Academic Performance and Perceptions of Educational Practices: A machine learning approach

Carlos Enrique Carrasco-Gutierrez[1]
Fernanda Silva dos Santos[2]
Thiago Christiano Silva[3]

Abstract

This study analyzes students' perceptions of higher education institutions (HEIs) regarding implementing unconventional educational practices and their impact on academic performance. To achieve this, we analyzed microdata from Brazil's National Student Performance Examination (ENADE), covering the three years leading up to the COVID-19 pandemic. The analysis encompassed all courses offered by HEIs, totaling a sample of $n = 765{,}923$ observations. We employed machine learning techniques to identify the most relevant attributes and econometric models to interpret the parameters. The results identified four educational practices that positively contribute to students' performance. These aspects are related to the use of information and communication technologies as a teaching strategy, the development of teamwork skills, the availability of teachers to assist students outside of class hours, and the conditions that allow students to participate in internal and external events at the institution. These findings underscore the importance of these teaching and learning strategies in Brazilian HEIs and support the development of public policies that expand their pedagogical use.

Keywords: performance, student, educational practices, ENADE, machine learning

JEL: A21, C13, I23

1. INTRODUCTION

Learning is a complex process that involves a variety of factors, with the interaction between teachers and students during the teaching and learning process being one of the key factors. Understanding students' learning methods and identifying the elements that affect their academic performance are crucial for the development of lesson planning and teaching strategies. This not only optimizes students' learning capabilities but also elevates their academic achievement.

Higher education institutions (HEIs) are increasingly compelled to reassess the educational process, acknowledge their societal responsibilities, and address challenges by proposing alternatives to traditional teaching models. Consequently, HEIs are restructuring their educational approaches by incorporating innovations that prioritize active student involvement throughout the learning process. This departure from traditional didactic models, where teachers impart knowledge passively, encourages students to actively engage in critical thinking and knowledge construction.

Active learning has emerged to enhance the teaching and learning process. This concept encompasses various classroom practices, all intended to put the student at the forefront of the

[1] Graduate Program in Economics at the Catholic University of Brasília (PPGE/UCB) and Graduate Program in Public Policy at the Catholic University of Brasília (MPPP/UCB), Brasília, DF, Brazil. E-mail address: carlosenrique@p.ucb.br
[2] Graduate Program in Economics at the Catholic University of Brasília (PPGE/UCB) at the Catholic University of Brasília (PPGE/UCB), Brasília, DF, Brazil. E-mail address: fernanda4.silva4@gmail.com
[3] Graduate Program in Economics at the Catholic University of Brasília (PPGE/UCB) at the Catholic University of Brasília (PPGE/UCB), Brasília, DF, Brazil. E-mail address: thiago.csilva@p.ucb.br

process. The focus shifts from the teacher to the students, who take responsibility for their own learning and actively participate in their educational journey. Consequently, contemporary academic initiatives prioritize a participatory approach, integrating active learning methods with traditional ones.

This study investigates the impacts of unconventional teaching and learning strategies on educational performance, as reported by students in higher education institutions and identified in the data from the National Student Performance Examination (ENADE) in Brazil. We seek to understand how these pedagogical strategies influence educational outcomes and drive students' academic and professional success. To explore this theme, we formulated ten questions related to active learning methods, derived from ENADE data:

1. Did the course program provide innovative learning experiences?
2. Did the course promote integrating theoretical knowledge with practical activities?
3. Were teachers available to assist students outside of class hours?
4. Did the course offer conditions for students to participate in internal and external events at the institution?
5. Were opportunities provided for students to engage in scientific research projects and activities that stimulate academic inquiry?
6. Were opportunities provided for students to participate in university extension programs, projects, or activities?
7. Did the institution promote cultural, recreational, and social interaction activities?
8. Did the course provide teaching assistants or tutors to assist students?
9. Did teachers use information and communication technologies (ICTs) as a teaching strategy (multimedia projectors, computer labs, virtual learning environments)?
10. Did the course provide opportunities for students to learn to work as teams?

Building on these questions, we formulated hypotheses for this study, aligned with the premise that the practices implemented by higher education institutions have positively impacted the academic performance of students. This study adopts an inductive thematic approach based on students' reports of their perceptions and concerns to investigate educational practices of higher education institutions. In the development of this study, we employed advanced machine learning techniques to identify the most relevant attributes in determining scores. Additionally, we applied econometric models for parameter estimation, incorporating the attributes previously identified as most important. The data used were obtained from the National Student Performance Examination, covering the three years before the COVID-19 pandemic and encompassing all courses (major programs) offered by higher education institutions in Brazil. The contribution of this work is multifaceted. First, provide empirical knowledge that demonstrates the effects of these ten pedagogical practices on students' performance. Second, we explore the wealth of information available in ENADE microdata to gain new insights into the primary determinants of educational performance in higher education institutions. Third, our approach goes beyond simple econometric analysis as it encompasses the intersection of two areas: machine learning in computer science and econometrics in economics. Using machine learning, we perform automatic attribute selection that can affect the variable of interest, reducing the subjectivity of the analyst in choosing the most relevant variables for the study. Machine learning does not replace analysts, but complements their analysis by identifying variables that might otherwise go unnoticed, avoiding arbitrary choices. The use of machine learning is a distinctive feature of the methodological procedure employed in this study compared to other similar studies, since it considers a broader set of information in the analysis, producing more robust results and making a significant contribution to the literature.

Our findings underscore four variables that positively influence academic performance: i) integration of information and communication technologies (ICTs); ii) development of teamwork skills; iii) availability of teachers to assist students outside of class hours; and iv) conditions allowing students to participate in internal and external events. These findings reveal the significance of these educational strategies in higher education institutions in Brazil, providing valuable support to the development of public policies aimed at optimizing the implementation of these innovative practices.

