *Expert Systems with Applications: An International Journal*, *37*(12), 7483-7488. https://doi.org/10.1016/j.eswa.2010.04.011

Jahanshahloo, G. R., Memariani, A., ., Hosseinzadeh Lotfi, F, & Rezai, H. Z. (2005). A note on some of DEA models and finding efficiency and complete ranking using common set of weights. *Applied Mathematics and Computation*, *166*(2), 265-281. https://doi.org/10.1016/j.amc.2004.04.088

Kao, H. Y., Chang, T. K., & Chang, Y. C. (2013). Classification of hospital web security efficiency using data envelopment analysis and support vector machine. *Mathematical Problems in Engineering*, 2013, 542314 .
Kuosmanen, T., & Johnson, A. L. (2010). Data envelopment analysis as nonparametric least-squares regression. *Operations Research*, *58*(1), 149-160. https://doi.org/10.1287/opre.1090.0722

Lam, K. F. (2010). In the determination of weight sets to compute cross-efficiency ratios in DEA. *Journal of the Operational Research Society*, *61*(1), 134-143. https://doi.org/10.1057/jors.2008.138

Liu, F. H. F., & Peng, H. H. (2008). Ranking of units on the DEA frontier with common weights. *Computers and Operations Research*, *35*(5), 1624-1637. https://doi.org/10.1016/j.cor.2006.09.006

Lozano, S., Soltani, N., & Dehnokhalaji, A. (2020). A compromise programming approach for target setting in DEA. *Annals of Operations Research*, *288*(1), 363-390. https://doi.org/10.1007/s10479-019-03486-7

Mecit, E. D., & Alp, I. (2012). A new restricted model using correlation coefficients as an alternative to cross-efficiency evaluation in Data Envelopment Analysis. *Hacettepe Journal of Mathematics and Statistics*, *41*(2), 321-335. Retrieved from https://dergipark.org.tr/en/pub/hujms/issue/7755/101371

Mecit, E. D., & Alp, I. (2013). A new proposed model of restricted data envelopment analysis by correlation coefficients. *Applied Mathematical Modelling*, *37*(5), 3407-3425. https://doi.org/10.1016/j.apm.2012.07.010

Michali, M., Emrouznejad, A., Dehnokhalaji, A., & Clegg, B. (2021). Noise-pollution efficiency analysis of European railways: A network DEA model. *Transportation Research Part D: Transport and Environment*, *98*, 102980. https://doi.org/10.1016/j.trd.2021.102980

Örkcü, H. H., Ünsal, M. G., & Bal, H. (2015). A modification of a mixed integer linear programming (MILP) model to avoid the computational complexity. *Annals of Operations Research*, *235*(1), 599-623. https://doi.org/10.1007/s10479-015-1916-3
Pecha, M., Horák, D. (2020). Analyzing *l*1-loss and *l*2-loss Support Vector Machines Implemented in PERMON Toolbox. In: Zelinka, I., Brandstetter, P., Trong Dao, T., Hoang Duy, V., Kim, S. (eds) AETA 2018 - *Recent*

Advances in Electrical Engineering and Related Sciences: Theory and Application. AETA 2018. Lecture Notes in Electrical Engineering, vol 554. Springer, Cham. https://doi.org/10.1007/978-3-030-14907-9_2

Petridis,K., Tampakoudis, I., Drogalas, G., Kiosses, N. (2022) A Support Vector Machine model for classification of efficiency: An application to M&A. *Research in International Business and Finance* 61,101633. https://doi.org/10.1016/j.ribaf.2022.101633

Podinovski, V. V. (1999). Side effects of absolute weight bounds in DEA models. *European Journal of Operational Research*, *115*(3), 583-595. https://doi.org/10.1016/S0377-2217(98)00124-6

Premachandra, I. M. (2001). A note on DEA vs principal component analysis: An improvement to Joe Zhu's approach. *European Journal of Operational Research*, *132*(3), 553-560. https://doi.org/10.1016/S0377-2217(00)00145-4

Ramasubramanian, K., Singh, A. Machine Learning Using R, 2nd Edition, Apress, Delhi India, 2019. https://doi.org/10.1007/978-1-4842-4215-5

Ramón, N., Ruiz, J. L., & Sirvent, I. (2012). Common sets of weights as summaries of DEA profiles of weights: With an application to the ranking of professional tennis players. *Expert Systems with Applications*, *39*(5), 4882-4889. https://doi.org/10.1016/j.eswa.2011.10.004

Rezaie, V., Ahmad, T., Daneshfard, C., Khanmohammadi, M., & Nejatian, S. (2013). An Integrated modeling based on data envelopment analysis and support vector machines: A case modeling of Tehran Social Security Insurance Organization. *World Applied Sciences Journal (Mathematical Applications in Engineering)*, 21, 138-142. https://doi.org/10.5829/idosi.wasj.2013.21.mae.99940

