Journal of Quality Measurement and Analysis *JQMA* 21(4) 2025, 167-194
e-ISSN: 2600-8602 *http://www.ukm.my/jqma*
https://doi.org/10.17576/jqma.2104.2025.10

COMPARING AN INTEGRATED DATA ENVELOPMENT ANALYSIS AND MACHINE LEARNING MODELS FOR ACCURATE ACADEMIC EFFICIENCY PREDICTION

(*Perbandingan Model Integrasi Analisis Kesimpulan Data dan Pembelajaran Mesin untuk Ramalan Kecekapan Akademik yang Tepat*)

NOR FAEZAH MOHAMAD RAZI*, NORHAYATI BAHARUN & SURAYA MASROM

ABSTRACT

The integration of data envelopment analysis (DEA) with machine learning (ML) offers a novel approach to evaluating academic efficiency beyond traditional measures like CGPA. This study develops an efficiency assessment framework combining DEA and ML to predict student academic achievement efficiency. The objectives are (1) to identify and validate input and output variables for DEA-based academic efficiency measurement and (2) to develop an integrated predictive model using DEA and ML for improved accuracy. A cross-sectional study was conducted on 1,099 final-year diploma students, collecting data on CGPA, satisfaction, and five competency domains (personal, adaptive, digital, social, and 21st-century skills). Efficiency scores were computed using the BCC and CCR DEA models, followed by ML predictions using random forest (RF), gradient boosting regressor (GBR), artificial neural networks (ANN), and AutoML via genetic programming. Performance was evaluated using RMSE, MAE, and R^2 metrics. The findings indicate that the DEA-GBR model achieved the highest predictive accuracy (RMSE = 0.0101, MAE = 0.0039, R^2 = 0.9889), outperforming other models. SHAP analysis identified digital competency as the most influential predictor, aligning with UiTM's digital transformation goals. The integration of DEA with ML significantly improved discriminatory power, reducing the number of efficient decision-making units (DMUs) from 134 to as low as 44. This study enhances academic efficiency assessment by integrating DEA with predictive ML models, providing a data-driven approach for student performance evaluation. Future research should expand datasets and explore additional ML techniques for further refinement.

Keywords: data envelopment analysis (DEA); machine learning (ML); measurement efficiency

ABSTRAK

Integrasi analisis penyampulan data (DEA) dengan pembelajaran mesin (ML) menawarkan pendekatan baru untuk menilai kecekapan akademik melangkaui ukuran tradisional seperti CGPA. Kajian ini membangunkan rangka kerja penilaian kecekapan yang menggabungkan DEA dan ML untuk meramal kecekapan pencapaian akademik pelajar. Justeru itu, objektif kajian ini adalah (1) untuk mengenal pasti dan mengesahkan pembolehubah input dan output untuk pengukuran kecekapan akademik berasaskan DEA dan (2) untuk membangunkan model ramalan bersepadu menggunakan DEA dan ML untuk ketepatan yang lebih baik. Kajian keratan rentas telah dijalankan ke atas 1,099 pelajar diploma tahun akhir, mengumpul data mengenai CGPA, kepuasan, dan lima domain kompetensi (peribadi, penyesuaian, digital, sosial, dan kemahiran abad ke-21). Markah kecekapan dikira menggunakan model BCC dan CCR DEA, diikuti dengan ramalan ML menggunakan hutan rawak (RF), regresor penggalak kecerunan (GBR), rangkaian neural buatan (ANN) dan AutoML melalui pengaturcaraan genetik. Prestasi dinilai menggunakan metrik RMSE, MAE dan R^2. Dapatan menunjukkan bahawa model DEA-GBR mencapai ketepatan ramalan tertinggi (RMSE = 0.0101, MAE = 0.0039, R^2 = 0.9889), mengatasi prestasi model lain. Analisis SHAP mengenal pasti kompetensi digital sebagai peramal yang paling berpengaruh, sejajar dengan matlamat transformasi digital UiTM. Penyepaduan DEA dengan ML telah meningkatkan kuasa diskriminasi dengan ketara, mengurangkan bilangan unit membuat keputusan (DMU) yang cekap daripada 134 kepada serendah 44. Kajian ini meningkatkan penilaian kecekapan akademik dengan menyepadukan DEA dengan model ML ramalan, menyediakan pendekatan dipacu data untuk penilaian prestasi pelajar. Penyelidikan masa depan harus mengembangkan set data dan meneroka teknik ML tambahan untuk pemurnian selanjutnya.

Kata kunci: analisis penyampulan data; pembelajaran mesin; kecemerlangan pelajar

1. Introduction

The increasing demand for evidence-based approaches in education has spurred interest in methods that can evaluate and enhance academic achievement. Academic achievement traditionally encompasses the knowledge, skills, and behaviors that students acquire in educational settings, often measured by quantifiable outcomes like Cumulative Grade Performance Average (CGPA). Historically, higher education success was primarily gauged through final exam performances; however, this measure is increasingly seen as inadequate for capturing the full spectrum of student competencies. Modern educational discourse now advocates for integrating broader competencies such as critical thinking, creativity, problem-solving abilities, and emotional intelligence into evaluations, recognizing these skills as essential for real-world readiness (Camanho *et al*. 2021; Yavuzalp & Bahcivan 2021). However, traditional techniques often fail to provide accurate or actionable insights into efficiency due to their inability to account for complex, multidimensional data. The integration of these competencies into academic assessments is vital for developing well-rounded individuals who can meet contemporary employers' expectations. By understanding graduates' comprehensive capabilities, such as problem-solving skills and adaptability, educators and administrators can tailor educational resources and support to better align academic outcomes with job market demands. This holistic evaluation approach ensures that higher education not only enhances the quality of its graduates but also prepares them effectively to meet professional challenges and fulfill employer expectations (Behle 2020; Datnow *et al*. 2022).

Forming a good academic achievement holistically means there is a need to evaluate all the resources that have been provided to the student and how well the student utilized it. Therefore, it involves a very complex measurement. Dealing with this issue, the non-parametric method (data envelopment analysis) has a better dispersion of results than the parametric method in measuring efficiency (Farantos 2015). Data envelopment analysis has been broadly used to evaluate efficiency in many areas such as financial institutions (de Abreu & Kimura 2020; Ebrahimi & Hajizadeh 2021; Tsolas *et al*. 2020), farming (Nandy & Singh 2021), hospitals (Cinaroglu 2021; Misiunas *et al*. 2016), airlines (Alcaraz *et al*. 2021; Özsoy & Örkcü 2021), and government agencies (Zhang & Shi 2019). Even though several research works have provided insights into the richness of DEA applications, many aspects of efficiency still need to be explored. Particularly in education field, previous literature measures achievement efficiency lacking in determining academic achievement. Most studies related to application of DEA in education field were to measure the performance of in the schools (Camanho *et al*. 2021; Esteve *et al*. 2020) and universities (Zhang & Shi 2019). It is due to the changes in the education landscape measuring student academic achievement will help higher institution management and educators evaluate the resources provided to students during their learning process and, in turn, can improve the quality of academic achievement.

DEA not only assesses the efficiency of educational entities by determining how effectively individuals or institutions utilize their inputs to produce outputs, but it also identifies inputs that are not being optimally used (Shero *et al*. 2022). This ability to pinpoint inefficiencies helps in refining educational strategies and resource allocation. Widely adopted across educational research, DEA has been instrumental in evaluating the performance of schools and universities, providing insights that guide improvements in educational practices. Despite the widespread use of DEA in educational efficiency analysis, its predictive potential remains underutilized. Similarly, while ML models are employed for predictive tasks, they lack the capacity to provide efficiency-specific insights. This gap highlights the need for a framework that combines DEA's evaluation strengths with ML's predictive power, enabling more nuanced and accurate analysis. This study addresses these limitations by integrating data envelopment analysis (DEA), a well-established efficiency evaluation tool, with machine learning (ML), renowned for its predictive capabilities.

Thus, this study aims to develop a full framework implementation of DEA model with ML approaches to predict the student's academic achievement efficiency. Therefore, in the objectives of the study, the following are the objectives that need to be achieved:

1. To evaluate the selection of input and output variables for measuring the academic achievement efficiency of final year diploma students using the DEA model.
2. To develop predictive models for the academic achievement efficiency score based on integrated DEA and ML approaches.

As new research on predicting efficiency score of students' academic achievement for DEA model using ML approaches will create a new research question that requires extensive empirical research work. All the research works need to be designed appropriately, considering many aspects based on the selection input and output variables and the ML configurations. The empirical works should observe how the ML algorithm can influence the efficiency score. Additionally, it is essential to look at the selection of input and output variables and its efficiency score to ML accuracy. It is also vital to observe ML performances with the other kind of ML algorithms. Thus, this research will fill the research gap on the development of predictive model using integrated method using DEA and ML approaches.

2. Literature Review

2.1. *An overview of efficiency analysis in education*

The primary objective of the literature review is to elucidate the development of Performance Evaluation method in measuring academic achievement in education area and prior research findings pertinent to this study, with a particular emphasis on data envelopment analysis (DEA). Efficiency in education occurs at a time when the output can be test results or added value produced at a minimum level or resources, such as finance or the natural ability of students (Johnes 2015). This refers to achieving maximum results (output) using minimum effort (input) in limited time. Whereas according to Ghaffarian Asl and Osam (2021), effectiveness can be seen as compatibility between the output that is the main goal and other criteria in relation to efficiency. This means that consider effectiveness and efficiency as two dimensions of institutional performance (Lindsay 1982). It can be said that when an organization or an individual has high efficiency, it will always increase the effectiveness of the achievement of that achievement. Table 1 provides a comprehensive review of various studies that assess the efficiency of educational institutions across different countries, employing a range of methodologies and efficiency models.

Numerous methodologies for efficiency analysis have been developed (Table 1), demonstrating valuable progress in supporting decision-makers in making informed choices. Pure technical in efficiency model dominantly its study globally such as in Spanish, Contreras & Lozano (2022) had analyzed the Spanish public university system in order to maximize its efficiency involving 84 universities. Studied in Germany by Zarrin (2021) and Gralka *et al.* (2019) involving 28 university hospitals in Germany. While Camanho *et al.* (2021) investigates the relationship between students' performance and the type of school attended during upper secondary education. Overall, looking at the assessment column, it was found that the focus of existing studies largely centers on teaching and research. This leaves other critical educational outputs, such as student employability, satisfaction, and post-graduation success, relatively underexplored. Expanding the focus on these areas could provide a broader understanding of educational outcomes. Lastly, the research predominantly relies on traditional DEA models. Exploring newer or unconventional efficiency models could yield fresh insights, particularly in managing complex educational environment.