The literature on pedagogical practices involving active learning methods presents a variety of options, including self-regulated learning (SRL), problem-based learning (PBL), the incorporation of learning styles, flipped classrooms and gamification, among others. Self-regulated learning (SRL), as defined by Zimmerman and Martinez-Pons (1986), embraces a constructivist approach where the primary goal of the classroom is to stimulate student learning and encourage active engagement. Within this framework, self-regulated learners are both aware of and in control of their own learning processes. Empirical evidence supports the idea that self-regulated individuals exhibit traits such as persistence, determination, strategic thinking, and the ability to self-assess their progress, distinguishing them from those who rely on cognitive dependence or possess low levels of self-regulation (Arias, Lozano, Cabanach, & Pérez, 1999; Zimmerman & Schunk, 2001; Xu, Benson, Mudrey-Camino, & Steiner, 2010; Garcia et al., 2018).

On the other hand, problem-based learning (PBL) is an experiential learning approach that empowers students to guide their own education by fostering the development of analytical skills (Torp & Sage, 2002; Bell, 2010; Blumenfeld et al., 1991). It places responsibility on students for their work, with teachers guiding the learning process (Herreid et al., 2011; Ngeow & Kong, 2001; Karabulut, 2002). Additional benefits include the transfer of knowledge and skills to real-life situations and a deeper understanding of scientific concepts relevant to everyday life (Hoffman & Ritchie, 1997; Duggan & Gott, 2002; Ketpichainarong et al., 2010). PBL involves students working in smaller groups to discuss challenges and find solutions, contributing to improved perception of learning and enhanced performance in assessments (Rosing, 1997; Rideout, 2001; Rideout & Carpio, 2001).

When it comes to learning styles, students employ diverse approaches, influenced by their unique personal characteristics, competencies and abilities. Various theories on learning styles, including Kolb's (1976) experiential learning cycle, highlight the importance of considering attitudes and feelings during the learning process. Incorporating learning styles into education enables teachers to employ suitable strategies tailored to individual characteristics, ensuring effective addressing of students' educational needs (Engels & Gara, 2010; Samarakoon et al., 2013; Nuzhat et al., 2013; Boström & Hallin, 2013; Boström, 2011; Smith, 2010).

Among other pedagogical practices of active learning method are flipped classrooms[4], initially popularized in secondary education in the United States (Lage & Platt, 2000; Bergmann & Sams, 2009) and the gamification process. Gamification involves employing game techniques such as challenges and rewards, with the primary aim of increasing engagement and stimulating users' interest. In education, gamification is becoming an important tool making classes more captivating and productive for both students and teachers, facilitating positive outcomes in the teaching-learning process.

Finally, Chickering and Gamson (1987, 1991) presented the "seven principles of good practice in undergraduate education". Based on decades of research on the educational experience in higher education, the authors observed declining student performance, student passivity, and poor teaching methods, among other factors. These principles encompass (a) encouraging contact between students and faculty, (b) developing reciprocity and cooperation among students, (c) encouraging active learning, (d) providing prompt feedback, (e) emphasizing time on tasks, (f) communicating

[4] Or called *inverted classrooms*.

high expectations, and (g) respecting diverse talents and ways of learning. Several studies have been developed based on these foundations, such as those that have evaluated and defined effective teaching in traditional classroom environments (Chickering & Gamson, 1987; Chickering & Ehrmann, 1996; McFadden, 2006; McCabe & Meuter, 2011; Gaižiūnienė, 2018).

Our paper connects with the empirical literature aimed at enhancing the understanding of factors influencing educational outcomes in the Brazilian higher education landscape. It establishes connections with several other studies in the field of higher education, particularly those conducted in the context of Brazil. For instance, Machado et al. (2024) explored the impact of accountability scores on Brazilian higher education. In complementing this study, their work emphasized the underlying effects on academic performance associated with specific educational practices. These practices encompass the integration of information and communication technologies, the cultivation of teamwork skills, teacher availability beyond regular class hours, and favorable course conditions. Vieira and Arends-Kuenning (2019) investigated the effects of affirmative action policies on college admission in Brazil. While their focus was on the enrollment of specific groups in higher education, our study delves into the impact of unconventional educational practices on academic performance, shedding light on potential factors influencing the success of affirmative action initiatives. Tavares (2015) examined the causal impacts of school management practices on educational outcomes. While both Tavares's study and ours highlight the importance of effective educational strategies, our paper concentrates on unconventional teaching practices in higher education across Brazil, whereas Tavares (2015) narrowed the discussion to the context of public schools in São Paulo.

Among other relevant works in the Brazilian context, we emphasize the significant contributions of Cornachione et al. (2010) and Malerva and Escorza (2018). The former authors investigated the factors influencing the academic performance of undergraduate students in accounting in four Brazilian universities, while the latter researchers conducted a study identifying the effects of learning strategies on the academic performance of medical students. Additionally, we highlight the contributions of Alencar and Fleith (2004), Santana and Araújo (2010), Oliveira et al. (2016), Signori et al. (2018), Lima et al. (2016), Heringer et al. (2019), Biffi et al. (2020), Corrêa et al. (2020), Cualheta et al. (2021), and Riccomini et al. (2021).

Following this introduction, our paper is structured as follows: Section 2 presents the methodology, including our description, database, and machine learning models. Section 3 discusses the obtained results, and Section 4 presents our conclusions.

