ASSESSING AND CLASSIFICATION OF ACADEMIC EFFICIENCY IN ENGINEERING TEACHING PROGRAMS

Enrique De La Hoz[1]✉
Rohemi Zuluaga[2]
Adel Mendoza[3]

[1]Universidad Tecnológica de Bolívar, Colombia

[2]Universidad del Sinú, Colombia

[3]Universidad del Atlántico, Colombia

✉ edelahoz@utb.edu.co

ABSTRACT

This research uses a three-phase method to evaluate and forecast the academic efficiency of engineering programs. In the first phase, university profiles are created through cluster analysis. In the second phase, the academic efficiency of these profiles is evaluated through Data Envelopment Analysis. Finally, a machine learning model is trained and validated to forecast the categories of academic efficiency. The study population corresponds to 256 university engineering programs in Colombia and the data correspond to the national examination of the quality of education in Colombia in 2018. In the results, two university profiles were identified with efficiency levels of 92.3% and 97.3%, respectively. The Random Forest model presents an Area under ROC value of 95.8% in the prediction of the efficiency profiles. The proposed structure evaluates and predicts university programs' academic efficiency, evaluating the efficiency between institutions with similar characteristics, avoiding a negative bias toward those institutions that host students with low educational levels.

KEYWORDS

HOW TO CITE

Highlights

- *A standardized framework for evaluating, calculating, and predicting the performance of engineering education is presented.*
- *Assessment of homogeneous universities makes it possible to correctly determine academic performance.*
- *Comparison of equivalent entities yields different average efficiency values for the global analysis.*

INTRODUCTION

The teaching of science, engineering, technology, and mathematics (STEM) is a critical aspect of countries' development. Different studies reveal positive association factors between economic growth and the number of professionals in STEM areas (Hoeg and Bencze, 2017; Sharma and Yarlagadda, 2018; Suter and Camilli, 2019). Bianchi and Giorcelli (2020) demonstrate how countries with better levels of science education have higher levels of innovation, represented in patent-for-invention registrations. Corlu and Aydin (2016) show that teaching in STEM areas generates higher levels of business creation. Therefore, it is essential to generate objective assessment tools for teaching STEM-related careers. Thus, this study presents a data-based model to analyze the fundamental characteristics and relationships of engineering education programs and the

results of a standardized assessment to achieve academic efficiency. However, it is crucial to highlight the inequalities in terms of access, resources, and opportunities in higher education. So, to avoid the biases that represent the different levels in the basic academic competencies with which students access university education, the comparison of the programs must be fair, that is, comparing between equals. Consequently, this study identifies homogeneous groups of engineering programs to analyze and forecast their level of efficiency within their reference group.

This research is aligned within the area of learning analytics, promoting the use of data as input to support decision-making in educational environments. Universities are traditionally characterized as generating large volumes of data. However, Long and Siemens (2014) show that strategic and operational decision processes are developed

under empirical and subjective schemes. In the educational field, data are mainly used to generate descriptive schemes such as the generation of reports and external and internal communication processes and transactions as accreditation and government oversight requirements. Thus, specific areas are identified where educational institutions have begun to implement data-based models to represent complex situations. First, student dropout has been studied from a predictive viewpoint. For example, Berens et al. (2019) and Suresh, Asokan and Vinodh (2016) developed models to predict student dropout using socioeconomic and academic variables. Heidrich et al. (2018) modeled student dropout using contextual variables obtained from the interaction of students in the educational process and monitoring the frequency of students' use of resources to support education, information such as the library, and complementary activities, among others (González et al., 2018).

From the efficiency analysis approach, several studies use machine learning techniques and Data Envelope Analysis to generate estimates of productivity and competitiveness. Most of these studies have been developed in the commercial and industrial field (Aldamak and Zolfaghari, 2017). Among these studies, the contributions of Granadillo, Gómez and Herrera (2019) stand out; the authors integrate financial items and levels of operational performance to estimate productivity indicators in the chemical sector in Colombia. Other studies develop multistage models in a similar approach, analyzing variables' performance and implementing supervised and unsupervised data learning models with efficiency analysis models (Fuentes, Fuster, and Lillo-Bañuls, 2016; Mikušová, 2017; Visbal-Cadavid, Mendoza and Hoyos, 2019).

This study analyzes how the differences between the study units can affect the results of the efficiency metrics in the Data Envelopment Analysis (DEA) models. These are applied in an educational context, considering the results of the state tests of quality of education in Colombia SABER PRO and SABER 11 as a study base. Cluster analysis, DEA, and a predictive model based on Random Forest (RF) are employed, respectively, allowing identification of the relationships between variables, creation of homogeneous groups, measurement of efficiency, and forecasting future efficiency categories. To the best of our knowledge, this efficiency analysis approach has not been previously developed in educational contexts. Therefore, one of this study's contributions is to propose an alternative approach for estimating educational efficiency, incorporating the creation of homogeneous groups to make a comparison between equals of efficiency levels. Simultaneously, the estimation of the efficiency levels is established using a direct estimation method as a reference base, considering all the observations in the database.

Consequently, it is necessary to organize a method in three phases that allow the following questions to be answered. How to define university profiles of engineering education considering state exams at the secondary and university level? What is the academic efficiency of the identified engineering profiles? How to predict through machine learning the efficiency category of a university program belonging to the engineering training profiles created? Therefore, the main objective of this study is to evaluate and forecast the academic performance of Colombian engineering programs, creating a replicable and reproducible method, offering objective guidelines for decision-making in a higher education environment.

The analysis of the efficiency of educational services is a field of great dynamism in the scientific community. From the first approach to the concept of efficiency applying linear programming techniques developed by Charnes, Cooper and Rhodes (1978), in recent decades there has been a greater dynamism in the measurement of efficiency in educational environments, schools, universities, and even systems countries education. DEA allows the incorporation of different combinations of input and output variables. Through a bibliographic review, Ferro and D'Elia (2020) classified the input variables in models of educational efficiency in teaching (dissemination of knowledge), research (production of basic or applied knowledge), and extension activities (also known as such as transfers, public, community or "third mission"), and input variables classified as human (teaching and research) and non-human (physical and financial resources).

Efficiency and Education

The type of educational data is a vital aspect in determining efficiency. Thus, there are different reports and databases where the results of large-scale tests are presented (e.g., PISA, SABER PRO, GMAT or TIMSS). These data can be the result of micro aggregations represented by average values of each institution or country. On the other hand, there are data at the individual level, which represent the performance of students in their interaction with a standardized test, the grades obtained in a study period, or external variables related to social, economic, and geographical aspects (Aparicio, Perelman and Santín, 2020; Thanassoulis et al., 2017; Visbal-Cadavid, Martínez-Gómez and Escorcia-Caballero, 2020). The primary consideration of these approaches is to assume that all study units have the same conditions, resources, and infrastructure, which can have fundamental implications for determining efficiency levels. Furthermore, standardized tests have limitations, such as the range of possible student responses, the context of each student to associate their reality with the questions and answers in predetermined categories, in addition to the difficulty of the test associated with the existence or lack of specific training on exam topics. The literature related to the measurement of efficiency in educational processes has shown increasing dynamics in recent years (Witte and López-Torres, 2017). Therefore, it is possible to find different approaches to evaluate efficiency in this sector (Agasisti, Munda and Hippe, 2019; Gralka, Wohlrabe and Bornmann, 2019; Khan, Khan and Hameed, 2019; Tran and Villano, 2019), in addition to studies applied to the Colombian context (Visbal-Cadavid, Mendoza and Hoyos, 2019).