Table 1. A listing of recent studies on university efficiency analysis

Research	Country	Methodology	Efficiency model	Assessment	Duration
(Contreras & Lozano 2022)	Spanish	DEA	Pure Technical	Overall	2016
(Camanho *et al*. 2021)	Italy	DEA	Malmquist index	Teaching and Research	2017-2018
(Tavares *et al*. 2021)	Brazil	Network DEA	Technical & Scale	Teaching and Learning	
(Zarrin 2021)	Germany	DEA	Pure Technical	Teaching and Learning	
(Tran *et al*. 2020)	Vietnam	Stochastic Frontier		Teaching and Learning	2013-2014
(Segovia-Gonzalez *et al*. 2020)	UK	DEA	Pure Technical	Teaching and Research	2018
(Gralka *et al*. 2019)	Germany	DEA SFA	Pure Technical	Research	2004-2013
(Yang *et al*. 2018)	China	Two-stage DEA	Not mentioned	Research	2010-2013
(Jauhar *et al*. 2017)	India	VRS	Pure Technical	Teaching and Learning	2001/02-2012/13
(Sagarra *et al*. 2017)	Mexico	VRS	Pure Technical	Teaching and Learning	2007-2012
(Munoz 2016)	Chile	CRS and VRS	Pure Technical	Research	2013–2014
(Pietrzak *et al*. 2016)	Poland	DEA	Scale	Research	2013-2014
(Aziz *et al*. 2013)	Malaysia	DEA	Technical	Overall	2011

2.2. *Data envelopment analysis*

Data envelopment analysis (DEA) is a non-parametric performance measurement tool introduced by Charnes, Cooper, and Rhodes in 1978 to evaluate the efficiency of Decision-Making Units (DMUs) such as businesses, government agencies, healthcare facilities, and educational institutions. Unlike traditional evaluation methods, DEA uses linear programming to compare the ratio of inputs (resources) to outputs (results) without requiring predefined weights or financial benchmarks (Ray 2022). The model assesses efficiency relative to a frontier, where DMUs on the frontier are efficient, while those below it are inefficient (Cooper *et al*. 2007). DEA has evolved with various models, including the CCR model (constant returns to scale) and the BCC model (variable returns to scale), making it adaptable for real-world efficiency assessments. In education, DEA is widely used to evaluate the efficiency of institutions, programs, and policies by analyzing how well resources such as faculty and funding contribute to student success (Zubir *et al*. 2024; Zubir *et al*. 2023).The ability of DEA to handle multiple inputs and outputs while accommodating complex educational structures makes it a valuable tool for assessing institutional performance (Abramo *et al*. 2018; Pokushko *et al*. 2020).

Despite its advantages, DEA has limitations, particularly its lack of predictive power and sensitivity to statistical noise, which can distort efficiency assessments (Jauhar *et al*. 2023; Zhong *et al*. 2021). Additionally, DEA's computational demands increase significantly with larger datasets containing numerous inputs and outputs, limiting its scalability (Emrouznejad & Shale 2009). While DEA has traditionally been used for institutional efficiency evaluations, recent studies have applied it at the individual level, treating students as DMUs to analyze academic efficiency. Efficient score predictions from DEA are essential for benchmarking performance and guiding resource allocation. Integrating machine learning (ML) with DEA has been proposed as a solution to enhance predictive capabilities, allowing for proactive and strategic decision-making in education. By addressing DEA's predictive limitations and computational challenges, future research can further optimize its role in assessing and improving educational efficiency.

2.3. *Machine learning*

Machine learning (ML), a subset of artificial intelligence (AI), enables systems to learn from data and make predictions with minimal human intervention. Integrating ML with data envelopment analysis (DEA) enhances DEA's ability to handle complex data patterns, improve predictive accuracy, and strengthen efficiency evaluations. DEA, traditionally used to assess the efficiency of Decision-Making Units (DMUs), is limited in managing non-linear relationships and large datasets. ML techniques, known for their flexibility and pattern-recognition capabilities, complement DEA by allowing for more advanced performance assessments (Avramidou & Tjortjis 2021; Sampath Kumar *et al*. 2023). Predictive modeling, a key aspect of ML, includes regression for continuous outcomes and classification for discrete categories. Regression models have been applied in areas such as environmental impact assessment and agricultural planning, while classification models have been used in education to predict student performance and early intervention strategies (Umer *et al*. 2017). The combination of DEA and ML enables more dynamic efficiency evaluations, expanding DEA's traditional applications beyond static measurements.

Integrating ML into DEA frameworks improves model functionality, interpretability, and robustness across various domains. ML techniques enhance feature selection, enabling DEA to focus on the most influential variables while handling large multidimensional datasets. This makes DEA more applicable in complex, data-rich environments such as healthcare, finance, and education (Abramo *et al*. 2018; Pokushko *et al*. 2020). The synergy between DEA and ML bridges gaps in traditional DEA models, increasing predictive power and supporting more data-driven decision-making. Empirical studies confirm the feasibility and effectiveness of this integration, demonstrating its ability to enhance performance evaluation and benchmarking (Zhong *et al*. 2021; Jauhar *et al*. 2023). This evolving integration offers valuable insights for further exploration, highlighting the potential of hybrid DEA-ML models in various industries. Table 2 provides a summary of key findings on the integration of DEA with ML techniques.

By combining the literature of DEA methods and machine learning integration method, we can notice that some of studies reflect the evolving integration of data envelopment analysis (DEA) with various ML techniques to enhance performance evaluation, efficiency measurement, and predictive accuracy across diverse fields. They have achieved good results through empirical studies and verified the reasonableness and feasibility of the integrated model. of DEA and ML. This provides valuable experience for further in-depth exploration.

Table 2: Previous findings on integrating DEA with machine learning approaches

References	Integrating Approach	Application
(De La Hoz *et al*. 2021)	DEA Clustering	To evaluate and forecast the academic efficiency of engineering programs in Colombia.
(Zhishuo Zhang *et al*. 2022)	DEA-SBM 11 ML Algorithm	Proposes a performance prediction method
(Singpai & Wu 2020)	DEA AutoML	To assess and predict performance in SDGs
(Zhong *et al*. 2021)	SE-SBM 15 ML algorithm	To construct the regression model.
(Jomthanachai *et al*. 2021)	DEA ML	To predict and determine the risk level based on the efficiency of DMUs.
(Liu *et al*. 2021)	DEA Regression	This study is to forecast annual fishery capacity.
(Xu *et al*. 2021)	DEA ML (CART, BT, RF, LR)	To predict the U.S COVID-19 response performance.
(Zhu *et al*. 2021)	DEA ML	To predict the DEA efficiency of DMUs.
(Akhavan Kharazian *et al*. 2019)	DEA CART	Determines the efficiency of individuals
(Tsaples *et al*. 2022)	DEA CART	To explore country sustainability composite indices under different perceptions and assumptions.

2.4. *Identified research gaps*

Gap 1: Subjectivity in input and output selection for DEA analysis

Through comprehensive literature exploration, it has been found that input and output selection in data envelopment analysis (DEA) is a critical step that significantly influences the scope and conclusions of efficiency analysis. In the context of education, outputs typically represent outcomes from various educational levels (e.g., primary, secondary, post-secondary), while inputs are the mechanisms by which decision-making units (DMUs) achieve these outcomes. Many research highlighted the importance of the selection of inputs and outputs in DEA, however, they found that they were inherently subjective. Variables such as Human Resources, Facilities, Financials, Equipment, Curriculum, Student Characteristics, and Community Resources are commonly categorized as input themes, while Student Achievement, Graduation Rates, Employment Outcomes, and Research Outcomes serve as output themes. This diversity reflects the wide-ranging factors influencing educational efficiency. Yet, the absence of universally accepted guidelines for defining these variables introduces inconsistency, limiting the comparability and validity of results. Misaligned or irrelevant variable selection can result in inaccurate evaluations of DMU performance, undermining the reliability of DEA analyses in education. The systematic studies conducted by Mohamad Razi *et al*. (2024) compiled seven input themes and four output themes, it is evident that the selection of input and output variables in efficiency studies lacks a standardized rule or benchmarking framework. This situation highlights a notable gap: an overemphasis on traditional inputs, such as human resources, financial resources, and student characteristics, while limited attention is given to broader, more comprehensive competencies. Based on current research, there were no studies that use student achievement and student satisfaction as input and student competencies (personal, adaptive, social, digital and 21st competencies) as output to be included in the potential selection input and output to measure student academic efficiency. This would be a significant literature gap that this study could fulfill. By introducing new input and output into the selection of input and output in measuring efficiency in education can lead to varying efficiency scores (Ahn *et al*. 2022) and allows more nuance understanding of efficiency (Mosbah *et al*. 2020).

Gap 2: Limitations in DEA's predictive capabilities and efficiency evaluation

Another significant methodological gap is the evaluation of efficiency scores through the integration of DEA with ML approaches, including conventional algorithms, artificial neural networks (ANN), and automated machine learning (AutoML) frameworks. While DEA is widely recognized for assessing efficiency, its reliance on deterministic models often limits its adaptability to dynamic and complex datasets. For instance, when a new DMU requires an efficiency score, the DEA analysis must be re-conducted, as noted by (Anouze & Bou-Hamad 2019; Zhu *et al*. 2021). To address these limitations, developing a predicted model for efficiency scores becomes essential. Integrating DEA with ML approaches mitigates many of DEA's inherent limitations, such as the inability to handle large, complex datasets and its lack of predictive capabilities. Moreover, ML integration enhances the discrimination capability (Jomthanachai *et al*. 2021) among decision-making units (DMUs), enabling more robust efficiency evaluations. Recent studies (Kannan *et al*. 2024; Khoubrane *et al*. 2024) have further demonstrated the effectiveness of hybrid DEA-ML models reinforcing the potential of these integrated approaches to provide scalable, predictive, and interpretable efficiency evaluation tools in educational and institutional settings.

These two research gaps (1) the subjectivity and lack of standardized guidelines for input-output selection in DEA and (2) the limitations of traditional DEA in handling dynamic datasets and predicting efficiency scores underscore the need for innovative approaches. Developing integrated predictive models that combine DEA with advanced ML techniques offers a pathway to address these challenges, streamlining efficiency evaluations while enhancing accuracy and reliability in educational contexts.

3. Methodology

3.1. *Research design*

This research design aims to develop a comprehensive framework for predicting students' academic achievement efficiency by integrating data envelopment analysis (DEA) and machine learning (ML) approaches as presented in Figure 5. The data analysis is structured into two main phases: first, efficiency scores are calculated using DEA with constant returns to scale (CCR/CRS) and variable returns to scale (BCC/VRS) models, and second, predictive models are developed using machine learning algorithms. By meeting its objectives, the study aims to enhance the accuracy and reliability of efficiency assessments, offering meaningful insights to improve educational outcomes. Key research objectives include evaluating the selection of input and output variables for measuring students' academic achievement efficiency, optimizing super-efficiency models within DEA to accurately identify the most efficient decision-making units (DMUs), and developing predictive models for academic achievement efficiency scores through the integration of DEA and ML approaches.

3.2. *Population and sample*

The target population for this study comprises all the final year diploma students who enroll in diploma level Universiti Teknologi Mara. A total of 24,074 final year diploma student been registered (Data Analytics and Statistics Unit, UiTM) in 15 campuses with various diploma programs in UiTM all around Malaysia. These campuses and branches been divided into five main regions. The Northern region had four main campuses. Central region had two campuses.

East Cost region, Southern region and East region had 3 campuses,4 campuses and 2 campuses respectively. Due to limited circumstances and time, this study decided to take a sample instead of the whole population. Since it involves a very large geographical region, multistage cluster sampling has been chosen for this study. The procedure involved in multistage cluster sampling is not very different from that in random sampling.

3.3. *Research instrument and data gathering*

Building upon our earlier publication on instrument development (Ahmad *et al*. 2023), this study utilizes the same validated tool, the Student's Competency Questionnaire (SCQ), to explore the integration of DEA and ML approaches. The SCQ consists of four parts: Part A (student demographics), Part B (academic background), Part C (student competencies based on Education 5.0@UiTM), and Part D (student satisfaction). The competencies in Part C are measured using a 10-point Likert scale ranging from Not Competent at All to Very Competent, providing greater granularity compared to the traditional 7-point scale. This choice aligns with recommendations by Awang *et al*. (2016), who found 10-point scales more effective for measurement models. Informed consent was obtained from all participants, and they were assured of their right to withdraw at any point without consequence.