2. METHODOLOGY

2.1 Regularized Regression Models

In this study, we applied the elastic net regularization technique in the context of supervised machine learning to identify the most informative and essential attributes for the predictive model. The technique also aims to eliminate variables that may be redundant, irrelevant, or detrimental to the model's accuracy. Additionally, this approach is especially valuable in situations where input variables exhibit significant correlations, potentially leading to multicollinearity. Elastic net combines two regularization techniques, lasso (L_1) and ridge (L_2), in an optimal way. Lasso regularization has the property of reducing the coefficients of less relevant regression variables to zero, retaining only those that are essential for predicting the target variable. On the other hand, ridge regularization penalizes the magnitude of larger coefficients, leading to their reduction but not complete elimination. We applied regularization to the following linear regression model:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_i^j + \varepsilon_i \quad (1)$$

In accordance with the elastic net procedure, we chose coefficients $\beta = [\beta_0, \beta_1, \dots, \beta_p]'$ that minimized the following loss function $L(\beta)$:

$$L(\beta) = \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_i^j)^2 + \lambda \left(\alpha \sum_{i=1}^{p} |\beta_j| + \frac{(1-\alpha)}{2} \sum_{i=1}^{p} {\beta_j}^2 \right) \tag{2}$$

In the context of this study, y_i represents the observed grade of the i-th student, p is the number of attributes in the model, n is the number of observations in the sample, and α and λ are hyperparameters. The first term of the loss function represents the mean squared error between the predicted values and the actual grade values. The second component is the regularization term, which involves two summations associated with L_1 and L_2 norms, respectively. The hyperparameter λ controls the importance of the regularization term; the higher the value of λ, the more intense the regularization is. When $\lambda = 0$, the loss function reduces to conventional linear regression. When $\lambda = 1$, the loss function contains elastic net regularization, which is an equal combination of lasso (L_1) and ridge (L_2) penalties. The hyperparameter α controls the intensity of the penalties applied to the regression coefficients, resulting in a convex combination of L_1 and L_2. When $\alpha = 0$, the application of elastic net becomes equivalent to ridge regularization, using only the L_2 penalty. When $\alpha = 1$, the elastic net technique becomes equivalent to lasso regularization, using only the L_1 penalty. Intermediate values of α, in the range of $0 < \alpha < 1$, allow for a combination of both penalties, enabling more flexible fitting. The appropriate selection of α and λ values is important for proper model fitting, and this choice will be made through cross-validation techniques, exploring different combinations, and selecting those that result in the best performance of previously unused data.

In summary, careful feature selection is crucial for several reasons. First, it can help improve computational efficiency by reducing the dimensionality of the dataset. Additionally, eliminating irrelevant or redundant features makes it possible to avoid overfitting, a situation where the model overly fits the training data. Thus, the resulting model will exhibit high predictive accuracy and the ability to generalize to unexplored datasets.

2.2 Econometric Model

We used the following empirical specification to analyze the impact of non-traditional teaching practices on educational performance:

$$y_{i,t} = \alpha_i + \alpha_{q(i),t} + X_{i,t}^{j}\beta' + Z\gamma' + \varepsilon_{i,t} \tag{3}$$

In this equation, i represents the student, $q(i)$ is a group to which the student belongs (such as school, city, state, etc.), and t is the time dimension. The term α_i captures individual characteristics that do not change over time, known as the individual fixed effect. Meanwhile, $\alpha_{q(i),t}$ captures fixed effects that allow for comparing similar groups in different aspects, represented by $q(i)$. The variable X_{it}^{j} considers the dimensions of interest related to educational practice j that we want to study. The variable Z is a vector of variables commonly used in empirical studies to explain academic performance, such as age, study hours and gender, etc.

The fixed effects α_i and $\alpha_{\ q(i),t}$ in the estimated model will be selected by the machine learning algorithm to capture the heterogeneity of students and different aspects represented by $q(i)$. We aim to determine whether the coefficients β_j in the vector β' are positive and statistically different from zero ($\beta_j > 0$). The dependent variable y_{it} in (3) represents the student's grade in year t, and $\varepsilon_{i,t}$ is the error term. The coefficient of interest, β_j, can be interpreted as the elasticity of the effect of educational practices on the grade.

2.3 Data

We used the microdata from ENADE[5] (National Student Performance Exam) for the pre-COVID-19 pandemic period, covering 2016 to 2018. This database encompasses all courses (majors) offered by higher education institutions in the in-person modality. ENADE plays a crucial role as an assessment tool, focusing on students in their final year of undergraduate programs. Its purpose extends beyond measuring these students' academic performance in relation to the established curriculum, also including evaluating the development of essential competencies and skills for their comprehensive education and professional training.

Tables 1 and 4 provide a description of the variables used in this study. Table 1 offers a detailed explanation of the variables associated with the ten core dimensions of our analysis, along with the dependent variable. Table 4 contains a comprehensive description of all attributes considered in this investigation.

Table 1 - Description of Variables Related to Educational Practices

Variables	Types of variables	Description
Dependent		
final_score	Integer	Overall raw score - Weighted average of the objective (60%) and discursive (40%) components in the general formation (value from 0 to 100)
Explanatory variables of interest		
team_work	ordinal (6 segments)	In the course, did you have the opportunity to learn how to work in a team? (strongly disagree = 1, ..., strongly agree = 6)
ext_projects	ordinal (6 segments)	Were opportunities offered for students to participate in university extension programs, projects, or activities? (strongly disagree = 1, ..., strongly agree = 6)
sci_init_prog	ordinal (6 segments)	Were opportunities provided for students to participate in undergraduate research projects and activities encouraging academic investigation? (strongly disagree = 1, ..., strongly agree = 6)
course_events	ordinal (6 segments)	Did the course provide conditions for students to participate in the institution's internal and/or external events? (strongly disagree = 1, ..., strongly agree = 6)
practical_activ	ordinal (6 segments)	Did the course facilitate the integration of theoretical knowledge with practical activities? (strongly disagree = 1, ..., strongly agree = 6)
outside_class	ordinal (6 segments)	Were the teachers available to assist students outside of class hours? (strongly disagree = 1, ..., strongly agree = 6)
ICT	ordinal (6 segments)	Did the teachers use information and communication technologies (ICTs) as a teaching strategy? (strongly disagree = 1, ..., strongly agree = 6)
teaching_assistant	ordinal (6 segments)	Did the course provide teaching assistants and/or tutors to assist students? (strongly disagree = 1, ..., strongly agree = 6)
culture_activ	ordinal (6 segments)	Did the institution promote cultural, recreational, and social interaction activities? (strongly disagree = 1, ..., strongly agree = 6)
Innov_learning	ordinal (6 segments)	Did the course provide innovative learning experiences? (strongly disagree = 1, ..., strongly agree = 6)