Using global management variables, articulating neural networks, and DEA models, Visbal-Cadavid, Martínez-Gómez

and Guijarro (2017) developed a model to predict the efficiency of public universities. In a similar approach, Aparicio, Cordero, and Ortiz (2019) make comparisons between efficiency analysis models using PISA tests as input data. Other authors have highlighted the relevance of grouping processes through cluster analysis to define complex association patterns related to the specific performance of variables generated in public reports, such as institutional budget, teacher salaries, campus area, and number of students, among others (Wolszczak-Derlacz, 2017). In addition, Nazir (2019) proposed limitations to carrying out the forecasting process based on the comparability and homogeneity of observations. To guarantee the efficiency of the prediction processes based on machine learning, it is essential to determine and characterize similarities between the study objects. It is also essential to highlight the investigations in which university institutions are defined and described in homogeneous groups. For example, Najadat, Althebyan and Al-Omary (2019) used non-hierarchical cluster techniques to create representative groups and identify leading institutions in the university context. Similarly, other investigations at the international level have estimated efficiency levels in large volumes of educational data. For example, Torres-Samuel et al. (2020) developed a Gaussian cluster model and a DEA

model to evaluate the technical efficiency of higher education institutions in Latin America considering macroeconomic variables and research, innovation, and development results.

MATERIALS AND METHODS

Data Collection

Through a rational analysis, the generic components of the state test for secondary education, known as SABER 11, were identified as input variables and the components of the state test for higher education, SABER PRO, as output variables of an efficiency model. With the previously refined and selected information, the following phases were carried out: i) A cluster analysis using the unsupervised learning algorithm k-means to identify the formation of homogeneous groups in the data, associated with the results of the SABER tests; ii) An academic efficiency analysis was developed under an exit optimization approach to determine academic efficiency profiles (AEP); iii) A predictive model was defined to classify and predict belonging to an academic efficiency profile of an engineering program through Random Forest (RF) and Decision Tree (DT). The process flow and the articulation of the techniques is shown in Figure 1.
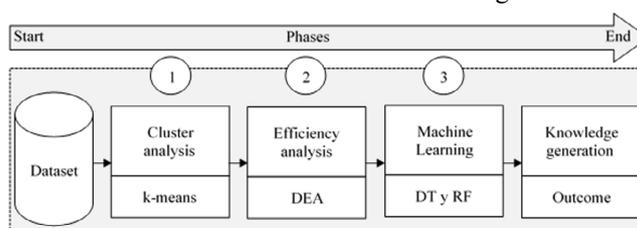

Figure 1: Research methodology

The components were identified through a rational analysis. The database used contains 12,411 observations, each of which represents a student. These observations come from 135 universities from Colombia (public: 30.37%, private: 69.63%) and eight academic degrees (civil, electromechanical, electrical, electronic, industrial, industrial automation, mechatronic, and chemical engineering). The base data were summarized by combining universities and undergraduate programs, leaving a total of 265 observations for analysis. It should be noted that universities do not have the same number of academic programs. In addition, the data

come from the databases of the Colombian Institute for the Evaluation of Education (ICFES).

The names, mean, and standard deviation of the study variables are reported in Table 1. The suffix for variables labeled S11 corresponds to the high school level test and SPRO corresponds to the college level assessment. It should be noted that the scale of academic competencies is 0-100, but the scale of the variable Formulation of engineering projects (FEP_SPRO) is 0-200. In addition, the mean and standard deviation belong to the results of the academic competencies evaluations of the school (S11) and the university (SPRO).

Variable	Full name	Average	Standard deviation
MATH_S11	Math	64.32	11.87
ENG_S11	English	60.78	10.03
NS_S11	Natural sciences	60.71	10.12
CS_S11	Citizenship skills	63.95	11.16
CR_S11	Critical reading	61.80	14.30
QR_SPRO	Quantitative reasoning	77.42	22.67
CS_SPRO	Citizenship skills	62.20	27.67
ENG_SPRO	English	59.19	28.99
WC_SPRO	Writing communication	67.50	25.49
CR_PRO	Critical reading	53.70	30.00
FEP_SPRO	Formulation of engineering projects	145.48	40.12

Table 1: Study variables, 2009-2018 (source: Delahoz-Dominguez, Zuluaga and Fontalvo-Herrera, 2020)

Academic Competences

SABER 11 is an evaluation of the level of secondary education in Colombia to provide educational institutions information on the development of basic skills that a student must develop during their time in school (ICFES, 2020). On the other hand, SABER PRO is an assessment aimed at higher education students close to graduation. Both evaluations are carried out by the Colombian Institute for the Evaluation of Education (ICFES) to measure the quality of all public or private educational institutions. The SABER PRO assessment is a mandatory requirement for all students who wish to acquire a professional degree in Colombia (ICFES, 2020). A student can also take the assessment only if they have passed 75% of the academic credits.

Cluster Analysis

For the development of the first phase of the proposed method, a non-hierarchical cluster analysis was carried out through the k-means algorithm (Clayman, Srinivasan and Sangwan, 2020; Oyelade et al., 2019). This algorithm randomly selects k points from the original data set to add as the initial clustering center. First, each unit in the data set is considered a point. Then, the distance between the data points and the core of the group is determined using the Euclidean equation, and the data set is preliminarily grouped by distance. Finally, the average distance of the observations in each group is calculated, the center of mass of the group is adjusted, and the final result of the grouping is obtained through multiple iterations. The Silhouette test (Menardi, 2011) assesses the quality of the membership of the observations, providing weights that oscillate between values of -1 and 1, where -1 is the evaluation of the observations better represented in another group; observations that are in the boundary between two clusters take the value of 0 and those that are well matched to the current group take the value of 1.

Data Envelopment Analysis

The key concept of DEA is the evaluation of the efficiency of the decision-making units that interact within a competition and development sector. Also known as border analysis, DEA has become the standard for the development of processes for comparing, measuring, and evaluating efficiency in productive organizations (Pawsey, Ananda and Hoque, 2018). Different approaches can be taken from the viewpoint of DEA analysis for educational purposes, for example, Amara, Rhaiem and Halilem (2020) evaluated research efficiency of Canadian scholars, considering aspects such as public funding seniority and university reputation.

The DEA-CCR model, known in the literature as technical efficiency, is the relationship between the weighted sum of the outputs and that of the inputs. The CCR model seeks to maximize the efficiency of a decision-making unit, within a group of reference organizations, through the optimal weights related to the input and output variables (Benicio and Mello, 2015). The optimization model associated with the DEA-CCR model endogenously calculates the weighting of the performance criteria and the result of the variables to achieve the maximum or minimum value of the objective function (Sinuany-Stern, Mehrez and Hadad, 2000). In addition, the result of the academic program results in the sum of the scores of the students who take the test for each institution adjusted by the arithmetic mean. However, the arithmetic mean is affected by outliers and, in particular, in the case of standardized tests, by the number of students taking the test. Therefore, when hypothetically considering two university programs A and B, with the same average results in the exams, different behaviors of the variability can be classified in the same technical efficiency category.