3.4. *Decision making units and input and output selection*

The selection of decision-making units (DMUs) is a crucial criterion in measuring relative efficiency. In this study, DMUs are defined as individuals, specifically the final year diploma students. A general rule of thumb, as stated by Raab and Lichty (2002) and Khezrimotlagh (2015) suggests that the number of DMUs should be greater than or equal to three times the sum of the inputs and outputs. Given that this study involves two inputs and five outputs, the minimum number of DMUs required is (2+5)*3=21. With a total of 1,282 final year diploma students initially considered, 183 samples were excluded due to incomplete information, leaving 1,099 decision-making units (DMUs) for this study. This sample size is deemed more than sufficient to proceed with the analysis. This study selected two inputs: student results (CGPA) and student satisfaction, and five output: five types of competencies (Personal, Adaptive, Digital, Century, and Social). The selection of these variables is based on the university's efficiency in terms of developing students' competencies, with the aim of identifying the skills students possess upon graduation and whether they correlate with their academic achievements

Input variables involved the Student Result determined by Cumulative Grade Point Average (CGPA). The CGPA is calculated by taking the weighted average of the grade points earned in all completed courses, where the weight of each course is determined by its credit hours. Next for the Satisfaction, it measure the level of contentment or satisfaction experienced by students during the diploma study. 16 satisfaction items consisting the to measure overall final year diploma student satisfaction during their studies. Meanwhile for output variables consisting five outputs variables which are Personal, Adaptive, Digital, Century, and Social competencies. Personal competency refers to the combination of skills, behaviors, and attitudes that enable an individual to navigate their personal and professional lives effectively. In an educational context, developing personal competency involves helping students cultivate a strong sense of self, resilience, and the ability to manage their emotions and relationships. This study involved nine item to measure final year diploma student with are self-control, trustworthy, Conscientiousness, Adaptability, Innovativeness, Achievement drive, Commitment, Initiative, Optimism.

3.5. S*tudent achievement efficiency based on different input and output selection*

Four models have been developed based on different selection inputs and outputs. Figure 1 illustrates a structured approach to variable selection across different model variants, all designed to evaluate how educational inputs (CGPA and satisfaction) can be optimized to enhance various student competencies. The selection of specific groups of variables for each model is based on thematic groupings that reflect distinct dimensions of student competencies.

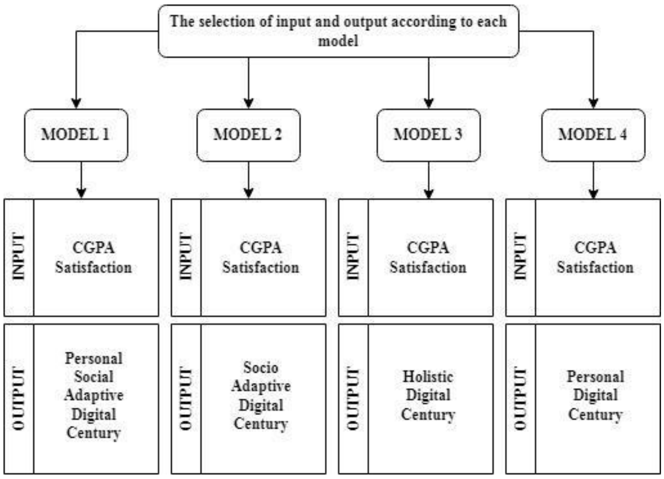

Figure 1: Input and output selection framework for DEA models evaluating student efficiency

3.6. *Development of an integrated DEA and ML predictive model*

Phase 1: Data preparation and preprocessing

Data preprocessing is performed on the student achievement dataset. In this phase, two main procedures have been conducted in order to produce an accurate result of efficiency score. The first is outliers. This study identifies the present of outliers using Mahanalobis Distance technique. A threshold of 18 for the Mahalanobis Distance was established as the cutoff for identifying and omitting outliers from the dataset. This threshold is well-supported in the literature, including studies by Barnett and Lewis (1994) and Leys *et al*. (2018), and is commonly used in practice. Therefore, in this study, Mahalanobis Distance values greater than 18 were considered as outliers and were excluded from further analysis. While the second issue was missing value. Dealing with self-reported data might face a missing value, ultimately will lead to biased result, thus, to overcome this, descriptive statistics been used to identify the extent and pattern of missingness, determining whether it is random or systematic. This approach, supported by studies like Ochieng' Odhiambo (2020), helps maintain the dataset's size and integrity while minimizing distribution distortion.

Phase 2: The development of DEA model

A radial model of CCR (CRS) and BCC (VRS) were used to measure the efficiency score. This study preferred output orientation model since the major focus on enhancing students' competencies (Adna *et al*. 2025; Toloo *et al*. 2021). The CCR Model (Charnes, Cooper, and Rhodes) model assumes constant returns to scale, meaning that an increase in inputs will lead to a proportional increase in outputs. By using the CRS model, the researchers could assess the overall efficiency of the programs, taking into account both technical and scale efficiency. In this context of study, using CRS model, the efficiency of each student in transforming their input (CGPA and Satisfaction) into output (personal competencies, adaptive competencies, digital competencies, 21st century competencies, social competency). The BCC(VRS) and CCR(CRS) model have been formulated as in Eq. (1) and Eq. (2) respectively.

$$
\begin{aligned}
&\text{Maximize } \theta_0 \\
&\text{subject to constraint:} \\
&\sum\nolimits_{j=1}^{n} \lambda_j x_{ij} \le x_{i0,(\forall_i)} \\
&\sum\nolimits_{j=1}^{n} \lambda_j y_{rj} \le \theta_0 y_{r0,(\forall_r)} \\
&\sum\nolimits_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0, \forall_j
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
&\text{Maximize } \theta_0 \\
&\text{subject to constraint:} \\
&\sum\nolimits_{j=1}^{n} \lambda_j x_{ij} \le x_{i0,(\forall_i)} \\
&\sum\nolimits_{j=1}^{n} \lambda_j y_{rj} \ge \theta_0 y_{r0,(\forall_r)} \\
&\lambda_j x \ge 0, \forall_j
\end{aligned}
\tag{2}
$$

where θ_0 is the efficiency score of the DMU being evaluated for both model (BCC and CCR); x_{ij} is the input I for DMU_j; y_{rj} is the output *r* for DMU_j; λ_j is the intensity variable (weights) associated with each DMU; $\sum\nolimits_{j=1}^{n} \lambda_j x_{ij} \le x_{i0,(\forall_i)}$ is the value of the *i*-th input for the DMU under evaluation, denoted by 0; $\sum\nolimits_{j=1}^{n} \lambda_j y_{rj} \le \theta_0 y_{r0,(\forall_r)}$ is the multiplication from x_j input and output values; *n* is the number of DMUs (1099) and $\sum\nolimits_{j=1}^{n} \lambda_j = 1$ is the convexity constraint which allows for variable return to scale (referring to BCC model).

Phase 3: The development of integrated predictive model using DEA+ML approaches

In this phase, the study presents the design and implementation of a predictive model to estimate efficiency scores using three machine learning approaches: conventional machine learning, artifical neural network (ANN) and metaheuristic AutoML genetic programming based. The conventional ML algorithm will serve as a baseline, utilizing traditional techniques such as regression, decision trees, and gradient boosting. On the other hand, metaheuristic AutoML will employ advanced optimization strategies to automatically select the best model and hyperparameters, aiming to enhance prediction accuracy and computational efficiency. Before proceed to develop the predictive model, certain keys of assumption need to be checked to ensure that the model is valid, reliable and interpretable. Normality, linearity, homocedasticity assumptions have been assessed. The normalized procedure also be done. Once all the assumptions have been identified, the next process is to develop the integrated predictive model using DEA and ML approaches.

a) Developing an integrated predictive model using DEA and ML approach (conventional ML regressor)

In developing predictive methods, selecting the appropriate ML algorithm is crucial to ensure the model aligns with the nature of the dataset and the study's objectives. Two primary types of ML algorithms are ML regressors and ML classifiers. The choice depends on whether the target variable is continuous or categorical.ML regressors aim to predict continuous numeric values based on input features, making them well-suited for datasets where the dependent variable can take any real number within a specific range (Sagar *et al*. 2023; Zhang *et al*. 2022). Regressor algorithms model the relationship between independent variables (features) and the continuous dependent variable (target) to provide accurate predictions. In this study, the dependent variable is the efficiency score, which is a continuous numeric value derived from DEA. Given this characteristic, ML regressor algorithms were chosen to predict efficiency scores effectively. The continuous nature of efficiency scores requires a regression approach to model the relationship between predictors (e.g., student performance metrics) and the target variable, ensuring precise and interpretable predictions. Common algorithm used by previous study was in regression model which linear regression (Ghildiyal *et al*. 2024), polynomial regression (Sagar *et al*. 2023) and multivariate regression (Zhang *et al*. 2019) and ridge regression and LASSO (Petrelli 2023).

When choosing between parametric and non-parametric algorithms for regression, the decision depends on the data's nature and the study's objectives. Parametric models assume a specific form for the relationship between variables, making them computationally efficient and easy to interpret. They work well when their assumptions hold but are limited in flexibility (Taylan 2020; Yavuz & Şahin 2022). Non-parametric models, on the other hand, do not assume a fixed functional form, allowing them to capture complex, nonlinear patterns in the data (Laksaci *et al*. 2023). They are more adaptable, particularly for high-dimensional datasets, but may require more computational resources. Given the continuous nature of the dependent variable (DEA efficiency scores), regression models were more suitable than classifiers, which are commonly found in earlier studies but are limited to categorical outcomes.

In this study, non-parametric regression approaches were prioritized due to their adaptability to high-dimensional educational data and ability to model non-linear student performance patterns. With a large sample size (n = 1099) and the goal of precise individual-level prediction, algorithms such as K-nearest neighbors (KNN), decision tree, random forest, support vector regression (SVR), and gradient boosting regressor (GBR) were employed for their robustness and proven performance in similar domains.

b) Developing integrated predictive model using DEA and ML approach (artificial neural network (ANN))

Artificial neural network (ANN) is composed of many artificial neurons that are linked together according to specific network architecture. ANN provides a new way for feature extraction using hidden layers and classification. Multilayer perceptron (MLP) is the most widely used ANN technique for data classification due to its most robust and special type of neural network (Isabona *et al*. 2022). Thus, in this study, we developed four models based on various activation function. This variety of activation functions (refer Table 3) allows a comparative analysis of how different non-linearities in the hidden layers impact the overall model performance. ANN was selected to model deeper and more abstract relationships, particularly where traditional tree-based methods may fall short.

Table 3: The choice of activation function of four predictive models using DEA+ANN approach

Model	Activation function (1^{st} hidden layer)-10	Activation function (2^{nd} hidden layer)-8	Activation function (Output layer)-1
1	ReLU	ReLU	Linear
2	Swish	Swish	Linear
3	Leaky_gut	Leaky_gut	Linear
4	elu	elu	Linear

c) Developing an integrated predictive model using DEA+ML approach (AutoML via genetic programming)

The integration of DEA with AutoML is a burgeoning area of research that aims to enhance the efficiency, accuracy, and applicability of DEA models. An innovative approach that leverages the principles of genetic algorithms to automate the process of model selection, feature engineering, and hyperparameter tuning is via genetic programming (Raglio *et al.* 2020; Schofield & Lensen 2021). The integration of genetic programming into AutoML frameworks can enhance the efficiency and effectiveness of machine learning pipelines by automating complex tasks and providing interpretable solutions. The process which in genetic programming typically includes steps (Figure 3) such as initialization, selection, crossover, mutation, and termination. These steps are iteratively applied to evolve programs that meet predefined fitness criteria.