2.4 Descriptive Statistics

Tables 2 and 3 provide a comprehensive overview of the descriptive statistics of our data. From Table 2, we observe that the performance indicator, represented by the final grade, spans a wide

[5] The quality indicators for higher education are determined based on the results of ENADE and the responses gathered from the student questionnaire. These indicators assign ratings to educational institutions, ranging from 1 to 5, with 5 representing the highest score.

range, from 0 to 99.4, with an approximate sample mean of 49. Additionally, we found that the average household income reached R$ 4710 (Brazilian reais). Our sample consists of individuals with ages ranging from 17 to 89 years, with an average age of around 27 years. Regarding the average daily time dedicated to studies, we identified a mean of 5.4 hours, while students, on average, read approximately 3 books from the course bibliography throughout the year. Finally, the average composition of households in the sample is approximately 2.5 individuals.

Table 3 presents the distribution of final grades among students, broken down by gender and region. It is noteworthy that the Southeast region concentrated the majority of individuals in this sample, accounting for 19.7% of men and 26.4% of women, while the North region had the lowest participation, with 2.8% of men and 4.2% of women. Regarding academic performance, the average grades were highest in the South region, both for men, with an average of 52.3, and women, recording an average of 49.4. On the other hand, the lowest average grades were identified in the North region, both for men, with an average of 48.4, and for women, with an average of 45.1. This disparity may be linked to lower income levels in the population of the North region compared to the South region, which in turn can negatively impact the educational performance of students.

Table 2 - Description of the numerical variables

Statistic	N	Mean	St. Dev.	Min	Pc(25)	Pc(75)	Max
final_score	765,923	49.03	17.79	0	36.8	61.9	99.4
family_income	765,923	4,710	5,430	703	2,108	4,919	28,110
age	765,923	27.42	7.06	17	23	30	89
study_hours	765,923	5.41	4.50	0	20	5.50	15
num_people	765,923	2.46	1.66	0	1	3	8
books_read	765,923	2.94	2.98	0	1	3	10

Table 3 - Description of final score by gender and region

Gender	Region	Obs	Freq	Mean	Median	St. Dev.	Pc(25)	Pc(75)	Min	Max
Male	North	21,717	2.8%	48.4	48.9	17.6	36.5	61.2	0	98.4
	Northeast	66,820	8.7%	50.2	51	18	37.5	63.5	0	98.8
	Southeast	150,758	19.7%	51.2	52	18	38.5	64.3	0	99.2
	South	54,356	7.1%	52.3	53	17.8	39.9	65.4	0	98.8
	Midwest	26,680	3.5%	49.5	49.9	18.3	37.1	62.8	0	99.2
Female	North	31,848	4.2%	45.1	45	17	33	57	0	98
	Northeast	101,784	13.3%	46.2	46.2	17.4	34	58.7	0	98.6
	Southeast	201,850	26.4%	48.6	48.7	17.5	36.4	61.2	0	99.4
	South	71,649	9.4%	49.4	49.7	17.5	37.5	62.1	0	98.8
	Midwest	38,461	5.0%	46.4	46.4	17.7	34.1	59	0	98.6
	Brazil	765,923	100%	49.03	49.3	17.79	36.8	61.9	0	99.4

Figure 1 displays the students' grades categorized into various ranges of average household income and grouped by regions[6]. It is evident that, across all regions of Brazil, as students' income increases,

[6] The minimum monthly wage in Brazil is adjusted annually through a presidential decree, leading to variations in the values over time. In the years 2016, 2017, and 2018, the established amounts were, respectively, R$ 880.00, R$ 937.00, and R$ 954.00.

their score also shows a corresponding increase. This pattern of behavior was consistent in all the analyzed regions.

Figure 2 highlights the distribution of students' grades in relation to regions, categorizing them according to the type of educational institution where they completed their studies before entering higher education (public, private, or other). It is noteworthy that in each region, the grades of students from private schools were higher than those of students from public schools. This trend can be understood by considering that students who attend private schools in Brazil generally have a more robust educational foundation compared to their peers from public schools, a fact that is also reflected in the final grades obtained by the students.