It should be noted that the DEA models assume that the input information of the models is accurate. However, in most cases, the input and output variables are imprecise, wrong, and biased. For example, when evaluating an individual student's performance through a standardized test, which was designed by experts who select the questions for each topic and dimension, adjusting the order and quantity of the questions according to standardized criteria of reading speed, comprehension, and analysis (Wolszczak-Derlacz, 2017). In summary, the process of creating the standardized test is a summation of subjective decision processes, which together generate noise in the final result. This is aligned with the main objective of this research to compare levels of efficiency among equals. Thus, although the exams are the same for all, the students do not come from the same context and, at a certain point, a university can be efficient in generating knowledge for its students, considering the student's initial learning inputs in relation to other universities (Duan, 2019; Ghasemi et al., 2020). Finally, the research model is presented below:

$$\max_{u,v} h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \quad (1)$$

subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \; j = 1, \ldots, n \quad (2)$$

where

n : Each of which uses the same inputs (in different quantities) to obtain the same outputs (in different quantities).

x_{ij} : Input quantities i consumed by the j-th DMU.

x_{io} : Amount of input i consumed by the DMU o.

y_{rj} : Observed quantities of outputs r produced by the j-th DMU.

y_{ro} : The observed quantity of the output r produced by the DMU o.

     : Weighting of the virtual output.

v_i : Weighting of the virtual input.

On the other hand, for the application of the DEA model, we start from the construction of the conceptual scheme that relates the input and output variables of the model. The inputs of the DEA model correspond to the competencies evaluated in SABER 11 (Math_S11, ENG_S11, NS_S11, CS_S11, and CR_S11) and the outputs of the model are the competences evaluated in SABER PRO (QR_SPRO, CS_SPRO, ENG_SPRO, WC_SPRO, CR_PRO, and FEP_SPRO).

Random Forest

The RF model is an assembly-type method based on the recurring and growing construction of multiple DTs through a bootstrapping aggregation process (Breiman, 2001). That is, multiple DTs are created with different composition of variables in such a way that each tree produces an independent result. A democratic process is then carried out where a category is assigned according to the resulting class with the most votes in general. This characteristic of generating separate responses for each Decision Tree and then joining them in a general prediction produces robust models that are less susceptible to extreme values and overfitting than a simple Decision Tree, thus improving the model's predictability and classification. The RF model presents a variable selection technique; in this way, it is possible to handle data sets with a large number of variables without using previous processes to reduce dimensions. The model also identifies the importance of the variables for the correct classification of the observations through a permutations test.

The success of the classification process occurs by minimizing the difference between the predicted value and the actual value. This relationship is described by the metrics True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), and F1 Score. The metrics used to assess performance will be correct classification rate (C), positive predictive value (PPV), negative predictive value (NPV), sensitivity (S) and specificity (E), and Area under ROC. It should be noted that: specificity indicates the ability of the estimator to detect negative cases; sensitivity indicates the ability of the estimator to detect positive cases; F1 score is the harmonic mean between precision and recall (Sun, Liu and Wang, 2020), the optimal value is 1, this indicates that there is a perfect precision and recall; and the Area under ROC represents the rate of PT and FP at various discrimination thresholds. A model with a perfect classification will have Area under ROC = 1. On the other hand, a totally random model will return a value of Area under ROC = 0.5.

Finally, for the training of the model, the cross-validation (10-folds) method was used and 70% of the data set, on the other hand, 30% of the data set is used for the evaluation of the model. The inputs of this model are the academic competencies of SABER 11 and SABER PRO, the output is the category of the efficiency group (Group 1 efficient, Group 1 Non-efficient, Group 2 efficient, Group 2 Non-efficient).

RESULTS

The proposed methodology consists of three phases. Detailed results for each phase are presented below.

First phase: Cluster Analysis

For the first phase of academic profiling, the k-means algorithm was used, varying the number of centroids from two to ten, to identify the optimal number of groups to which the university programs belonged. The test's highest value corresponds to the formation of two groups with a Silhouette test value of 0.49 (see Figure 2).
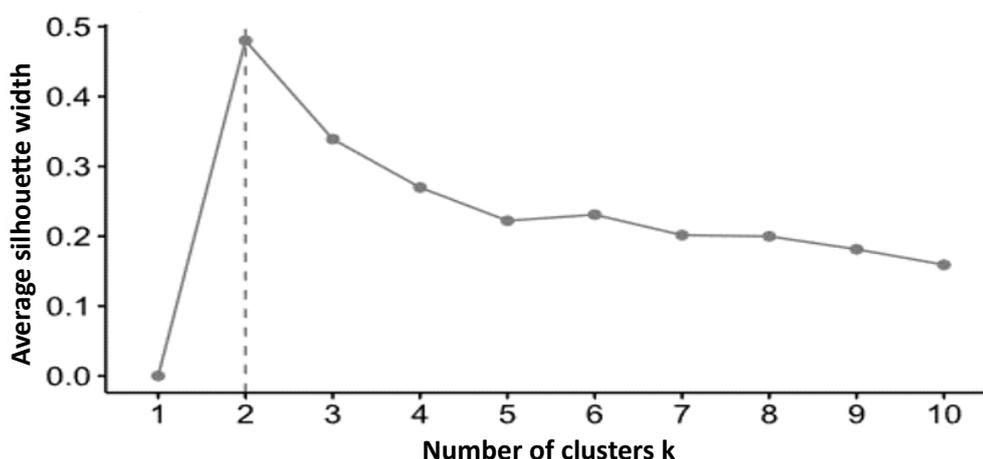

Figure 2: Analysis of the Silhouette test for optimal group conformation

Consequently, the representative elements of each profile are analyzed as shown in Figure 3. The university programs of cluster 1 are observed to have on average a higher score in all academic competences than the programs of Group 2. Thus, the competencies that will characterize Group 1 are those with an average higher than 70 points: ENG_SPRO, CS_SPRO, CR_SPRO, QR_SPRO, and MATH_S11. On the other hand, the characterizing competencies of Group 2 are those with an average higher than 55 points: QR_SPRO, MATH_S11, CR_S11, CS_S11, and NS_S11. Therefore, profile one can be contextualized as university programs with a high level of university entrance competencies and profile two as those with a medium-low level of university entrance competencies.
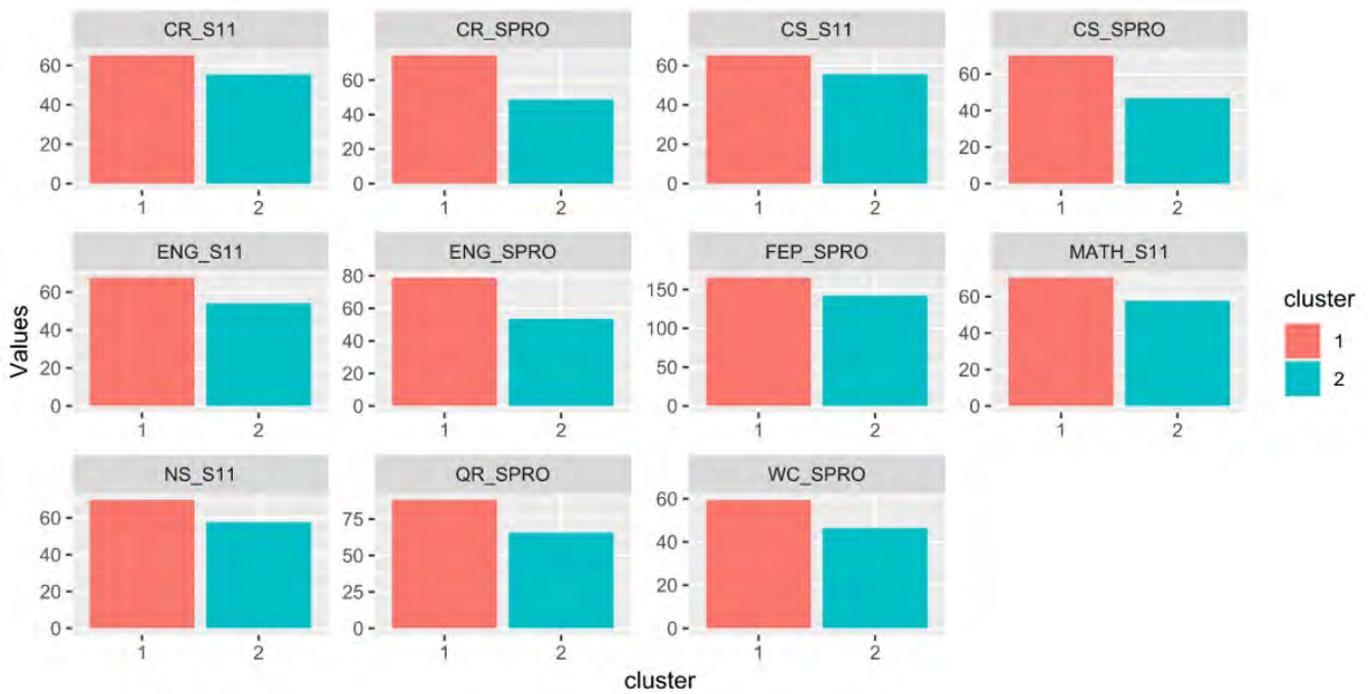
Figure 3: Contribution of academic competencies in clusters