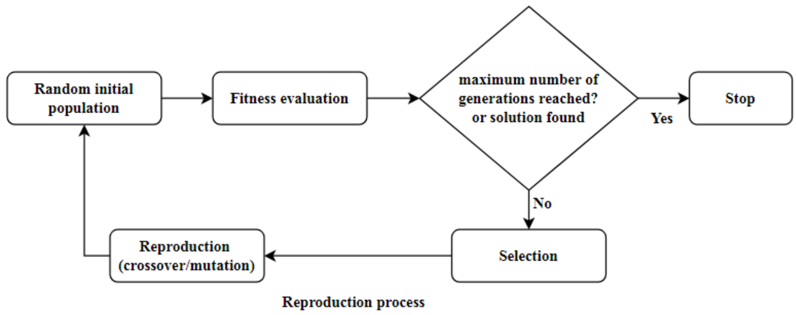

Figure 2: General flowchart for genetic programming

On this study, initially, a DEA model is employed to calculate input-output efficiency scores. All indicators for input and output variables are defined, and the efficiency score is used as the target variable. Before integrated with AutoML via GP, the dataset undergoes preprocessing, including normalization to ensure consistency in data values and assumption verification earlier. At the beginning of the genetic programming process, a population of individual pipelines is randomly generated. This population serves as the starting point for the evolutionary process. The population size is defined as 1099, which represents the number of individual pipelines that will evolve over the generations. From the current population, two individuals are selected based on their fitness scores. The selection mechanism often favors individuals with higher fitness scores, ensuring that better-performing models have a higher likelihood of passing their characteristics to the next generation.

The implementation method of GP machine learning is executed with Python Google Colab platform. The AML TPOT machine learning model has been set to employ 70:30 percent ratio between cross validation for all experiments. TPOT used cross validation for the machine learning training approach and the number of default value for the training and validation chunks is defined as 5. The number of iterations for the GP to set the final optimized pipelines is determined by the number of generations in such that the algorithm should work better with more generations. Since the computational resources constrain cannot handle longer time would be required to complete with bigger number of generations therefore smaller number of generations (5) been used to quickly evaluate the model's general behavior. Population size is the number of individuals or potential pipelines that can be stored in the GP selection pool. Previous experiments by Masrom *et al*. (2020) found that the default population size 10 is fit enough to the algorithm to produce best optimal result. However, since no rule of thumb for determining population size, therefore this study decided to start population size of 10 as the default population size. Six population sizes have been chosen in this experiment (10, 25, 50, and 75, 100, 200) to obtain optimal accuracy.

Mutation rate is a small value used to control the GP exploration search by applying random changes to some of the pipelines in the selection pool. The crossover furthermore is used to tell the GP how many pipelines to reproduce at each of the generation. The value for both must be in between 0.1 to 0.9 and not exceed 1.0. The mutation setting will be 0.7 and crossover rates setting is 0.3. The selection of mutation and crossover rates in genetic programming is a critical factor since these parameters determine how solutions evolve over generations, impacting convergence speed and solution quality. Study by Masrom *et al*. (2020) and Hassanat *et al*. (2019) found that higher crossover rates have been found to improve algorithm accuracy, while lower rates may lead to premature convergence. Thus, this study chooses to conduct an experiment with various mutation rates and crossover rate as stated in Table 4.

Table 4. The Important Parameters in AutoML via GP

Parameters	Configuration value
Generation	5
Population_size	10,25,50,75,100,200
Mutation rate	0.1,0.2,0.3,0.5,0.9,0.8,0.7
Crossover rate	0.1,0.2,0.3,0.5,0.9,0.8,0.7

Every experiment method in Figure 3 was repeated with data split 0.3 and each split ratio is repeated with six population sizes (10, 25, 50, 75, 100, and 200) and 3 validations from the 3 GP generations. Thus, the total experiment run was (7 mutation rates X 7 crossover rate X 6 population size X 1 split ratio (0.3) equal to 294. For each experimental run, three validations from five GP generations in total of 1470 (total experiment runs X GP generation).

The total row of the student achievement efficiency dataset is 1109, which is divided into training and testing set as depicted in Figure 3. If the split ratio is 0.7, 776 out of 1109 were deployed for training set and the rest of 332 data left for testing. For validation, TPOT used cross-validation to divide the training dataset into training and validation set according to the number of k-folds. This research used 5 k-folds. New individuals with the highest predictive value are added to the population after each generation. This process continues until the population size limit is reached (i.e., the number of individuals equals the predefined limit) or the maximum number of generations (j = max) is achieved. Once the genetic programming process reaches the specified stopping criteria, the best-performing pipelines are selected based on their fitness scores and these selected pipelines are then applied to the test dataset to evaluate their predictive accuracy and performance.

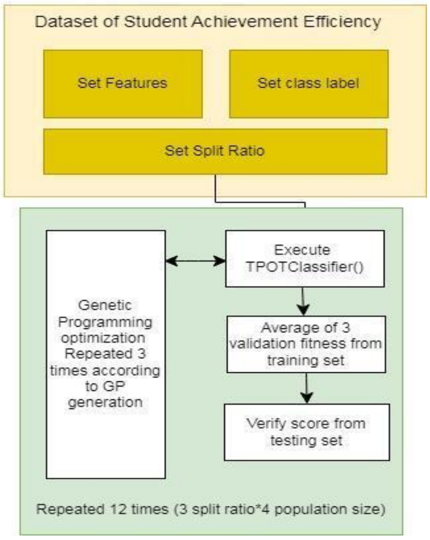

Figure 3. The experimental methodology on GP

Phase 4: Model validation

When evaluating the performance of predictive models, especially in regression tasks, it's crucial to quantify the model's accuracy and reliability. Commonly used metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). Each of these metrics provides different insights into the model's performance and understanding them comprehensively is essential for interpreting results accurately.

Phase 5: Identifying most importance features of integrated predictive model using permutation test

This study employs a comprehensive methodology to identify the most significant features influencing student achievement efficiency. The approach integrates permutation tests and SHAP (SHapley Additive exPlanations) analysis to evaluate feature importance across the best predictive models selected from three ML approaches: conventional ML algorithms, ANN, and AutoML via genetic programming. By combining permutation tests and SHAP, the study ensures a robust assessment of feature significance, enhancing the interpretability of predictive models and the understanding of the factors driving student efficiency. It consists of two main phases: First phases Conducting Permutation Test involves procedures such as (1) Baseline performance measurement, (2) feature vale shuffling, (3) performance evaluation post shuffling and(4)ranking feature importance. While Phase 2 SHAP analysis involved (1) Computing SHAP values and (2) Generating SHAP summary plots.

A larger increase in the error metric (MSE) after shuffling a feature indicates a more significant contribution to the model's predictive power. The permutation test results provide an initial ranking of features based on their importance. This methodology ensures a thorough analysis of feature importance for predicting student achievement efficiency, combining permutation tests for assessing model reliance on each feature and SHAP for explaining feature contributions. This dual approach not only enhances interpretability but also supports data-driven educational strategies.

Phase 6: Comparison on Efficiency score between traditional DEA model and Integrated Predictive models

Once the best predictive model is identified based on the three approaches (Conventional ML, ANN, and AutoML via genetic programming), the next step involves comparing the predicted efficiency scores from this optimal model with the original efficiency scores obtained from the traditional DEA model. This comparison aims to evaluate the effectiveness of integrating ML with DEA in enhancing the prediction accuracy and discriminative power for assessing student achievement efficiency. This study evaluated the best model based on performance metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2), and Spearman's rank correlation coefficient (rho), the most accurate model from the three approaches is selected as the best predictive model. A descriptive statistical analysis to evaluate key measures such as mean, standard deviation, median, maximum, minimum values, and the number of efficient units identified by each approach and will highlight any differences in the distribution of efficiency scores, showing whether the integrated DEA+ML approach provides a more refined or accurate assessment. Spearman Rank Correlation Coefficient used to determine whether the differences between the original DEA efficiency scores and the predicted scores from the best predictive model are significant. Next, the discriminatory power been assessed using Evaluate the number of efficient units (i.e., DMUs with efficiency scores equal to or greater than 1.0) identified by both approaches. The integrated model may be able to better differentiate between efficient and inefficient units, reducing the number of units with perfect efficiency scores and thus improving the model's discriminative ability. The best integrated predictive model using DEA +ML approaches has been identified based on these comparison tests. Overall, the process of developing integrated DEA+ML approaches to predict efficiency scores has been compiled into Figure 4.

4. Results

4.1. *Summary statistics on input and output variables*

The main drawback ohas beenA were its sensitivity to outliers (Dharmapala 2021) and missing values (Chen *et al*. 2020). Especially when using real-life data, outliers are common and decrease the precision of the DEA. To address this issue, comprehensive outlier detection and handling procedure was implemented. To ensure the accuracy and reliability of the DEA model, an outlier detection procedure was implemented using Mahalanobis Distance. This method identifies multivariate outliers by measuring the distance between a point and the center of the dataset (Rajamani & Iyer 2023; Sari *et al*. 2021). A cutting point of Mahalanobis Distance greater than or equal to 18 was established as the threshold for omitting outliers from the dataset. Addressing missing values is crucial because they can lead to biased or invalid results, significantly distorting statistical analyses and model predictions. The missing value analysis revealed some gaps in the data collection process, necessitating the exclusion of certain records. The decision to exclude these records was based on the extent and pattern of missing data, aiming to minimize any potential biases or distortions. Consequently, data points exceeding this threshold were removed, resulting in a cleaned dataset comprising 1099 observations. This sample size is sufficient, as it exceeds the calculated required sample size of 1075 as shown in Table 5. This cleaned dataset ensures a more robust and accurate DEA analysis by mitigating the effects of outliers and addressing the issue of missing data comprehensively. This comprehensive approach ensures that the DEA model's results are accurate and reflective of the true performance of the decision-making units (DMUs), providing valuable insights for improving educational efficiency.