Figure 1 – Distribution of final grades by income and region

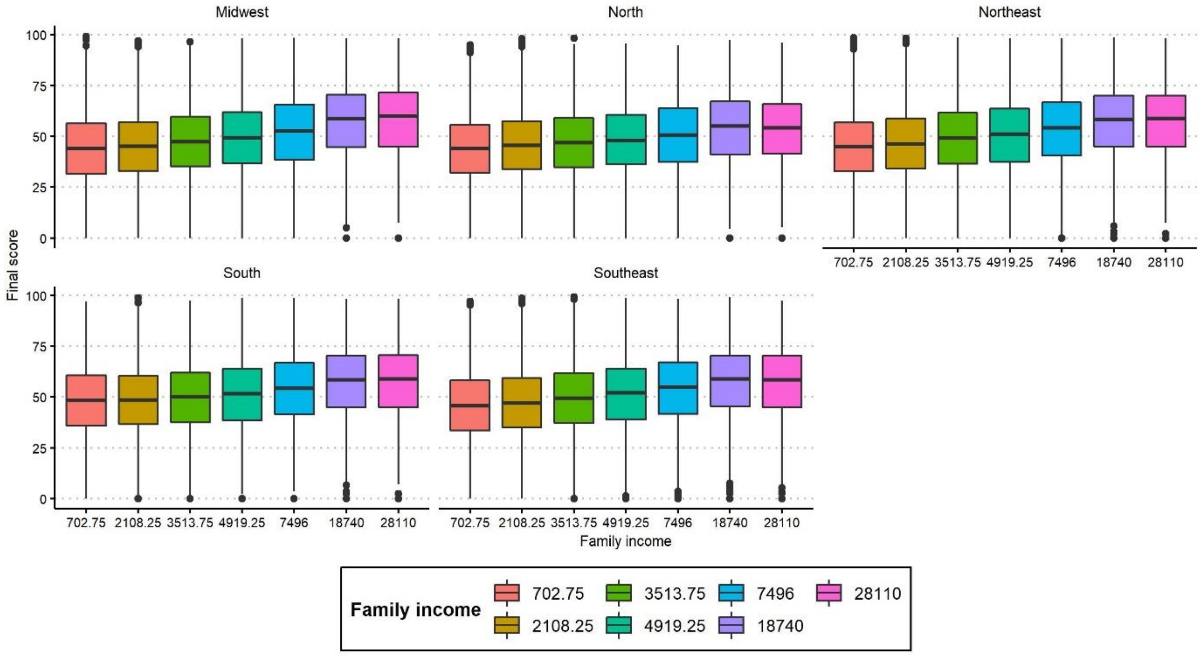
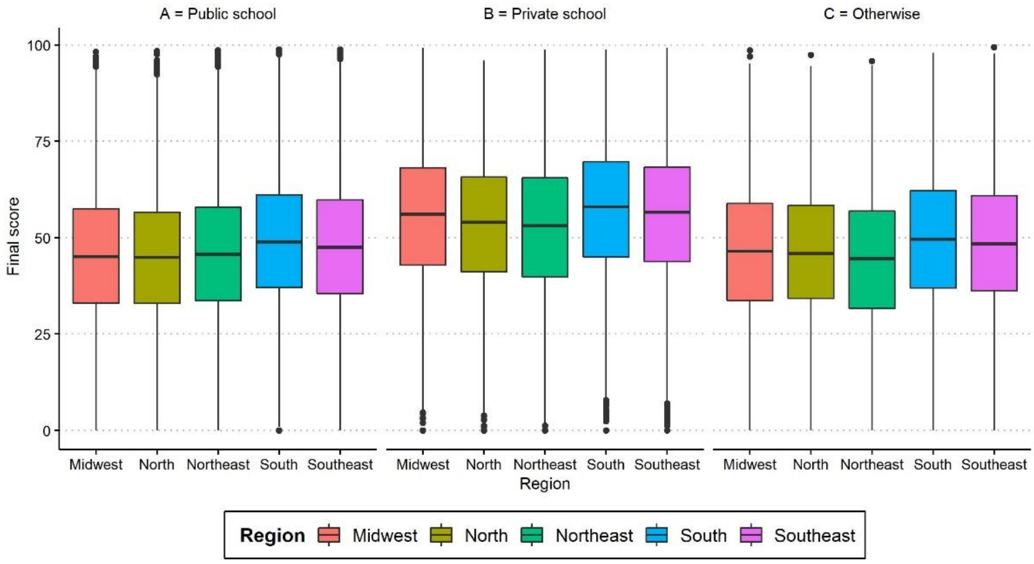


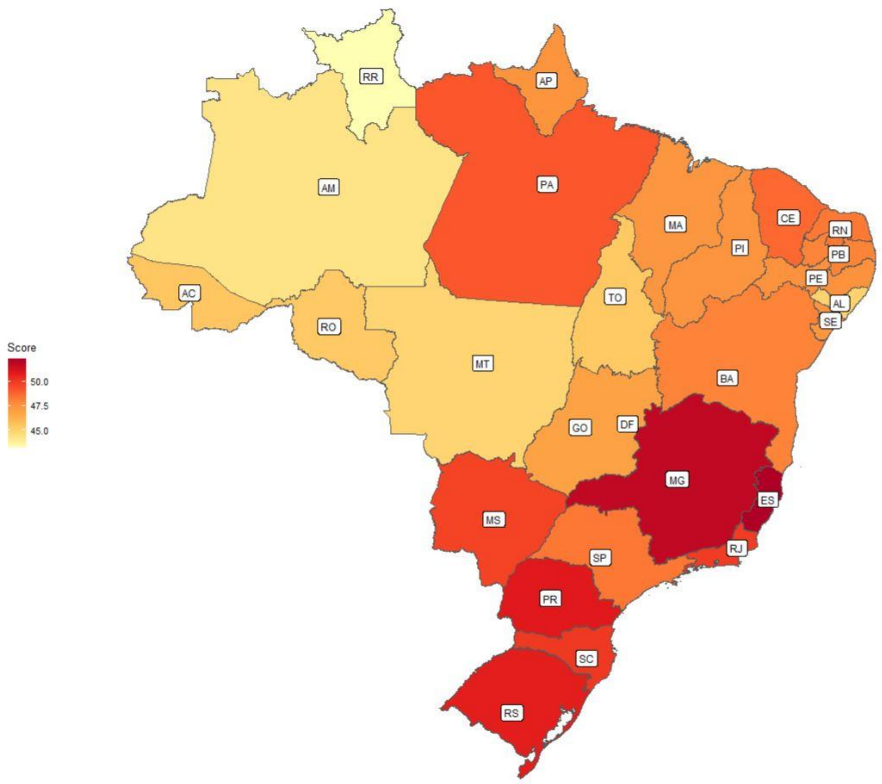
Figure 2 – Distribution of students' final score, categorized by geographic region and educational institution where they completed their studies before entering higher education.



Note: Course operating region codes are designated as follows: 1 = North, 2 = Northeast, 3 = Southeast, 4 = South, 5 = Midwest. Categories include A = all in public school, B = all in private school, C = a combination of all abroad, mostly in public school, mostly in private school, and partly in Brazil and partly abroad.

Figure 3 displays the heat map representing the average grades of students by state. It is noteworthy that the states of Minas Gerais (MG) and Espírito Santo (ES), located in the Southeast region, along with Paraná and Rio Grande do Sul, in the South region, exhibited the highest average grades. Coincidentally, these regions are also the most economically developed in Brazil, with the highest gross domestic product (GDP). In contrast, the state of Roraima (RR), located in the North region of the country, recorded the lowest average, standing out as the least economically favored region in Brazil.