Results by Cluster

In this phase, to improve the interpretation of academic profiles, the two-dimensional representation of the groups generated by the k-means algorithm was developed (see Figure 4). There is a wide separation of both groups, cluster 1 located on the left side is represented as a homogeneous group, showing low variability in the results of the SABER tests among its members. Group 2 is located on the right side of the map and presents greater variability among its members, represented in the area they occupy.
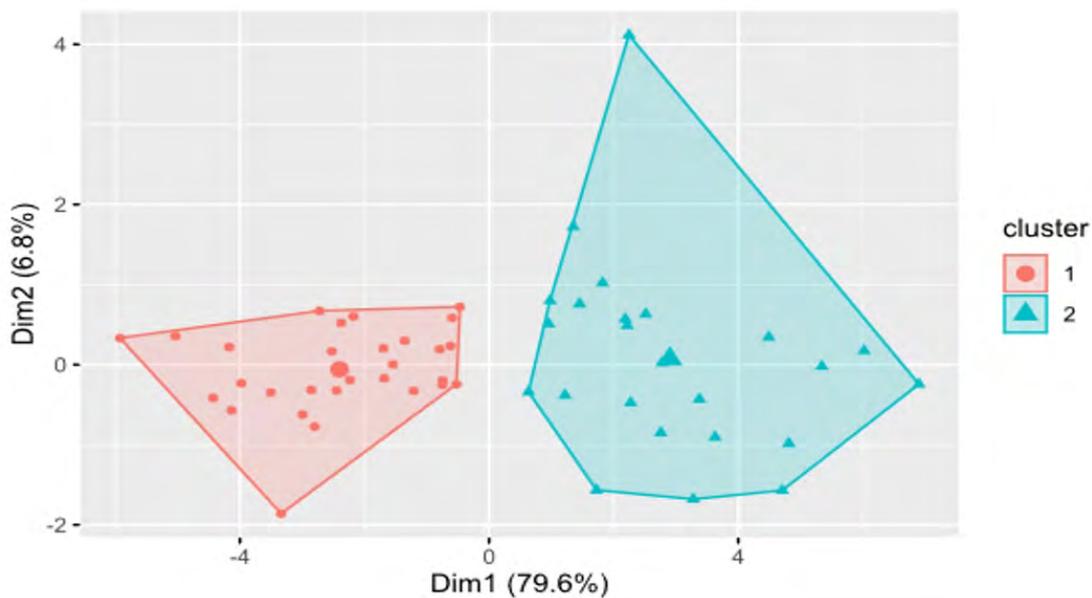
Figure 4: Graphic representation of the *k*-means cluster

Second phase: Efficiency Analysis

For the second phase of the method, the efficiency of the academic programs was calculated, adjusted by the previously determined efficiency profiles. The efficiency results by groups are presented in Table 2. The average efficiency of Profile 1 is 0.973 with a standard deviation of 0.029, 28 programs are determined as efficient out of a total of 111 members of this profile. In contrast, the average efficiency of Profile 1 is 0.941, with a deviation of 0.052, considering 34 university programs out of 154 members of this profile as efficient. The average efficiency for institutions without clustering by the cluster is 0.832, and the number of inefficient academic programs ranges from 223 in the global analysis to 203 in the group-adjusted analysis.

Variable	No cluster	Clustering	
	General	Profile 1	Profile 2
Efficient programs	42 (15.85%)	28 (25.23%)	34 (22.08%)
Non-efficient programs	223 (84.15%)	83 (74.77%)	120 (77.92%)
Minimum level of efficiency	0.681	0.853	0.731
Average efficiency	0.832	0.973	0.941
Standard deviation	0.097	0.029	0.052
Number of programs (DMUs)	265	111	154

Table 2: Summary of the efficiency model

The integration of the results of phase one and two allows the creation of four AEP: 1) Efficient institutions in group one, 2) Non-Efficient institutions in group one, 3) Efficient institutions in group two, 4) Non-Efficient institutions in group two. Table 3 presents the value of the efficiency of some units of the study concerning the group to which they belong. The institutions that present a value equal to one in their level of efficiency are be classified as efficient and the rest as "Non-efficient.".

No cluster		Cluster			
		G1		G2	
DMU	Efficiency	DMU	Efficiency	DMU	Efficiency
U1	0.94	U24	1	U1	1.00
U2	0.91	U36	0.99	U2	0.99
U3	0.82	U37	0.98	U3	0.82
U4	0.85	U68	0.97	U4	0.87
U5	1.00	U69	0.99	U5	1.00
U6	0.86	U70	0.99	U6	0.91
U7	0.89	U71	1.00	U7	0.91
U8	0.93	U72	1.00	U8	0.96
U9	0.88	U73	0.99	U9	0.91
U10	1.00	U83	0.96	U10	1.00
U11	0.89	U84	1.00	U11	0.91
U12	0.93	U87	0.95	U12	1.00
U13	0.87	U88	0.98	U13	0.90
U14	0.87	U90	1.00	U14	0.91
U15	0.95	U105	1.00	U15	0.98

Table 3: Selected DMU Efficiencies

To contextualize the results, the variable "High-quality accreditation" of the academic program is used as an adjustment factor to provide tools to justify the differences in the levels of efficiency between the profiles. In Table 4, when comparing the results of the global efficiency analysis with the program accreditation variable, there is a higher proportion of accredited universities in the efficient universities profile.

Group	Quality Accreditation		
	Not	Yes	Total
Efficient	30.95%	69.05%	42 (15.85%)
Non-Efficient	60.09%	39.91%	223 (84.15%)

Table 4: Distribution of efficient DMUs without adjusting for academic profiles

Results of the efficiency levels adjusted by the quality accreditation variable are presented in Table 5. The results reveal a difference in the behavior of the groups. In the AEP of efficient units of Group 1, 92.86% of the programs are accredited, while in the AEP of efficient units of Group 2, only 15.83% have the accreditation.