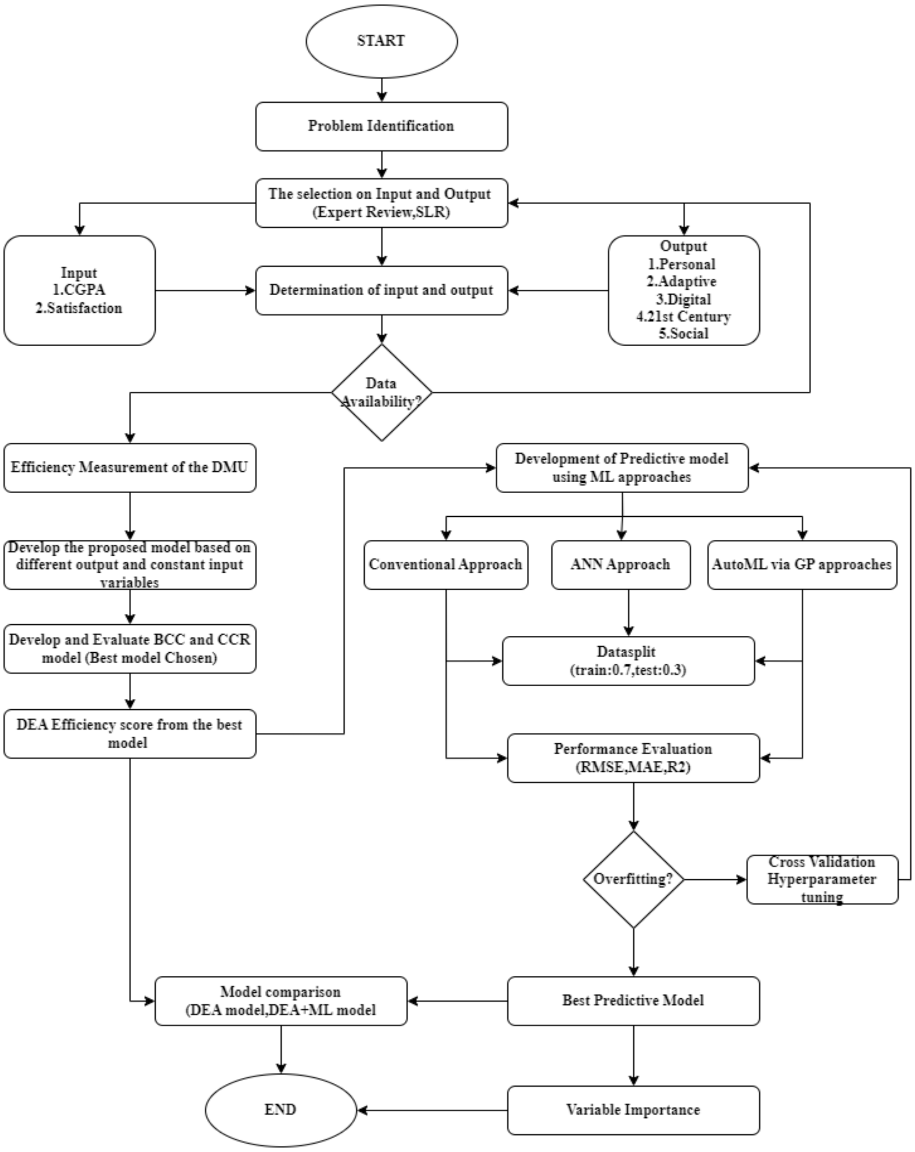

Figure 4. Overall research process for developing predictive efficiency model

Table 5 Descriptive findings on input and output variables

Variables	Mean	SD	Minimum	Maximum
Variable Input:				
CGPA	3.39	0.34	2.06	4.00
Satisfaction	7.99	1.22	1.00	10.00
Variable Output:				
Personal	8.16	1.16	3.89	10.00
Adaptive	8.21	1.20	4.00	10.00
Digital	8.30	1.22	4.00	10.00
21^{st} Century	8.01	1.33	3.40	10.00
Social	8.23	1.24	4.00	10.00

4.2. *Interrelationship between input and output variables*

The correlation matrix presented in Figure 5 shows the relationships between various input and output variables in the study, including CGPA, Satisfaction, Personal Competencies, Adaptive Competencies, Digital Competencies, Social Competencies, and 21st Century Competencies. The values in the matrix represent the Spearman correlation coefficients, which range from -1 to 1. Positive values indicate a positive correlation, while negative values indicate a negative correlation. The strength of the correlation is indicated by the color intensity, with deeper red colors representing stronger positive correlations and deeper blue colors representing weaker or negative correlations.

Figure 5: Correlation analysis for input and output selection

The correlation analysis shows that CGPA has low correlations with other variables (ranging from 0.10 to 0.29), indicating its independence from satisfaction and competencies. Satisfaction is moderately correlated with all competencies (r = 0.62–0.67), suggesting that higher satisfaction is linked to higher competency levels. Strong interrelationships exist among competencies, with Personal, Adaptive, Digital, Social, and 21st Century Competencies all highly correlated (r = 0.81–0.90). These findings validate the selection of input and output variables in the DEA model, as strong correlations enhance the model's ability to assess efficiency accurately (Naseri *et al*. 2020; Dobos & Vörösmarty 2024). The results confirm that student competencies are interdependent, reinforcing the importance of a holistic approach in efficiency analysis. Given the high interrelationships observed among input and output variables in our study, these input and output selections are thus validated.

4.3. *Identifying the optimal student achievement efficiency*

Table 6 presents the efficiency results of four data envelopment analysis (DEA) models, using both the CCR (Charnes, Cooper, and Rhodes) and BCC (Banker, Charnes, and Cooper) approaches. The data includes minimum and maximum efficiency scores, the range, mean, standard deviation, and the number of efficient decision-making units (DMUs) for each model. For the CCR model, the minimum efficiency scores range from 0.4581 (Model 4) to 0.4673

(Model 1), with the maximum efficiency score consistently at 1.0000 across all models. The range of efficiency scores varies slightly, from 0.5333 in Model 1 to 0.5419 in Model 4. The mean efficiency scores show a slight decline from Model 1 (M = 0.7832, SD = 0.0836) to Model 3 (M = 0.7636, SD = 0.0854), with Model 4 at M = 0.7719 (SD = 0.0865). The number of efficient DMUs in the CCR model decreases from 17 in Model 1 to 13 in Model 3, with Model 4 having 15 efficient DMUs.

In contrast, the BCC model shows higher minimum efficiency scores, ranging from 0.4818 (Model 4) to 0.5060 (Model 1). The maximum efficiency score remains at 1.0000 for all models. The range of efficiency scores varies from 0.4940 in Model 1 to 0.5182 in Model 4. The mean efficiency scores are higher than those in the CCR model, ranging from 0.8654 in Model 3 (SD = 0.1024) to 0.8820 in Model 1 (SD = 0.0959). The standard deviations are slightly higher than those in the CCR model, indicating more variability, with values ranging from 0.0959 to 0.1024. The number of efficient DMUs in the BCC model is significantly higher, decreasing from 134 in Model 1 to 110 in Model 3, with Model 4 having 113 efficient DMus. Overall, the BCC model consistently shows higher mean efficiency scores and a greater number of efficient DMUs compared to the CCR model. This difference highlights the BCC model's ability to account for variable returns to scale, providing a more nuanced and often more favorable efficiency assessment. The variability in efficiency scores, as indicated by the standard deviations, suggests that the BCC model captures a broader range of efficiency levels among DMUs.

Table 6. Summary statistics for proposed CCR and BCC model

Item	CCR Model Efficiency				BCC Model Efficiency			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Minimum	0.4673	0.4612	0.4595	0.4581	0.5060	0.5033	0.5022	0.4818
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Range	0.5333	0.5388	0.5405	0.5419	0.4940	0.4967	0.4978	0.5182
Mean	0.7832	0.7736	0.7636	0.7719	0.8820	0.8738	0.8654	0.8703
Std. deviation	0.0836	0.0841	0.0854	0.0865	0.0959	0.0988	0.1024	0.1012
Number of efficient DMUs	17	16	13	15	134	121	110	113

Furthermore, this analysis compares the discriminatory capabilities of the Banker, Charnes, and Cooper (BCC) model and the Charnes, Cooper, and Rhodes (CCR) model in evaluating the efficiency of decision-making units (DMUs). The BCC model exhibits superior discriminatory power, as evidenced by its ability to recognize a larger number of efficient units and a greater standard deviation in efficiency scores. This suggests that the BCC model excels at differentiating DMUs based on their operational scales and efficiency characteristics, which is essential for nuanced analysis and effective policy making. In contrast, the CCR model demonstrates moderate discriminatory power. It effectively distinguishes between fully efficient units and others to a certain extent; however, its lower standard deviation and fewer identified efficient units imply a reduced sensitivity to variations among DMUs, especially those not operating at optimal scales. This characteristic suggests that the CCR model may not capture the full range of operational efficiencies in environments where variable returns to scale are significant.

The enhanced discriminatory power of the BCC model makes it particularly suitable for detailed efficiency analyses. By identifying a broader range of efficient units, the BCC model facilitates the discovery of best practices and supports benchmarking efforts, thus fostering operational improvements across units. The application of the BCC model in academic research enhances the credibility and relevance of findings Dellnitz *et al*. (2018), as demonstrated in Table 4.13. The model's thorough approach to assessing efficiency under variable returns to scale, as highlighted by Dellnitz & Rödder (2020), ensures that the research conclusions are robust, practical, and accurately represent the operational dynamics at play. This is crucial for studies that aim to influence policy or organizational strategies. The selection of the BCC model for this study is based on its proven ability to handle variable operational conditions and its exceptional capacity to differentiate performance among DMUs. By accurately identifying the sources of inefficiencies whether they stem from underutilization of resources or diminishing returns the BCC model supports more targeted recommendations for improving student performance, aligning closely with educational strategies that emphasize personalized and competency-based learning approaches. Sun (2017) supported that in educational context, different models such as the exam review model and talent cultivation model highlight various aspects of learning efficiency. Therefore, the BCC model's flexibility, its ability to capture variable returns to scale, and its demonstrated effectiveness in complex settings make it a highly appropriate and theoretically robust tool for evaluating student academic achievement. Since Model 1 under BCC model identifies the largest number of efficient DMUs (134) and may offers the broadest insight into students efficiency based on the chosen input and output relationship, thus this study decided to choose as the best model since it comprehensively assesses efficiency and identify subtle performance differences among students.

4.4. *Integrated prediction model using DEA and ML approaches*

4.4.1. *DEA and ML approaches (conventional approach)*

The dataset exhibited deviations from normality, heteroscedasticity, and linearity, requiring adjustments such as robust standard errors and nonparametric methods. After evaluating multiple regressor, gradient boosting regressor (GBR) demonstrated the best generalization with minimal overfitting, followed by random forest as a reliable alternative. KNN and decision tree faced overfitting issues, while SVR improved after hyperparameter tuning, as shown in the performance summary Table 7.

4.4.2. *DEA with ML (artificial neural network approach)*

The comparative analysis of multi-layer perceptron artificial neural networks (MLP-ANN) using different activation functions revealed that activation choice significantly impacts model performance. Model 1 (ReLU) showed moderate generalization with a test RMSE of 0.0450 and R^2 of 0.7687. Model 2 (Swish) improved generalization, achieving a lower test RMSE of 0.0391 and R^2 of 0.8249. Model 3 (Leaky ReLU) performed best in training (RMSE = 0.0336, R^2 = 0.8779) but had slightly lower test accuracy (RMSE = 0.0404, R^2 = 0.8130). Model 4 (ELU) demonstrated the most balanced performance, achieving the lowest test RMSE (0.0370) and highest R^2 (0.8434), indicating superior generalization. These results suggest that ELU activation provides smoother convergence and enhanced learning, making it ideal for educational datasets that require robust handling of noise (Cococcioni *et al*. 2020; Maurya *et al*. 2023).