Figure 3 – Average final scores by state



3. RESULTS

3.1 Maching learning results

After preprocessing the data from the entire dataset provided by the ENADE, we obtained a set of 40 variables, as presented in Table 4. Next, we employed the elastic net regularization method to identify attributes with the highest predictive capability to explain student grades. To conduct this selection, we divided the dataset into two distinct sets: the training set, comprising 70% of the observations, and the test set, comprising the remaining 30%. The application of the Bernoulli sampling procedure ensured randomness in the allocation of these sets. To obtain a more robust and accurate assessment of the model's performance on unseen data, we employed the k-fold cross-validation method. In this procedure, we divided the dataset into $k = 5$ equally sized folds[7], and each fold was further subdivided into $k = 5$ additional subsets. Consequently, the $k - 1 = 4$ new subsets formed in this way were allocated for model training, while the remaining subset was designated as the validation set. This sequence of steps was iterated $k = 5$ times, ensuring that each

[7] We chose the value of k=5 due to the sample size.

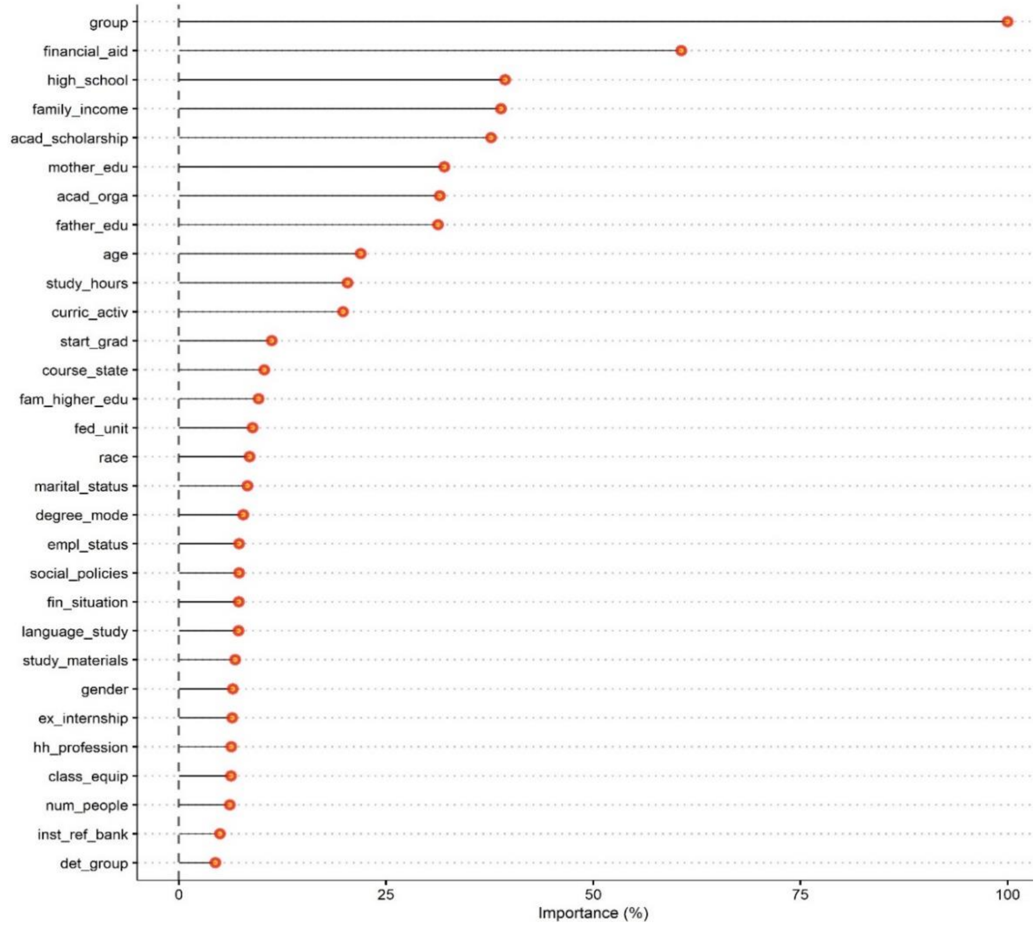
partition acted as the validation set exactly once. Finally, evaluation metrics were computed for each iteration, resulting in a comprehensive measure of the overall model performance.

The sampling structure adopted by ENADE allows for the assumption that the data are independent and identically distributed, validating the use of k-fold cross-validation as an appropriate model selection procedure. Additionally, we applied Z-score standardization to all data points. This standardization was carried out using only the predefined values extracted from the training data to prevent any information leakage from the test set. The results of the loss function optimization showed adherence statistical values of RMSE = 0.91 and MAE = 0.73, providing the hyperparameter values $\alpha = 0.987$ and $\lambda = 0.016$.

Visually, Figure 4 illustrates the results of applying the elastic net regularization technique, highlighting the importance of the initial 30 variables in determining the overall grade. We observed that the four most relevant variables for grades corresponded to the institutional administrative category (group[8]), the type of scholarship received to partially or fully cover tuition fees (financial_aid), the type of high school attended (high_school), and household income (family_income).

Table 4 provides the complete ranking of variables by importance. Additionally, it is worth noting that traditional variables, such as "gender" and "race", commonly used in regression studies, did not show great relevance in determining the dependent variable (final_score). This further strengthens the previous analysis of the machine learning method in our study.

Figure 4 – Feature selection results using an elastic net procedure with L_1 and L_2 regularization.



Note: Results of feature selection using an elastic net procedure with L_1 and L_2 regularization. Coefficients are normalized in terms of the most important attribute "group". Average ranking of the most important attributes. The lower the ranking, the more important the attribute. This table presents the selection of the top 30 most important attributes chosen by the elastic net procedure.