AEP	Group distribution	Quality Accreditation	
		No	Yes
Group 1 Efficient	10.57%	7.14%	92.86%
Group 1 Non-efficient	31.32%	22.89%	77.11%
Group 2 Efficient	12.83%	73.53%	26.47%
Group 2 Non-efficient	45.28%	84.17%	15.83%

Table 5: Efficiency distribution for DMUs adjusted by academic profiles

A similar analysis is presented in Table 6. The academic program, considering the distribution by type of specific program, is taken into account using the global estimate of efficiency.

Engineering program	Efficient	Non-efficient	Total
Civil	13.04%	86.96%	17.36%
Electromechanical	20.00%	40.00%	1.89%
Electrical	22.22%	77.78%	3.40%
Electronic	22.64%	77.36%	20.00%
Industrial	13.54%	86.46%	36.23%
Industrial automation	8.57%	91.43%	13.21%
Mechatronic	25.00%	75.00%	3.02%
Chemistry	7.69%	92.31%	4.91%

Table 6: Distribution of DMUs in efficient groups by program without adjustment for profiles

Table 7 reports the distribution considering the adjustment by academic group. There are differences in the distribution of efficiency categories concerning the results of Table 6, for example, the Industrial Engineering program went from having a total of 13.54% of efficient units under the total efficiency measurement scheme to 22.92% considering the sum of the categories of efficient units of Profiles 1 and 2.

One of the greatest advantages of the DEA model is determining the number of resource units that a DMU will increase/decrease to reach the efficiency frontier. In Tables 8 and 9, these values are presented as objectives that an academic program should achieve to reach the efficiency frontier established by proposed model. For example, in Table 8, which reports the results for the global efficiency model, the most significant opportunity for improvement is associated with proficiency in the English language. On the other hand, when analyzing the results of the model adjusted by profiles (Table 9), the most significant opportunity for improvement is different for each of the groups.

ERIES Journal
volume 14 issue 1

Electronic ISSN
1803-1617

Printed ISSN
2336-2375

47

Engineering program	Group 1 Efficient	Group 1 Non-Efficient	Group 2 Efficient	Group 2 Non-Efficient	Total
Civil	10.87%	34.78%	4.35%	50.00%	17.36%
Electromechanical	0.00%	20.00%	40.00%	40.00%	1.89%
Electrical	22.22%	55.56%	0.00%	22.22%	3.40%
Electronic	15.09%	30.19%	18.87%	35.85%	20.00%
Industrial	6.25%	20.83%	16.67%	56.25%	36.23%
Industrial automation	8.57%	42.86%	2.86%	45.71%	13.21%
Mechatronic	12.50%	12.50%	25.00%	50.00%	3.02%
Chemistry	15.38%	76.92%	7.69%	0.00%	4.91%

Table 7: Distribution of DMUs in efficient groups by academic program adjusted by profiles

	MATH_S11	ENG_S11	NS_S11	CS_S11	CR_S11
Mean	1.92	2.24	1.55	1.50	1.11
Standard deviation	1.77	1.96	1.61	1.50	1.22
Number of programs	150	124	110	76	104
Maximum	9.55	11.4	7.74	9.77	6.77

Table 8: Improvement targets for input variables under full efficiency estimation

Cluster 1					
	MATH_S11	ENG_S11	NS_S11	CS_S11	CR_S11
Mean	2.28	1.55	1.28	1.50	0.93
Standard deviation	1.87	1.91	1.76	1.91	1.37
Number of programs	85	46	44	44	33
Maximum	7.44	7.57	8.43	11.61	5.65
Cluster 2					
	MATH_S11	ENG_S11	NS_S11	CS_S11	CR_S11
Mean	1.10	2.07	1.41	1.41	1.23
Standard deviation	1.41	2.51	1.39	1.85	1.21
Number of programs	30	71	76	31	72
Maximum	6.97	16.89	7.17	9.84	6.48

Table 9: Improvement objectives for input variables under adjustment for academic profile

Group 1 presents more significant opportunities for improvement in mathematical competence. In Group 2, the improvement is mainly due to mastery of English, with mathematics being only the fourth competence with the most significant room for improvement for this profile. These results indicate that the comparison between equals allows a better conceptualization of the efficiency results and, therefore, objective, and specific tools for developing improvement strategies in the higher education sector.

Third phase: Machine Learning

In the third phase, the RF model predicts the academic efficiency profile to which an academic program belongs. The predictors of the model are the results of academic competencies and the result will correspond to one of the four AEPs. Thus, the model obtained a mean precision of (0.833) during the training phase using 10-folds in cross-validation. Consequently, Table 10 shows the performance metrics of the RF model training, the mean, lower limit of the confidence interval (LL.CI – 5% significance) and upper limit of the confidence interval (UL.IC – 5% significance) of Sensitivity, Specificity and F1 Score.

	RF Model								
	Sensitivity			Specificity			F1 Score		
1	0.76	0.75	0.77	0.93	0.92	0.94	0.73	0.71	0.74
2	0.56	0.55	0.57	0.9	0.89	0.91	0.67	0.65	0.68
3	0.75	0.74	0.76	0.96	0.95	0.97	0.78	0.77	0.80
4	0.58	0.57	0.59	0.91	0.90	0.92	0.68	0.66	0.70
5	0.57	0.56	0.58	0.90	0.89	0.91	0.61	0.6	0.63
6	0.58	0.57	0.59	0.92	0.91	0.93	0.68	0.66	0.70
7	0.75	0.74	0.76	0.96	0.95	0.97	0.78	0.77	0.80
8	0.88	0.87	0.89	0.98	0.97	0.99	0.93	0.91	0.94
9	0.63	0.62	0.64	0.94	0.93	0.95	0.71	0.7	0.73
10	0.75	0.74	0.76	0.95	0.94	0.96	0.83	0.82	0.85

Table 10: Improvement objectives for input variables under adjustment for academic profiles

Then, the Area under ROC value was equal to 95.8% for the RF predictions (ROC curve). Consequently, cross-validation is performed to generate coherence on the model. In this case, Table 10 shows a reduction in the standard deviation measurement for the precision and Area under ROC metric results with values of 0.739 and 0.035, respectively (see Table 11).

Based on the model results, the importance of the predictors can be determined (see Table 12). Thus, what role academic competencies play in efficiency can be observed for each group. For example, in cluster 1, to detect the efficient study

units, the variables that have a positive relationship in the model correspond to QR_SPRO, CR_SPRO, CS_SPRO, WC_SPRO, and FEP_SPRO. On the other hand, the variables that have a negative relationship with the model correspond to MATH_S11, CR_S11, CS_S11, NS_S11, ENG_S11, and ENG_SPRO.

On the other hand, Table 13 presents the Random Forest model's performance metrics, evaluating the model's ability to identify group membership and associated efficiency accordingly. The results show that the model can mainly identify the academic efficiency profile to which each program belongs.