Table 7. The findings of regression model performance (DEA + conventional ML approach)

Regressor	Hyperparameter Tuning	Data Split	MSE	RMSE	MAE
KNN	Cross Validation	Train	0.0007	0.0279	0.0166
		Test	0.0014	0.0382	0.0232
	GridSearch	Train	0.0000	0.0000	0.0000
		Test	0.0008	0.0292	0.0177
Decision Tree	Cross Validation	Train	0.0000	0.0000	0.0000
		Test	0.0008	0.0289	0.0129
	GridSearch	Train	0.0001	0.0102	0.0049
		Test	0.0008	0.0286	0.0145
Random Forest	Cross Validation	Train	0.00008	0.0091	0.0036
		Test	0.0006	0.0247	0.0115
	GridSearch	Train	0.00008	0.0091	0.0036
		Test	0.0006	0.0247	0.0115
SVR	Cross Validation	Train	0.0022	0.0474	0.0349
		Test	0.0028	0.0532	0.0368
	GridSearch	Train	0.00005	0.0077	0.0069
		Test	0.0002	0.0166	0.0110
GBR	Cross Validation	Train	0.0000	0.0021	0.0014
		Test	0.0003	0.0181	0.0094
	GridSearch	Train	0.0000	0.0035	0.0026
		Test	0.0002	0.0158	0.0099

Figure 6 describes the architecture and performance of the best MLP-ANN model using backpropagation. This model consists of an input layer with 7 neurons, two hidden layers, and an output layer with 1 neuron. The first hidden layer contains 40 neurons, and the second hidden layer contains 38 neurons. The activation function used for both hidden layers is the Exponential Linear Unit (ELU), while the output layer uses a linear activation function. Based on the Model Architecture this MLP-ANN model consists of 7 neurons (corresponding to the seven input features for input Layer, two hidden layers where the first hidden input layer consists of 40 neurons with ELU activation function while second hidden layer consists of 38 neurons with ELU activation function. The output Layer consists of 1 neuron with linear activation function which is found suitable for regression tasks. Throughout the experiment task, activation function, ELU (Exponential Linear Unit) used in the hidden layers to introduce non-linearity, helping the model learn complex patterns while avoiding issues like vanishing gradients meanwhile Linear activation function used in the output layer to predict continuous values, which is appropriate for regression problems.

Table 8. Comparison of MLP ANN with different activation functions

Model	Activation function (1st hidden layer)-10	Activation function (2nd hidden layer)-8	Activation function (Output layer)-1	Performance Train dataset			Performance Test dataset		
				RMSE	MAE	R^2	RMSE	MAE	R^2
1	ReLU	ReLU	Linear	0.037	0.025	0.862	0.045	0.028	0.769
2	Swish	Swish	Linear	0.037	0.026	0.851	0.039	0.026	0.824
3	Leaky_gut	Leaky_gut	Linear	0.033	0.023	0.878	0.040	0.025	0.813
4	**elu**	**elu**	**Linear**	**0.032**	**0.023**	**0.889**	**0.037**	**0.024**	**0.843**

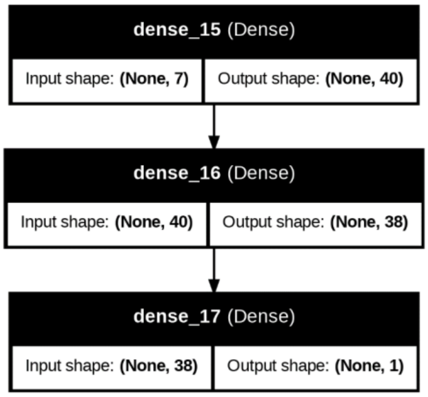

Figure 6: DEA-MLP network

4.4.3. *DEA and ML (AutoML via genetic programming approach)*

As a summary presented in Table 9, Model DEA+XGB1, the ΔRMSE is 0.0137 and ΔMAE is 0.0073, indicating a slight increase in error from training to testing, but overall, the model retains an extremely high R^2 value of 0.9995, showing excellent predictive performance. However, Model DEA+XGB2, the ΔRMSE is 0.0116 and ΔMAE is 0.0058, with a slightly higher test error compared to training, but the R^2 value of 0.9559 still indicates strong predictive power, albeit with slightly more variability in the test set. This suggests that both configurations perform exceptionally well, but the Model DEA+XGB1 achieves a higher level of precision and minimal error across training and testing datasets.

Table 10 compares efficiency scores across DEA integrated with ML models: random forest (DEA-RF), gradient boosting regressor (DEA-GBR), artificial neural networks (DEA-ANN), and AutoML (genetic programming) (DEA-AutoML (GP)). DEA-GBR (0.8808) and DEA-AutoML (GP) (0.8799) have the closest mean efficiency to the original DEA, while DEA-RF (0.8786) and DEA-ANN (0.8781) are slightly lower. Standard deviation is smallest for DEA-RF (0.0936), indicating less variability, while DEA-GBR (0.0957) and DEA-AutoML (GP) (0.0950) closely resemble DEA. Median efficiency is highest for DEA-AutoML (GP) (0.9002), followed by DEA-GBR (0.9001). Overall, DEA-GBR and DEA-AutoML (GP) best align with DEA in efficiency distribution, while DEA-RF is more conservative with lower variability. DEA-ANN offers more flexibility but predicts moderate efficiency. DEA-GBR and DEA-AutoML (GP) are the most reliable in replicating DEA results, while DEA-RF provides a stable alternative.

Table 9. DEA+XGB regressor performance result

Model	Best Pipelines	RMSE	MAE	R^2
1	XGBRegressor(learning_rate=0.1, max_depth=8, min_child_weight=2, n_estimators=100)	0.0182	0.0087	0.9622
2	XGBRegressor(learning_rate=0.1, max_depth=7, min_child_weight=2)	0.0186	0.0110	0.9605
3	XGBRegressor(learning_rate=0.1, max_depth=6, min_child_weight=4)	0.0185	0.0103	0.9606
4	XGBRegressor(learning_rate=0.1, max_depth=6, min_child_weight=6)	0.0202	0.0107	0.9536
5	XGBRegressor(XGBRegressor(learning_rate=0.1, max_depth=7, min_child_weight=2, n_estimators=100, n_jobs=1, objective=reg:squarederror, subsample=0.3, verbosity=0), learning_rate=0.5, max_depth=8, min_child_weight=14,)	0.0188	0.0112	0.9596

Table 10: Descriptive summaries of four integrated predictive DEA+ ML models

Descriptive	DEA	DEA-RF	DEA-GBR	DEA-ANN	DEA-GP
Mean	0.8820	0.8786	0.8808	0.8781	0.8807
Std.deviation	0.0959	0.0936	0.0957	0.0956	0.0950
Median	0.9000	0.9000	0.9001	0.8989	0.9002
Maximum	1.0000	1.0000	1.0200	1.0484	1.0093
Minimum	0.5060	0.5368	0.5081	0.5113	0.5177
No of efficient	134	47	44	74	50

Further statistical analysis has been conducted on testing datasets and found that Table 11 revealed the result on performance metrices on predicted models. DEA_GBR consistently outperforms the other models, achieving the lowest RMSE (0.0101) and MAE (0.0039), which indicates it provides the most accurate predictions. Additionally, it has the highest R^2 (0.9889) and Spearman's rho (0.995), meaning it explains the most variance in the data and ranks the efficiency scores most accurately. DEA_RF also performs well, with strong R^2 (0.9736) and Spearman's rho (0.987), but it has slightly higher RMSE and MAE compared to DEA_GBR, indicating somewhat less accurate predictions. DEA_GP (XGB) demonstrates good performance, with an R^2 of 0.9622 and a high Spearman's rho (0.995), but it has a higher RMSE (0.0182) and MAE (0.0087) than DEA_GBR. DEA_ANN, however, performs the worst among the models, with the highest RMSE (0.0242) and MAE (0.0158), and the lowest R^2 (0.9361), making its predictions less reliable compared to the other models.

Table 11. Common statistical parameters of four DEA-ML algorithms for testing datasets

Model	RMSE	MAE	R^2	Spearman's rho
DEA_RF	0.0156	0.0060	0.9736	0.987**
DEA_GBR	0.0101	0.0039	0.9889	0.995**
DEA_ANN	0.0242	0.0158	0.9361	0.969**
DEA_GP(XGB)	0.0182	0.0087	0.9622	0.995**

DEA_GBR emerges as the best-performing model, excelling in predictive accuracy with the lowest RMSE and MAE, the highest R^2 and Spearman's rho, and its ability to maintain a distribution of efficiency scores similar to the original DEA model. DEA_RF is also a strong contender, offering reliable predictions with minimal variability, though it is more conservative in identifying efficient DMUs. DEA_GP (XGB) is a close alternative to DEA_GBR but has slightly higher error rates. In contrast, DEA_ANN, while allowing for the greatest over-efficiency, performs the worst in terms of prediction accuracy and consistency. Based on both the descriptive statistics and performance metrics, DEA_GBR is the best overall model, followed by DEA_RF and DEA_GP (XGB), with DEA_ANN being the least favorable due to its higher error rates and lower predictive accuracy. This result consistent with findings by (Burnaev & Boldyreva 2024; Langenberger *et al*. 2023; Sukiasyan 2023) that gradient boosting regressor (GBR) and random forest (RF), have shown significant effectiveness in predicting efficiency scores across various domains such as healthcare, education, and manufacturing. These models are valued for their ability to handle complex datasets and provide accurate predictions, making them suitable for diverse applications.

Not only that, the involvement genetic programming approach in this study shows promise in improving prediction accuracy due to its performance in this study. GP has been shown to effectively balance the trade-off between interpretability and accuracy in predictive modeling. The combination of DEA and GP has been particularly effective. These have been proven by

Panigrahi *et al*. (2018) produced a DEA-based evolutionary computation model for stock market forecasting, demonstrating better performance by efficiently selecting input variables, leading to improved prediction accuracy.

5. Discussion

This study compared four integrated DEA-ML models (DEA-RF, DEA-GBR, DEA-ANN, and DEA-GP) using descriptive statistics and performance metrics (Table 5X). Among them, DEA-GBR demonstrated the highest median efficiency score (0.9001) and slightly higher mean efficiency (0.8808), making it the best-performing model, followed by DEA-RF and DEA-GP. DEA-GP exhibited stable efficiency predictions with a narrower range, while DEA-ANN had the highest variability, aligning with Zhang *et al*. (2019), who noted that ANN models can produce outlier predictions when not optimally tuned. The integration of ML significantly enhanced DEA's discriminatory power, reducing the number of efficient DMUs from 134 (BCC model) to 47 (DEA-RF), 44 (DEA-GBR), 74 (DEA-ANN), and 50 (DEA-GP), supporting findings from (Kordrostami & Mirmousavi 2013; Lam 2018; Pendharkar 2013; Peyrache & Silva 2024) on the role of ML in refining DEA assessments.

Gradient boosting regressor (GBR) and random forest (RF) outperformed other models in predictive accuracy, consistent with Burnaev and Boldyreva (2024), Langenberger *et al*. (2023), and Sukiasyan (2023), who highlighted their effectiveness across healthcare, education, and manufacturing domains. Genetic programming (GP) also showed promise in balancing accuracy and interpretability, aligning with Panigrahi *et al.* (2018), who demonstrated the potential of DEA-based evolutionary computation for improved prediction accuracy. Feature importance analysis using SHAP (refer Table 12) revealed that Digital Competency was the strongest predictor across all models, aligning with UiTM's digital transformation vision as emphasized by Md Zain (2020). Social Competency ranked second in most models, highlighting its role in student efficiency, while Adaptive Competency was moderately important. CGPA and Satisfaction had limited predictive impact, ranking lower across models, and 21st Century Competency was consistently the least influential. These findings suggest that non-academic factors particularly digital and social skills play a crucial role in shaping academic efficiency. This underscores the importance of investing in digital infrastructure, competency-based curriculum design, and adaptive learning strategies to enhance student outcomes. Overall, the integration of ML techniques with DEA provides a robust, interpretable, and scalable framework for evaluating individual student performance and informing institutional policy and academic decision-making.