[8] Administrative category of the HEI (higher education institution). Refer to Table 7 in the appendix.

Table 4 – Description of the variables used in variable selection by machine learning and sorted by order of importance

Fixed Effects	Variables (ordered by ML)	Type of variable	Questions/ Description
1	group	Nominal (94 segments)	ENADE course framing area code (all courses - 2016, 2017, and 2018 triennial). See Table 7 in the appendix for the complete description.
2	financial_aid	Character (11 segments)	*What type of scholarship or course financing did you receive to cover all or most of the tuition fees? (*A = None, as my course is free, B = None, even though my course isn't free, C = Full ProUni scholarship, D = Partial ProUni scholarship only, E = FIES loan only, F = Partial ProUni scholarship and FIES loan, G = Scholarship offered by state, district, or municipal government, H = Scholarship offered by the institution itself, I = Scholarship offered by another organization (company, NGO, other), J = Financing offered by the institution itself, K = Bank financing).
3	high_school	Character (6 segments)	*What type of school did you attend for high school?* (A = All in public school, B = All in private school, C = All abroad, D = Mostly in public school, E = Mostly in private school, F = Partly in Brazil and partly abroad).
	family_income	Integer	*What is your total family income, including your own earnings?*
4	acad_scholarship	Character (6 segments)	*Throughout your academic journey, have you received any type of academic scholarship?* In case there is more than one option, select only the scholarship with the longest duration. (A = None, B = Undergraduate research scholarship, C = Extension scholarship, D = Teaching assistantship scholarship, E = PET scholarship, F = Other type of academic scholarship).
5	mother_edu	Character (5 segments)	*Up to which level of education did your mother complete? (A = None, B = Elementary School: 1st to 5th grade, C = Elementary School: 6th to 9th grade, D = High School, E = Higher Education - Undergraduate, F = Postgraduate)*
6	acad_orga	Character (6 segments)	*Code of the academic organization of the Higher Education Institution (HEI). (A = Federal Center for Technological Education, B = University Center, C = College, D = Federal Institute of Education, Science and Technology, E = University)*
7	father_edu	Caracter (6 segmentos)	*Up to which level of education did your father complete? (A = None, B = Elementary School: 1st to 5th grade, C = Elementary School: 6th to 9th grade, D = High School, E = Higher Education - Undergraduate, F = Postgraduate)*
	age	Numeric	*Age of the applicant on November 26, 2017.*
	study_hours	Numeric	*How many hours per week, approximately, did you dedicate to studying, excluding class hours?*
8	curric_activ	Character (6 segments)	*During your undergraduate course, did you participate in programs and/or curricular activities abroad? (A = I did not participate, B = Yes, Science Without Borders Program, C = Yes, exchange program funded by the Federal Government (Marca; Brafitec; PLI; other), D = Yes, exchange program funded by the State Government, E = Yes, exchange program offered by my institution, F = Yes, other non-institutional exchange)*
	start_grad	Numeric	*Year of start of undergraduate studies.*
9	course_state	Caracter (27 segmentos)	*Code of the state where the course is located (11 = Rondônia, ... , 53 = Federal District).*

10	fam_higher_edu	Binary (2 segments)	*Has anyone in your family completed a higher education course? (A = Yes, B = No)*
11	fed_unit	Nominal (6 segments)	*In which state (UF) did you complete high school? (11 = Rondônia, ..., 53 = Federal District)*
12	race	Character (6 segments)	*What is your color or race? (A = White, B = Black, C = Asian, D = Mixed race, E = Indigenous, F = I do not wish to declare)*
13	marital_status	Nominal (5 segments)	*What is your marital status? (A = Single, B = Married, C = Legally separated/divorced, D = Widowed, E = Other)*
14	degree_mode	Character (5 segments)	*Which type of high school did you complete? (A = Traditional high school, B = Technical vocational (electronics, accounting, agricultural, other), C = Teacher training vocational, D = Adult and Youth Education (EJA) and/or GED, E = Other type)*
15	empl_status	Nominal (5 segments)	*Which of the following options best describes your work situation (excluding internships or scholarships)? (A = I am not working, B = I work occasionally, C = I work up to 20 hours per week, D = I work 21 to 39 hours per week, E = I work 40 hours per week or more)*
16	social_policies	Character (6 segments)	Did your admission to the undergraduate course occur through affirmative action or social inclusion policies? (A = No, B = Yes, based on ethnic-racial criteria, C = Yes, based on income criteria, D = Yes, because you studied at a public school or private school with a scholarship, E = Yes, through a system that combines two or more of the previous criteria, F = Yes, through a different system than the previous ones)
17	fin_situation	Character (6 segments)	*Which of the following options best describes your financial situation (including scholarships)? (A = I have no income and my expenses are funded by government programs, B = I have no income and my expenses are funded by my family or other people, C = I have income but receive help from family or others to fund my expenses, D = I have income and do not need help to fund my expenses, E = I have income and contribute to family support, F = I am the main provider for my family)?*
18	language_study	Character (5 segments)	Did you have the opportunity for foreign language learning at the Institution? (A = Yes, only in face-to-face mode, B = Yes, only in semi-face-to-face mode, C = Yes, part in face-to-face mode and part in semi-face-to-face mode, D = Yes, in distance learning mode, E = No)
19	study_materials	Character (6 segments)	*The equipment and materials available for practical classes were adequate for the number of students. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
20	gender	Character (2 segments)	*Gender (M = Male, F = Female)*
21	ex_internship	Nominal (6 segmentos)	*Opportunities for students to participate in exchanges and/or internships within the country were offered. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
22	hh_profession	Nominal (6 segmentos)	*The teachers demonstrated mastery of the content covered in the subjects. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
23	class_equip	Nominal (6 segmentos)	*The environments and equipment provided for practical classes were suitable for the course. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
	num_people	Inteiro	*How many people from your family live with you? Consider your parents, siblings, spouse, children, and other relatives who live in the same house with you.*
24	inst_ref_bank	Nominal (6 segmentos)	*The institution provided a cafeteria, snack bar, and bathrooms in suitable conditions that met the needs of its users. (Strongly Disagree = 1, ..., Strongly Agree = 6)*