Model	Accuracy				Area under ROC			
	Min	Mean	Max	sd	Min	Mean	Max	sd
RF Model	0.737	0.833	1	0.073	0.896	0.958	1	0.035

Table 11: Performance metrics for cross-validation training

Competence	Importance			
	Group 1 efficient	Group 1 Non efficient	Group 2 efficient	Group 2 Non efficient
MATH_11	-0.062	0.178	-0.011	0.070
CR_11	-0.034	0.172	-0.016	0.127
CS_11	-0.067	0.166	0.014	0.030
NS_11	-0.042	0.175	0.017	0.054
ENG_11	-0.043	0.099	0.012	0.017
QR_PRO	0.023	0.047	-0.039	0.127
CR_PRO	0.071	0.042	-0.044	0.223
CS_PRO	0.011	0.061	-0.034	0.162
ENG_PRO	-0.012	0.030	-0.031	0.135
WC_PRO	0.007	0.003	0.004	0.075
FEP_PRO	0.079	0.010	-0.019	0.068

Table 12: Importance of variables for the Random Forest classification

Metric	Group 1 efficient	Group 1 Non efficient	Group 2 efficient	Group 2 Non efficient
Sensitivity	0.250	0.960	0.250	0.829
Specificity	0.986	0.818	0.941	0.844
F1 Score	0.364	0.814	0.315	0.817

Table 13: Summary of the test process metrics

Finally, this research consists of three phases: cluster analysis, efficiency analysis, and machine learning. Exploratory cluster analysis analyzes the data by identifying clustering patterns between programs and characterizing academic profiles between university degrees. The efficiency measurement phases are performed for the profiles resulting from the cluster analysis and also for the raw data to provide a basis for comparison and contrast. The findings in Table 2 reveal the highest efficiency for non-group analysis, but this restricts the scope and complexity of the analyses that can be performed. It is also challenging to compare a university with a high level of reputation, popularity, experience, and positioning with a university for which these characteristics are low.

The efficiency analysis was carried out considering the academic program and its quality accreditations, allowing estimation of how the accreditations influence the level of efficiency in both study groups. The results highlighted how efficient Group 1 (G1_EFF) is made up of 92.86% of accredited universities, in contrast to efficient Group 2 (G2_EFF), with only 26.47% of accredited universities (see Table 5). Finally,

in this phase, it is possible to determine the weak and strong competencies of each efficient group. In addition, the score that must be increased to reach the efficiency threshold for each study unit can be established (see Table 9). This makes it possible to identify the competencies that higher education institutions must strengthen within their teaching curriculum to improve academic performance and, consequently, the level of efficiency.

In the third phase, an RF model predicts the membership of an efficiency profile (G1_EFF, G1_Not_EFF, G2_EFF, G2_Not_EFF) of the academic programs studied. This is very useful because if the university predicts a student's performance in advance, it could take steps to improve or maintain their efficiency.

DISCUSSION

It is essential to compare the results with other studies that use machine learning and artificial intelligence techniques to predict the efficient group (Group 1, Group 2, etc.) and/or the type of efficiency (efficient or not efficient). The Random

Forest model output used to classify academic efficiency levels was Area under curve ROC=0.94; this value is at the same level in another educational research. For example, the authors Durairaj and Vijitha (2014) presented a group analysis using k-means and predicting students' performance from the probabilistic algorithm and Naïve Bayes DT, obtaining a sensitivity of 0.94, a specificity of 0.47, and an F1 score of 0.93. On the other hand, de Morais, Araújo and Costa (2014) developed a methodology that consists of analyzing groups using k-means and then generating academic performance prediction through multivariate regression, obtaining a significant adjustment of 99%. Torabi, Moradi and Khantaimoori (2012) predicted the results of student evaluations using Bayesian networks, obtaining a precision of 66%. Kolo, Adepoju Solomon and Alhassan (2015) used a DT algorithm and obtained an accuracy of 66.8% in academic evaluations. Singh, Sabitha and Bansal (2016) used the k-means group analysis algorithm, showing an excellent performance of up to 80% precision in the academic evaluation results.

Consequently, Alsabawy, Cater-Steel and Soar (2016), points out that the improvement of efficiency levels is not an easy task, since there are no "automatisms" for efficiency, identifying that it is made to believe that the improvement in the educational contexts is associated with technological change. Considering the previous approach, our research uses the outputs of the DEA linear programming model, specifically the slack variables as an objective element to objectively identify potential areas for improvement in the institutions.

The comparison of institutions with similar characteristics was one of the objectives of our study, this aspect can be understood by the total range of the efficiency scores, as presented in Table 4, the minimum value of the efficiency score for Cluster 1 and 2 is 85.3% and 73.1% respectively, these values are higher than the minimum score in the global scenario without grouping, which was 68.1%. When comparing with the research (Johnes, 2006), where they analyze 130 universities in the UK using six inputs and three outputs, the minimum efficiency score was estimated at around 60%. Similarly, Klumpp (2018) in the research of 17 European universities identified a minimum efficiency score of 61.60%; The minimum threshold of 60% for the efficiency score increases in our research when institutions are grouped by similarity factors.

Kuah and Wong (2011) evaluated universities' efficiency through a DEA model. They affirm that the efficiency of a university is made up of two dimensions: teaching efficiency and research efficiency. Their research indicates an alternative to measure efficiency. However, our research's advantage is that it uses standardized tests as inputs, which are objective measures. However, one limitation of our study is that only one aspect of a university's efficiency is measured. Ramzi, Afonso and Ayadi (2016) developed an efficiency analysis of primary and secondary education in Tunisia using a DEA model, highlighting the need for clustering (cluster) and the importance of calculating the educational efficiency. By comparison, our research measures the relative impact colleges have on students when evaluating high school and college exam results. Like Agasisti, Munda and Hippe (2019), our study evaluates the university's contribution to students' professional achievement. Therefore, the proposed methodology produces good results and is relevant for the educational context when comparing our research results with similar approaches.

The results of our work become an objective tool to evaluate the academic performance of university institutions. In university management, it could be useful for independent regulatory entities as a mechanism to identify representative institutions and determine objective evaluation criteria. In the specific case of university decision-makers, the structure proposed in the research allows strategically mapping the position of a program or university in an academic context, thus supporting decision-making in investments, curricular designs, and new academic programs. Finally, students have a tool to support the career's decision to study, associating their interests with the efficiency results delivered by our efficiency analysis structure. The results indicate that Quality Accreditations support higher academic efficiency for engineering programs. However, the cluster analysis isolates the quality accreditation effect, evidencing that quality accreditations have a greater impact on universities that receive students with better abilities from high school.

CONCLUSIONS

This study comprehensively evaluated the educational efficiency of 265 academic engineering programs. A three-phase method was proposed that assesses the effect that large universities have on the sector's overall efficiency performance. This research's key contribution is the specific description of a method to evaluate and forecast academic efficiency in university education. The first phase (cluster analysis) groups universities with similar academic characteristics in clearly defined profiles. Consequently, the efficiency analysis is carried out through DEA (second phase), first without considering cluster analysis and then calculating each profile's efficiency. The evaluation of homogeneous universities makes it possible to correctly determine academic performance. Finally, the third phase corresponds to the machine learning model's application to predict an academic efficiency profile.

From the empirical evidence, the following criteria are the research findings. The first phase results show the formation of two groups: the first with high results in basic professional skills and the second group with high results in secondary basic skills. The second phase reveals that the average efficiency value for Groups 1 and 2 is 0.973 and 0.941, respectively. Finally, in the third phase, the RF model was trained and validated, which obtained a high percentage of success for predicting the academic efficiency category. A structured method for analyzing, measuring, and forecasting efficiency in engineering education is presented to the scientific community and the education sector internationally. The proposed structure enables a decision-making process for continuous improvement in educational contexts.