Saati, S., Hatami-Marbini, A., Agrell, P. J., & Tavana, M. (2012). A common set of weight approach using an ideal decision making unit in data envelopment analysis. *Journal of Industrial & Management Optimization*, *8*(3), 623-637. http://dx.doi.org/10.3934/jimo.2012.8.623

Saradhi, V.V., & Girish, K.R. (2015). Effective Parameter Tuning of SVMs Using Radius/Margin Bound Through Data Envelopment Analysis, *Neural Process Letters* 41:125–138. https://doi.org/10.1007/s11063-014-9338-9

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, *1986*(32), 73-105. https://doi.org/10.1002/ev.1441

Suthaharan, S. (2016). Support Vector Machine. In: Machine Learning Models and Algorithms for Big Data Classification. Integrated Series in Information Systems, vol 36. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7641-3_9

Thompson, R. G., Singleton Jr, F. D., Thrall, R. M., & Smith, B. A. (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces*, *16*(6), 35-49. https://doi.org/10.1287/inte.16.6.35

Tsionas, M. (2022). Efficiency estimation using probabilistic regression trees with an application to Chilean manufacturing industries. *International Journal of Production Economics*, 249, 108492. https://doi.org/10.1016/j.ijpe.2022.108492

Ünsal, M. G. & Orkcu, H. H. (2017). Ranking decision making units with the integration of the multi-dimensional scaling algorithm into PCA-DEA. *Hacettepe Journal of Mathematics and Statistics*, 46(6), 1187-1197. Retrieved from https://dergipark.org.tr/en/pub/hujms/issue/38350/444401

Ünsal, M. G., & Nazman, E. (2020). Investigating socio-economic ranking of cities in Turkey using data envelopment analysis (DEA) and linear discriminant analysis (LDA). *Annals of Operations Research*, *294*(1), 281-295. https://doi.org/10.1007/s10479-017-2748-0

Ünsal, M. G., & Örkcü, H. H. (2016). A note for weight restriction in the ARIII model based on correlation coefficients. *Applied Mathematical Modelling*, *40*(23-24), 10820-10827. https://doi.org/10.1016/j.apm.2016.06.052

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. https://doi.org/10.1016/j.omega.2021.102490

Valero-Carreras, D., Aparicio, J., & Guerrero, N. M. (2022). Multi-output Support Vector Frontiers. *Computers & Operations Research*, *143*, 105765. https://doi.org/10.1016/j.cor.2022.105765

Wang, Y. M., Chin, K. S., & Luo, Y. (2011). Cross-efficiency evaluation based on ideal and anti-ideal decision making units. *Expert Systems with Applications*, *8*(38), 10312-10319. https://doi.org/10.1016/j.eswa.2011.02.116

Wang, Y. M., Luo, Y., & Liang, L. (2009). Ranking decision making units by imposing a minimum weight restriction in the data envelopment analysis. *Journal of Computational and Applied Mathematics*, *223*(1), 469-484. https://doi.org/10.1016/j.cam.2008.01.022

Wanke, P., & Barros, C.P. (2017) Efficiency thresholds and cost structure in Senegal airports, *Journal of Air Transport Management* 58, 100-112. https://doi.org/10.1016/j.jairtraman.2016.10.005

Yang, X., & Dimitrov, S. (2017). Data envelopment analysis may obfuscate corporate financial data: using support vector machine and data envelopment analysis to predict corporate failure for nonmanufacturing firms. *INFOR: Information Systems and Operational Research*, 55(4), 295-311. https://doi.org/10.1080/03155986.2017.1282290

Yeh, C. C., Chi, D. J., & Hsu, M. F. (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37(2), 1535-1541.
Zhang, Q., Wang, C. (2019). DEA efficiency prediction based on IG–SVM. *Neural Computing and Applications*, 31:8369–8378 https://doi.org/10.1007/s00521-018-3904-4

Zhao, L., Wang, B., He, J., Jiang, X. (2022) SE-DEA-SVM evaluation method of ECM operational disposition scheme. *Journal of Systems Engineering and Electronics* 33 (3), 600 – 611. https://doi.org/10.23919/JSEE.2022.000058

Zhu, N., Zhu, C., Emrouznejad, A. (2021) A combined machine learning algorithms and DEA method for measuring and predicting the efficiency of Chinese manufacturing listed companies, *Journal of Management Science and Engineering* 6(4), 435-448. https://doi.org/10.1016/j.jmse.2020.10.001

Zhu, J. (1998). Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities. *European Journal of Operational Research*, *111*(1), 50-61. https://doi.org/10.1016/S0377-2217(97)00321-4