Table 12 Variable importance based on DEA+ML model

Variable	DEA-RF	DEA-GBR	DEA-ANN	DEA-GP
CGPA	6	6	5	6
Satisfaction	5	5	2	5
Personal	3	3	6	4
Adaptive	4	4	3	3
21^{st} Century	7	7	7	7
Digital	1	1	1	1
Social	2	2	4	2

6. Conclusion

This study aimed to develop and validate a DEA model for assessing the academic achievement efficiency of final-year diploma students and to enhance predictive accuracy by integrating

DEA with ML techniques. The Student Competencies Questionnaire (SCQ), guided by Pillar 1 of the Education 5.0 Framework, was developed and validated to ensure reliable data collection. Among the four DEA models analyzed, BCC Model 1 emerged as the most effective for evaluating student efficiency. To further enhance the model's predictive power, DEA was integrated with ML approaches, including random forest, gradient boosting, ANN, and GP, leading to a significant reduction in the number of efficient decision-making units (DMUs) from 134 (using DEA alone) to 47 (DEA-RF), 44 (DEA-GBR), 74 (DEA-ANN), and 50 (DEA-GP). This integration improved the discriminatory power of efficiency assessments, offering a more refined evaluation of student performance.

Despite the promising outcomes of this study, several limitations merit careful consideration. First, the dataset was confined to final-year diploma students within a specific academic setting, which may limit the generalizability of the findings to broader and more diverse student populations. Second, although the integration of DEA with machine learning significantly improved the discriminatory power of efficiency assessments, it introduced substantial computational demands particularly with the GP approach which posed challenges related to processing time and resource optimization. Third, while the selection of input variables (CGPA and satisfaction) and output variables (student competencies) was theoretically justified, it may not fully encapsulate the complex and multidimensional nature of academic achievement at the individual student level, which was the central focus of this study. In contrast to traditional DEA applications that assess institutional or organizational units, this research evaluated each student as a distinct decision-making unit (DMU), thereby introducing additional complexity arising from personal learning trajectories, motivation levels, and behavioral variability. Moreover, the use of self-reported data through the Student Competencies Questionnaire (SCQ), despite undergoing thorough validation procedures, may be subject to response bias and subjective interpretation. To address these limitations, future research should incorporate more diverse and objective indicators, such as continuous assessment marks and data from student entrance and exit surveys, to better capture longitudinal learning outcomes and competency development. Expanding the dataset to include multiple institutions and academic programs, along with the deployment of more advanced computational infrastructures, would further enhance the validity, scalability, and generalizability of the proposed DEA-ML framework for assessing academic efficiency at the individual level.

Acknowledgments

Thank you to the Universiti Teknologi MARA for the full support of this research.

References

Abramo G., Cicero T. & D'Angelo C.A. 2011. A field-standardized application of DEA to national-scale research assessment of universities. *Journal of Informetrics* **5**(4): 618-628.

Adna N., Yusoff N.S.M. & Sheikh Zolknain S.S.A. 2025. Measuring the performance of public universities Malaysia using CCR model and ranking with super efficiency model. *Malaysia Journal of Invention and Innovation* **4**(5): 54–60.

Ahmad S., Umirzakova S., Mujtaba G., Amin M.S. & Whangbo T. 2023. *Education 5.0: Requirements, enabling technologies, and future directions*. arXiv. https://arxiv.org/abs/2307.15846.

Ahn H., Clermont M. & Langner J. 2022. The impact of selected input and output factors on measuring research efficiency of university research fields: Insights from a purpose-, field-, and method-specific perspective. *Journal of Business Economics* **92**(8): 1303-1335.

Akhavan Kharazian M., Shahbazi M.M. & Fatehi M. 2019. Exploring the optimum pattern for knowledge workers selection using DEA and CART compilation approach. *Research in Production and Operations Management* **10**(1): 65–82.

Alcaraz J., Anton-Sanchez L., Aparicio J., Monge J.F. & Ramón N. 2021. Russell Graph efficiency measures in data

envelopment analysis: The multiplicative approach. *European Journal of Operational Research* **292**(2): 663–674.
Anouze A.L.M. & Bou-Hamad I. 2019. Data envelopment analysis and data mining to efficiency estimation and evaluation. *International Journal of Islamic and Middle Eastern Finance and Management* **12**(2): 169–190.
Avramidou A. & Tjortjis C. 2021. Predicting CO_2 emissions for buildings using regression and classification. *IFIP Proceedings of the International Conference on Artificial Intelligence Applications and Innovations*, pp. 543-554.
Awang Z., Afthanorhan A. & Mamat M. 2016. The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences* **4**(1): 13-21.
Aziz N.A.A., Janor R.M. & Mahadi R. 2013. Comparative departmental efficiency analysis within a university: A DEA approach. *Procedia - Social and Behavioral Sciences* **90**: 540–548.
Barnett V. & Lewis T. 1994. *Outliers in Statistical Data*. 3rd Ed. John Wiley & Sons. Kluwer Academic Publishers, Boston/Dordrecht/London.
Barnett V. & Lewis T. 1994. *Outliers in Statistical Data*. 3rd Ed. Chichester: John Wiley & Sons.
Behle H. 2020. Students' and graduates' employability. A framework to classify and measure employability gain. *Policy Reviews in Higher Education* **4**(1): 105–130.
Burnaev M.A. & Boldyreva E.A. 2024. Machine learning methods for predicting the performance of online courses. *Proceedings of the V International Conference on Neural Networks and Neurotechnologies (NeuroNT)*, pp. 19-21.
Camanho A.S., Varriale L., Barbosa F. & Sobral T. 2021. Performance assessment of upper secondary schools in Italian regions using a circular pseudo-Malmquist index. *European Journal of Operational Research* **289**(3): 1188–1208.
Chen C., Ren J., Tang L. & Liu H. 2020. Additive integer-valued data envelopment analysis with missing data: A multi-criteria evaluation approach. *PLoS ONE* **15**(6): e0234247.
Cinaroglu S. 2021. Oncology services efficiency in the age of pandemic: A jackknife and bootstrap sensitivity analysis for robustness check of DEA scores. *Journal of Cancer Policy* **27**: 100262.
Cococcioni M., Rossi F., Ruffaldi E. & Saponara S. 2020. A novel posit-based fast approximation of ELU Activation function for deep neural networks. *Proceedings of the 2020 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 244-246.
Contreras I. & Lozano S. 2022. Size efficiency, splits and merger gains, and centralized resource reallocation of Spanish public universities. *Socio-Economic Planning Sciences* **81**: 101190.
Cooper W., Seiford L.M. & Tone K. 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software.* 2nd Ed. New York: Springer.
Datnow A., Park V., Peurach D.J. & Spillane J.P. 2022. Transforming education for holistic student development: Learning from education system (re)building around the world [Report]. The Brookings Institution, Center for Universal Education. https://eric.ed.gov/?id=ED626329
de Abreu E.S. & Kimura H. 2020. Determinants of efficiency in state-chartered financial institutions: Why financial education and freedom matter. *Heliyon* **6**(12): e05795.
De La Hoz E., Zuluaga R. & Mendoza A. 2021. Assessing and classification of academic efficiency in engineering teaching programs. *Journal on Efficiency and Responsibility in Education and Science* **14**(1): 41–52.
Dellnitz A., Kleine A. & Rodder W. 2018. CCR or BCC: What if we are in the wrong model? *Journal of Business Economics* **88**(7): 831-850.
Dellnitz A. & Rödder W. 2020. Returns to scale as an established scaling indicator: Always a good advisor? *Jahrbücher für Nationalökonomie und Statistik* **241**(2): 173–186.
Dharmapala P.S. 2021. Detecting outliers in DEA after correcting biases in efficiency scores using simulated beta samples. *International Journal of Mathematics in Operational Research* **19**(3): 387-400.
Dobos I. & Vörösmarty G. 2024. Input and output reconsidered in supplier selection DEA model. *Central European Journal of Operations Research* **32**(1): 67-81.
Ebrahimi B. & Hajizadeh E. 2021. A novel DEA model for solving performance measurement problems with flexible measures: An application to Tehran Stock Exchange. *Measurement* **179**: 109444.
Emrouznejad A. & Shale E. 2009. A combined neural network and DEA for measuring efficiency of large scale datasets. *Computers & Industrial Engineering* **56**(1): 249–254.
Esteve M., Aparicio J., Rabasa A. & Rodriguez-Sala J.J. 2020. Efficiency analysis trees: A new methodology for estimating production frontiers through decision trees. *Expert Systems with Applications* **162**: 113783.
Farantos G.I. 2015. The data envelopment analysis method and the influence of a phenomenon in organizational Efficiency: A literature review and the Data Envelopment Contrast Analysis new application. *Data Envelopment Analysis and Decision Science* **2015**(2): 101–117.
Ghaffarian Asl S. & Osam N. 2021. A study of teacher performance in English for academic purposes course: Evaluating efficiency. *Sage Open* **11**(4).
Ghildiyal S., Garg R., Dhanik S., Sarkar A., Sharma G. & Kharayat P.S. 2024. A comprehensive review analysis of

supervised machine learning techniques. *Proceedings of the 2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 925-930.

Gralka S., Wohlrabe K. & Bornmann L. 2019. How to measure research efficiency in higher education? Research grants vs. publication output. *Journal of Higher Education Policy and Management* **41**(3): 322–341.

Hassanat A., Almohammadi K., Alkafaween E.A., Abunawas E., Hammouri A. & Prasath V.S. 2019. Choosing mutation and crossover ratios for genetic algorithms—A review with a new dynamic approach. *Information* **10**(12): 390.

Isabona J., Imoize A.L., Ojo S., Karunwi O., Kim Y., Lee C.C. & Li C.T. 2022. Development of a multilayer perception neural network for optimal predictive modeling in urban microcellular radio environments. *Applied Sciences* **12**(11): 5713.

Jauhar S.K., Pant M. & Nagar A.K. 2017. Sustainable educational supply chain performance measurement through DEA and differential evolution: A case on Indian HEI. *Journal of Computational Science* **19**: 138–152.

Jauhar S.K., Zolfagharinia H. & Amin S.H. 2023. A DEA-ANN-based analytical framework to assess and predict the efficiency of Canadian universities in a service supply chain context. *Benchmarking: An International Journal* **30**(8): 2734–2782.

Johnes J. 2015. Operational research in education. *European Journal of Operational Research* **243**(3): 683–696.

Jomthanachai S., Wong W.P. & Lim C.P. 2021. An application of data envelopment analysis and machine learning approach to risk management. *IEEE Access* **9**: 85978–85994.

Kannan K.R., Abarna K.T.M. & Vairachilai S. 2024. Enhancing student academic performance forecasting in technical education: A cutting-edge hybrid fusion method. *SSRG International Journal of Electronics and Communication Engineering* **11**(12): 146–153.

Khezrimotlagh D. 2015. How to deal with numbers of decision making units and variables in data envelopment analysis. arXiv. https://arxiv.org/abs/1503.02306

Khoubrane Y., Ramli N.A. & Khairi S.S.M. 2024. 1. Performance measurement: Machine Learning as a complement to DEA for continuous efficiency estimation. *Malaysian Journal of Fundamental and Applied Sciences* **20**(2): 288-301.

Kordrostami S. & Mirmousavi A. 2013. A new method for improving the discrimination power and weights dispersion in the data envelopment analysis. *Journal of Mathematical Extension* **7**(2): 49–56.