REFERENCES

Agasisti, T., Munda, G. and Hippe, R. (2019) 'Measuring the efficiency of European education systems by combining Data Envelopment Analysis and Multiple-Criteria Evaluation', *Journal of Productivity Analysis*, Vol. 51, No. 2, pp. 105–124. https://doi.org/10.1007/s11123-019-00549-6

Aldamak, A. and Zolfaghari, S. (2017) 'Review of efficiency ranking methods in data envelopment analysis', *Measurement*, Vol. 106, pp. 161–172. https://doi.org/10.1016/j.measurement.2017.04.028

Alsabawy, A. Y., Cater-Steel, A. and Soar, J. (2016) 'Determinants of Perceived Usefulness of e-Learning Systems', *Computers in Human Behavior*, Vol. 64, pp. 843–858. https://doi.org/10.1016/j.chb.2016.07.065

Amara, N., Rhaiem, M. and Halilem, N. (2020) 'Assessing the research efficiency of Canadian scholars in the management field: Evidence from the DEA and fsQCA', *Journal of Business Research*, Vol. 115, pp. 296–306. https://doi.org/10.1016/j.jbusres.2019.10.059

Aparicio, J., Cordero, J. M. and Ortiz, L. (2019) 'Measuring efficiency in education: The influence of imprecision and variability in data on DEA estimates', *Socio-Economic Planning Sciences*, Vol. 68, 100698. https://doi.org/10.1016/j.seps.2019.03.004

Aparicio, J., Perelman, S. and Santín, D. (2020) 'Comparing the evolution of productivity and performance gaps in education systems through DEA: an application to Latin American countries', *Operational Research: An International Journal.* https://doi.org/10.1007/s12351-020-00578-2

Benicio, J. and Mello, J. C. S. de (2015) 'Productivity Analysis and Variable Returns of Scale: DEA Efficiency Frontier Interpretation', *Procedia Computer Science*, Vol. 55, pp. 341–349. https://doi.org/10.1016/j.procs.2015.07.059

Berens, J., Schneider, K., Gortz, S., Oster, S. and Burghoff, J. (2019) 'Early Detection of Students at Risk - Predicting Student Dropouts Using Administrative Student Data from German Universities and Machine Learning Methods', *Journal of Educational Data Mining*, Vol. 11, No. 3, pp. 1–41. https://doi.org/10.5281/zenodo.3594771

Bianchi, N. and Giorcelli, M. (2020) 'Scientific Education and Innovation: From Technical Diplomas to University Stem Degrees', *Journal of the European Economic Association,* Vol. 18, No. 5, pp. 2608-2646. https://doi.org/10.1093/jeea/jvz049

Breiman, L. (2001) 'Random Forests', *Machine Learning*, Vol. 45, No. 1, pp. 5-32. https://doi.org/10.1023/A:1010933404324

Charnes, A., Cooper, W. W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Clayman, C. L., Srinivasan, S. M. and Sangwan, R. S. (2020) 'K-means Clustering and Principal Components Analysis of Microarray Data of L1000 Landmark Genes', *Procedia Computer Science*, Vol. 168, pp. 97-104. https://doi.org/10.1016/j.procs.2020.02.265

Corlu, M. A. and Aydin, E. (2016) 'Evaluation of Learning Gains Through Integrated STEM Projects', *International Journal of Education in Mathematic*s, *Science and Technology*, Vol. 4, No. 1, pp. 20–29. https://doi.org/10.18404/ijemst.35021

de Morais, A. M., Araújo, J. M. F. R. and Costa, E. B. (2014) 'Monitoring student performance using data clustering and predictive modelling', *In 2014 IEEE Frontiers in Education Conference* (FIE) Proceedings, pp. 1–8. https://doi.org/10.1109/FIE.2014.7044401

Delahoz-Dominguez, E., Zuluaga, R. and Fontalvo-Herrera, T. (2020) 'Dataset of academic performance evolution for engineering students', *Data in Brief*, Vol. 30, 105537. https://doi.org/10.1016/j.dib.2020.105537

Duan, S. X. (2019) 'Measuring university efficiency: An application of data envelopment analysis and strategic group analysis to Australian universities', *Benchmarking: An International Journal*, Vol. 26, No. 4, pp. 1161–1173. https://doi.org/10.1108/BIJ-10-2017-0274

Durairaj, M. and Vijitha, C. (2014) 'Educational data mining for prediction of student performance using clustering algorithms', *International Journal of Computer Science and Information Technologies*, Vol. 5, No. 4, pp. 5987–5991.

Ferro, G. and D'Elia, V. (2020) 'Higher Education Efficiency Frontier Analysis: A Review of Variables to Consider', Vol. 13, No. 3, 140-153. https://doi.org/10.7160/eriesj.2020.130304

Fuentes, R., Fuster, B. and Lillo-Bañuls, A. (2016) 'A three-stage DEA model to evaluate learning-teaching technical efficiency: Key performance indicators and contextual variables', *Expert Systems with Applications*, Vol. 48, pp. 89–99. https://doi.org/10.1016/j.eswa.2015.11.022

Ghasemi, N., Najafi, E., Hoseinzadeh Lotfi, F. and Movahedi Sobhani, F. (2020) 'Assessing the performance of organizations with the hierarchical structure using data envelopment analysis: An efficiency analysis of Farhangian University', *Measurement*, Vol. 156, 107609. https://doi.org/10.1016/j.measurement.2020.107609

Gralka, S., Wohlrabe, K. and Bornmann, L. (2019) 'How to measure research efficiency in higher education? Research grants vs. publication output', *Journal of Higher Education Policy and Management*, Vol. 41, No. 3, pp. 322–341. https://doi.org/10.1080/1360080X.2019.1588492

Granadillo, E. D. L. H., Gomez, J. M. and Herrera, T. J. F. (2019) 'Methodology with multivariate calculation to define and evaluate financial productivity profiles of the chemical sector in Colombia', *International Journal of Productivity and Quality Management*, Vol. 27, No. 2, pp. 144–160. https://doi.org/10.1504/IJPQM.2019.100141

Heidrich, L., Victória Barbosa, J. L., Cambruzzi, W., Rigo, S. J., Martins, M. G. and dos Santos, R. B. S. (2018) 'Diagnosis of learner dropout based on learning styles for online distance learning', *Telematics and Informatics*, Vol. 35, No. 6, pp. 1593–1606. https://doi.org/10.1016/j.tele.2018.04.007

Hoeg, D. G. and Bencze, J. L. (2017) 'Values Underpinning STEM Education in the USA: An Analysis of the Next Generation Science Standards', *Science Education*, Vol. 101, No. 2, pp. 278–301. https://doi.org/10.1002/sce.21260

ICFES (2020) *Estructura general del examen Icfes Saber Pro - Portal Icfes*, [Online], Available at: https://www.icfes.gov.co/web/guest/acerca-examen-saber-11 [26 Oct 2020].

Johnes, J. (2006) 'Data envelopment analysis and its application to the measurement of efficiency in higher education', *Economics of Education Review*, Vol. 25, No. 3, pp. 273–288. https://doi.org/10.1016/j.econedurev.2005.02.005

Khan, S. U., Khan, Z. and Hameed, G. (2019) 'Efficiency Assessment of Public Education & Health Sector in Selected Middle-Income Countries with Special Reference to Millennium Development Goals (MDGs)', *Journal of Applied Economics and Business Studies*, Vol. 3, No. 1, pp. 41–60. https://doi.org/10.34260/jaebs.313

ERIES Journal
volume 14 issue 1

Electronic ISSN
1803-1617

Printed ISSN
2336-2375

51

Klumpp, M. (2018) 'The Index Number Problem with DEA: Insights from European University Efficiency Data', *Education Sciences*, Vol. 8, No. 2, pp. 79. https://doi.org/10.3390/educsci8020079

Kolo, K. D., Adepoju Solomon A. and Alhassan, J. K. (2015) 'A Decision Tree Approach for Predicting Students Academic Performance', *International Journal of Education and Management Engineering*, Vol. 5, No. 5, pp. 12-19. https://doi.org/10.5815/ijeme.2015.05.02

Kuah, C. T. and Wong, K. Y. (2011) 'Efficiency assessment of universities through data envelopment analysis', *Procedia Computer Science*, Vol. 3, pp. 499–506. https://doi.org/10.1016/j.procs.2010.12.084

González, L., Minguillón, J., Martínez, T. and Meneses, J. (2018) 'Institutional support to provide freshmen with flexible learning paths at course and semester level in open higher education', *Working Paper UOC*.

Long, P. and Siemens, G. (2014) 'Penetrating the fog: analytics in learning and education', *Italian Journal of Educational Technology*, Vol. 22, No. 3, pp. 132–137.

Menardi, G. (2011) 'Density-based Silhouette diagnostics for clustering methods', *Statistics and Computing*, Vol. 21, No. 3, pp. 295–308. https://doi.org/10.1007/s11222-010-9169-0

Mikušová, P. (2017) 'Measuring the efficiency of the Czech public higher education institutions: an application of DEA', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 10, No. 2, pp. 58–63. https://doi.org/10.7160/eriesj.2017.100204

Najadat, H., Althebyan, Q. and Al-Omary, Y. (2019) 'Higher Education Units Assessment Based on Data Envelopment Analysis and Clustering Techniques', *In 2019 International Arab Conference on Information Technology* (ACIT), pp. 171–177. https://doi.org/10.1109/ACIT47987.2019.8991103

Nazir, A. (2019) 'A critique of imbalanced data learning approaches for big data analytics', *International Journal of Business Intelligence and Data Mining*, Vol. 14, No. 4, pp. 419–457. https://doi.org/10.1504/IJBIDM.2019.099961

Oyelade, J., Isewon, I., Oladipupo, O., Emebo, O., Omogbadegun, Z., Aromolaran, O., Uwoghiren, E, Olaniyan, D., Olawole, O. (2019) 'Data Clustering: Algorithms and Its Applications', *In 2019 19th International Conference on Computational Science and Its Applications* (ICCSA), pp. 71–81. https://doi.org/10.1109/ICCSA.2019.000-1

Pawsey, N., Ananda, J. and Hoque, Z. (2018) 'Rationality, accounting and benchmarking water businesses: An analysis of measurement challenges', *International Journal of Public Sector Management*, Vol. 31, No. 3, pp. 290–315. https://doi.org/10.1108/IJPSM-04-2017-0124

Ramzi, S., Afonso, A. and Ayadi, M. (2016) 'Assessment of efficiency in basic and secondary education in Tunisia: A regional analysis', *International Journal of Educational Development*, Vol. 51, pp. 62–76. https://doi.org/10.1016/j.ijedudev.2016.08.003

Sharma, J. and Yarlagadda, P. K. (2018) 'Perspectives of 'STEM education and policies' for the development of a skilled workforce in Australia and India', *International Journal of Science Education*, Vol. 40, No. 16, pp. 1999–2022. https://doi.org/10.1080/09500693.2018.1517239

Singh, I., Sabitha, A. S. and Bansal, A. (2016) 'Student performance analysis using clustering algorithm', *In 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence)*, pp. 294–299. https://doi.org/10.1109/CONFLUENCE.2016.7508131

Sinuany-Stern, Z., Mehrez, A. and Hadad, Y. (2000) 'An AHP/DEA methodology for ranking decision making units', *International Transactions in Operational Research*, Vol. 7, No. 2, pp. 109–124. https://doi.org/10.1111/j.1475-3995.2000.tb00189.x

Sun, Q., Liu, H., & Wang, W. (2020). 'Evaluating the Quality of GANs Based on F1 Harmonic Mean', *Procedia Computer Science*, Vol. 166, pp. 463–468. https://doi:10.1016/j.procs.2020.02.064

Suresh, K. M., Asokan, P. and Vinodh, S. (2016) 'Application of design for Six Sigma methodology to an automotive component', *International Journal of Six Sigma and Competitive Advantage,* Vol. 10, No. 1, pp. 1. https://doi.org/10.1504/IJSSCA.2016.080446

Suter, L. E. and Camilli, G. (2019) 'International Student Achievement Comparisons and US STEM Workforce Development', *Journal of Science Education and Technology*, Vol. 28, No. 1, pp. 52–61. https://doi.org/10.1007/s10956-018-9746-0

Thanassoulis, E., Dey, P. K., Petridis, K., Goniadis, I. and Georgiou, A. C. (2017) 'Evaluating higher education teaching performance using combined analytic hierarchy process and data envelopment analysis', *Journal of the Operational Research Society*, Vol. 68, No. 4, pp. 431–445. https://doi.org/10.1057/s41274-016-0165-4

Torabi, R., Moradi, P. and Khantaimoori, A. R. (2012) 'Predict Student Scores Using Bayesian Networks', *Procedia - Social and Behavioral Sciences*, Vol. 46, pp. 4476–4480. https://doi.org/10.1016/j.sbspro.2012.06.280

Torres-Samuel, M., Vásquez, C. L., Luna, M., Bucci, N., Viloria, A., Crissien, T. and Manosalva, J. (2020) 'Performance of Education and Research in Latin American Countries through Data Envelopment Analysis (DEA)', *Procedia Computer Science*, Vol. 170, pp. 1023–1028. https://doi.org/10.1016/j.procs.2020.03.079

Tran, C.-D. T. T. and Villano, R. A. (2019) 'Environmentally-adjusted efficiencies of Vietnamese higher education institutions: a multi-stage bootstrap DEA method', *International Journal of Operational Research,* Vol. 36, No. 3, pp. 413–439. https://doi.org/10.1504/IJOR.2019.103126

Visbal-Cadavid, D., Martínez-Gómez, M. and Escorcia-Caballero, R. (2020) 'Exploring University Performance through Multiple Factor Analysis: A Case Study', *Sustainability*, Vol. 12, No. 3, 924. https://doi.org/10.3390/su12030924

Visbal-Cadavid, D., Martínez-Gómez, M. and Guijarro, F. (2017) ,Assessing the efficiency of public universities through DEA. A case study', *Sustainability*, Vol. 9, No. 8, 1416. https://doi.org/10.3390/su9081416

Visbal-Cadavid, D., Mendoza, A. M. and Hoyos, I. Q. (2019) 'Prediction of efficiency in Colombian Higher Education institutions with Data Envelopment Analysis and Neural Networks', *Pesquisa Operacional,* Vol. 39, No. 2, pp. 261–275. https://doi.org/10.1590/0101-7438.2019.039.02.0261

Witte, K. D. and López-Torres, L. (2017) 'Efficiency in education: a review of literature and a way forward', *Journal of the Operational Research Society*, Vol. 68, No. 4, pp. 339–363. https://doi.org/10.1057/jors.2015.92

Wolszczak-Derlacz, J. (2017) 'An evaluation and explanation of (in)efficiency in higher education institutions in Europe and the U.S. with the application of two-stage semi-parametric DEA', *Research Policy,* Vol. 46, No. 9, pp. 1595–1605. https://doi.org/10.1016/j.respol.2017.07.010