Laksaci A., Kaid Z., Alahiane M., Ouassou I. & Rachdi M. 2023. Non parametric estimations of the conditional density and mode when the regressor and the response are curves. *Communications in Statistics-Theory and Methods* **52**(13): 4659–4674.

Lam K.F. 2018. A discriminant analysis and goal programming approach to solve the multiple criteria data envelopment analysis model. *Journal of Advanced Management Science* **6**(4): 222-226.

Langenberger B., Schulte T. & Groene O. 2023. The application of machine learning to predict high-cost patients: A performance-comparison of different models using healthcare claims data. *PLoS ONE* **18**(1): e0279540.

Leys C., Klein O., Dominicy Y. & Ley C. 2018. Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology* **74**: 150–156.

Lindsay A.W. 1982. Institutional performance in higher education: The efficiency dimension. *Review of Educational Research* **52**(2): 175-199.

Liu S., Sun J.X., Lyu C., Chu T.J. & Zhang H.X. 2021. Evaluating fishing capacity based on DEA and regression analysis of China's offshore fishery. *Journal of Marine Science and Engineering* **9**(12): 1402.

Masrom S., Mohamad M., Hatim S.M., Baharun N., Omar N. & Rahman A.S.A. 2020. Different mutation and crossover set of genetic programming in an automated machine learning. *IAES International Journal of Artificial Intelligence* **9**(3): 402-408.

Maurya R., Aggarwal D., Gopalakrishnan T. & Pandey N.N. 2023. Enhancing deep neural network convergence and performance: A hybrid activation function approach by combining ReLU and ELU activation function. *Proceedings of the 2023 Second International Conference on Informatics (ICI)*, pp. 1-5.

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

Mohamad Razi N.F., Baharun N. & Masrom S. 2024. Efficiency in academic evaluation: Data envelopment analysis systematic review. *International Journal of Evaluation and Research in Education* **13**(4): 2450–2460.

Mosbah E., Zaibet L. & Dharmapala P.S. 2020. A new methodology to measure efficiencies of inputs (outputs) of decision making units in Data Envelopment Analysis with application to agriculture. *Socio-Economic Planning Sciences* **72**: 100857.

Munoz D.A. 2016. Assessing the research efficiency of higher education institutions in Chile: A data envelopment analysis approach. *International Journal of Educational Management* **30**(6): 809–825.

Nandy A. & Singh P.K. 2021. Application of fuzzy DEA and machine learning algorithms in efficiency estimation of paddy producers of rural Eastern India. *Benchmarking: An International Journal* **28**(1): 229–248.

Naseri H., Najafi S.E. & Saghaei A. 2020. DEA model considering outputs with stochastic noise and a heavy-tailed

(stable) distribution. *INFOR: Information Systems and Operational Research* **58**(1): 87-108.
Ochieng' Odhiambo F. 2020. Comparative study of various methods of handling missing data. *Mathematical Modelling and Applications* **5**(2): 87-93.
Özsoy V.S. & Örkcü H.H. 2021. Structural and operational management of Turkish airports: A bootstrap data envelopment analysis of efficiency. *Utilities Policy* **69**: 101180.
Panigrahi S.S., Mantri J.K. & Gahan P. 2018. A DEA-based evolutionary computation model for stock market forecasting. *Proceedings of the International Conference on Microelectronics, Computing & Communication Systems*, pp. 139–148.
Pendharkar P.C. 2013. Ensemble based ranking of decision making units. *INFOR: Information Systems and Operational Research* **51**(3): 151–159.
Petrelli M. 2023. Supervised machine learning methods. In *Machine learning for Earth Sciences: Using Python to Solve Geological Problems*: 99–116. Cham: Springer.
Peyrache A. & Silva M.C.A. 2024. The decomposition of efficiency in parallel network production models. *Omega* **127**: 103084.
Pietrzak M., Pietrzak P. & Baran J. 2016. Efficiency assessment of public higher education with the application of data envelopment analysis: The evidence from Poland. *Online Journal of Applied Knowledge Management* **4**(2): 59–73.
Pokushko M., Stupina A., Medina-Bulo I. & Dresvianskii E. 2020. Data envelopment analysis in performance assessment of fuel and energy complex enterprises. *Bulletin of Kemerovo State University. Series: Political, Sociological and Economic Sciences* **5**(2): 251–262.
Raab R.L. & Lichty R.W. 2002. Identifying subareas that comprise a greater metropolitan area: The criterion of county relative efficiency. *Journal of Regional Science* **42**(3): 579–594.
Rajamani S.K. & Iyer R.S. 2023. Enhancing outlier detection in healthcare data through Mahalanobis distance metric analysis. In *Digitalization and The Transformation of The Healthcare Sector*: pp. 237–260. Hershey, PA: IGI Global Scientific Publishing.
Ray S.C. 2022. Data envelopment analysis: A nonparametric method of production analysis. In Ray S.C., Chambers R.G. & Kumbhakar S.C. (eds.). *Handbook of Production Economics*: 409–470. Singapore: Springer Nature Singapore.
Raglio A., Imbriani M., Imbriani C., Baiardi P., Manzoni S., Gianotti M., Castelli M., Vanneschi L., Vico F. & Manzoni L. 2020. Machine learning techniques to predict the effectiveness of music therapy: A randomized controlled trial. *Computer Methods and Programs in Biomedicine* **185**: 105160.
Ruwaida Md Zain. 2020. Bangga, UiTM menuju era kampus digital. *My Metro*. https://www.hmetro.com.my/mutakhir/2020/02/547002/bangga-uitm-menuju-era-kampus-digital
Sagar G.V., Ambareesh S., Augustine P.J. & Gayathri R. 2023. Classification and regression algorithms. In de Albuquerque V.H.C., Raj P. & Yadav S.P. (eds.). *Toward Artificial General Intelligence: Deep Learning, Neural Networks, Generative AI*: 53–86. Berlin: Walter de Gruyter GmbH.
Sagarra M., Mar-Molinero C. & Agasisti T. 2017. Exploring the efficiency of Mexican universities: Integrating data envelopment analysis and multidimensional scaling. *Omega* **67**: 123–133.
Sampath Kumar S., Manjunatha Reddy B.N. & Parameshwara M.C. 2023. Regression-based approach for paddy crop assists for atmospheric data. *Proceedings of the International Conference on Emerging Research in Computing, Information, Communication and Applications*, pp. 415-429.
Sari P.P., Herdiani E.T. & Sunusi N. 2021. Outlier detection using minimum vector variance algorithm with depth function and Mahalanobis distance. *Jurnal Matematika, Statistika dan Komputasi* **17**(3): 418-427.
Schofield F. & Lensen A. 2021. Using genetic programming to find functional mappings for UMAP embeddings. *2021 IEEE Congress on Evolutionary Computation (CEC)*, pp. 704-711.
Segovia-Gonzalez M.M., Dominguez C. & Contreras I. 2020. An assessment of the efficiency of Spanish schools: evaluating the influence of the geographical, managerial, and socioeconomic features. *International Transactions in Operational Research* **27**(4): 1845–1868.
Shero J.A., Al Otaiba S., Schatschneider C. & Hart S.A. 2022. Data envelopment analysis (DEA) in the educational sciences. *The Journal of Experimental Education* **90**(4): 1021–1040.
Singpai B. & Wu D. 2020. Using a DEA–AutoML approach to track SDG achievements. *Sustainability* **12**(23): 10124.
Sukiasyan A. 2023. Evaluating machine learning models for healthcare services efficiency. *Cadernos de Educação Tecnologia e Sociedade* **16**(4): 1280-1289.
Sun B. 2017. Mathematical models of learning efficiency. *Eurasia Journal of Mathematics, Science and Technology Education* **13**(7): 4261-4270.
Tavares R.S., Angulo-Meza L. & Sant'Anna A.P. 2021. A proposed multistage evaluation approach for Higher Education Institutions based on network Data envelopment analysis: A Brazilian experience. *Evaluation and Program Planning* **89**: 101984.
Taylan P. 2020. On foundations of estimation for nonparametric regression with continuous optimization. In

García Márquez F.P. (ed.). *Handbook of Research on Big Data Clustering and Machine Learning*: 177–203. Hershey, PA: IGI Global.

Toloo M., Keshavarz E. & Hatami-Marbini A. 2021. Selecting data envelopment analysis models: A data-driven application to EU countries. *Omega* **101**: 102248.

Tran C.D.T.T., Battese G.E. & Villano R.A. 2020. Administrative capacity assessment in higher education: The case of universities in Vietnam. *International Journal of Educational Development* **77**: 102198.

Tsaples G., Papathanasiou J. & Georgiou A.C. 2022. An exploratory DEA and machine learning framework for the evaluation and analysis of sustainability composite indicators in the EU. *Mathematics* **10**(13): 2277.

Tsolas I.E., Charles V. & Gherman T. 2020. Supporting better practice benchmarking: A DEA-ANN approach to bank branch performance assessment. *Expert Systems with Applications* **160**: 113599.

Umer R., Susnjak T., Mathrani A. & Suriadi S. 2017. On predicting academic performance with process mining in learning analytics. *Journal of Research in Innovative Teaching & Learning* **10**(2): 160–176.

Xu Y., Park Y.S. & Park J.D. 2021. Measuring the response performance of U.S. States against COVID-19 Using an Integrated DEA, CART, and logistic regression approach. *Healthcare* **9**(3): 268.

Yang G.L., Fukuyama H. & Song Y.Y. 2018. Measuring the inefficiency of Chinese research universities based on a two-stage network DEA model. *Journal of Informetrics* **12**(1): 10–30.

Yavuz E. & Şahin M. 2022. Investigation of parametric, non-parametric and semiparametric methods in regression analysis. *Sakarya University Journal of Science* **26**(6): 1111–1116.

Yavuzalp N. & Bahcivan E. 2021. A structural equation modeling analysis of relationships among university students' readiness for e-learning, self-regulation skills, satisfaction, and academic achievement. *Research and Practice in Technology Enhanced Learning* **16**(1): 15.

Zarrin M. 2021. Advancing efficiency analysis using data envelopment analysis: The case of German health care and higher education sectors. PhD Thesis. Universität Augsburg.

Zhang X. & Shi W. 2019. Research about the university teaching performance evaluation under the data envelopment method. *Cognitive Systems Research* **56**: 108–115.

Zhang Z., Li Y., Li L., Li Z. & Liu S. 2019. Multiple linear regression for high efficiency video intra coding. *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1832-1836.

Zhang Z., Xiao Y. & Niu H. 2022. DEA and machine learning for performance prediction. *Mathematics* **10**(10): 1776.

Zhong K., Wang Y., Pei J., Tang S. & Han Z. 2021. Super efficiency SBM-DEA and neural network for performance evaluation. *Information Processing & Management* **58**(6): 102728.

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.

Zubir M.Z., Noor A.A., Mohd Rizal A.M., Harith A.A., Abas M.I., Zakaria Z. & Bakar A.F.A. 2024. Approach in inputs & outputs selection of data envelopment analysis (DEA) efficiency measurement in hospitals: A systematic review. *PLoS ONE* **19**(8): e0293694.

Faculty of Computer and Mathematical Sciences
University Teknologi MARA, Cawangan Perak
Kampus Tapah
34500 Tapah Road
Perak , MALAYSIA
Email: norfa122@uitm.edu.my[*]*, norha603@uitm.edu.my, soray078@uitm.edu.my*

Received: 10 April 2025
Accepted: 15 June 2025

[*]Corresponding author