25	det_group	Caracter (11 segmentos)	*Were any of the following groups instrumental in helping you overcome challenges during your college education and completing it? (A = I did not face difficulties, B = I did not receive support to overcome difficulties, C = Parents, D = Grandparents, E = Siblings, cousins, or uncles, F = Religious leader or representative, G = Coursemates or friends, H = Course professors, I = Student support service professionals from the institution, J = Work colleagues, K = Other group)*
26	course_school	Caracter (9 segmentos)	*What was the main reason for you to choose this course? (A = Job market entry, B = Family influence, C = Professional advancement, D = Social prestige, E = Vocational calling, F = Offered in distance learning mode, G = Low competition for admission, H = Other reason)*
27	infra_condition	Nominal (6 segments)	*The classroom infrastructure conditions were suitable. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
28	course_region	Nominal (5 segments)	*Course operating region code (1 = North, 2 = Northeast, 3 = Southeast, 4 = South, 5 = Midwest)*
29	library	Nominal (6 segments)	*The library provided the necessary bibliographic references for the students. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
30	num_employees	Caracter (6 segmentos)	*The institution had a sufficient number of staff members for administrative and academic support. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
31	aid_type	Nominal (6 segments)	*Throughout your academic journey, did you receive any form of financial assistance for living expenses? In case of having more than one option, mark only the type of assistance with the longest duration. (A = None, B = Housing allowance, C = Food allowance, D = Housing and food allowance, E = Living expenses assistance, F = Other type of assistance)*
	books_read	Integer	*Excluding the books listed in your course bibliography, how many books have you read this year?*
32	stud_ded_org	Nominal (6 segments)	*The course required organization and consistent dedication to studies. (Strongly Disagree = 1, ..., Strongly Agree = 6)*
33	birthplace	Caracter (3 segmentos)	*What is your nationality? (A = Brazilian, B = Naturalized Brazilian, C = Foreign)*

Note: Self-compiled based on ENADE questionnaires for the years 2016, 2017, and 2018. Variables marked with (*) have been transformed into quantitative variables. The variable "family_income" has been transformed into a numerical variable considering the average values per income bracket. For each income bracket in the ENADE questionnaire, the sample mean was calculated. The column corresponding to the "order" variable indicates the ranking order determined by feature selection using the elastic net approach.

3.2 Econometric results

Given the set of variables identified in the previous analysis by Machine Learning, the empirical specification to be estimated is configured as follows:

$$final\_score_{i,t} = \alpha_i + \alpha_{g(i),t} + X_{i,t}^{j}\beta' + \gamma_1 family\_income_i + \gamma_2 age_i + \gamma_3 study\_hours_i + \gamma_4 books\_read_i + \gamma_5 gender_i + \gamma_6 num\_people_i + \gamma_7 start\_grad_i + \varepsilon_{i,t} \quad (3)$$

Where:

$final\_score_{it}$: the overall score on the ENADE exam for student *i* in year *t*.
α_i: individual fixed effects chosen by elastic net regularization*;*
$\alpha_{q(i),t}$: fixed effects that ensure the comparison of similar groups in different aspects, denoted by $q(i)$ and chosen by elastic net regularization.*;*
X_{it}^{j}: considers the dimensions of interest related to educational practices *j* that we wish to study, which belong to the set{team_work, ext_projects, sci_init_prog, course_events, practical_activ, outside_class, ICT, teaching_assistant, culture_activ, innov_learning}
β: a vector containing the parameters of interest $\beta = [\beta_1, \beta_2, \dots, \beta_{10}]'$;
$family\_income_i$: the total income of the student i's household;

age_i: the age of the student;
$study\_hours_i$: the number of hours student i dedicates to studying daily;
$books\_read_i$: the number of books read by student i;
num_people_i: the number of people living with the student;
$start\_grad$: represents the year of the start of the course.

We included quantitative control variables in the regression analysis, specifically: $family\_income$, age, $study\_hours$, $books\_read$, num_people and $start\_grad$. Since thehousehold income and study hours variables were among the top 10 variables selected by the elastic net method, we extended the analysis to order 12 to complete the set of the 10 most relevant variables. Table 5 displays the results of the model estimates as described in equation (3), considering the top 10 variables selected by the elastic net regularization method. These variables include "group", "financial_aid", "high_school", "acad_scholarship", "mother_edu", "acad_orga", "father_edu", "curric_activ", "course_state" and "fam_higher_edu" (see the complete description of these variables in Table 4).
The students' grades reflected the results of various assessments conducted in each course and different years. Therefore, before conducting econometric estimations to ensure comparability of results, we proceeded with the standardization of the quantitative variables that would be used in the regression. This standardization involved centering each variable around its sample mean and dividing it by the standard deviation of the corresponding course. The results of the estimation of model (3) are detailed in Table 5.

Table 5 – Econometric Results

Dependent variable	**final_score**		
Model	(1)	(2)	(3)
Independent variables			
innov_learning	-0.046*** (0.003)	-0.046*** (0.005)	-0.046*** (0.004)
practical_activ	-0.005 (0.004)	-0.005 (0.005)	-0.005 (0.004)
outside_class	0.018*** (0.003)	0.018*** (0.005)	0.018*** (0.002)
course_events	0.039*** (0.003)	0.039*** (0.006)	0.039*** (0.005)
sci_init_prog	-0.038*** (0.004)	-0.038*** (0.003)	-0.038*** (0.005)
ext_projects	0.006 (0.004)	0.006** (0.003)	0.006 (0.007)
teaching assistant	-0.032*** (0.003)	-0.032*** (0.004)	-0.032*** (0.003)
ICT	0.040*** (0.003)	0.040*** (0.009)	0.040*** (0.003)
team_work	0.021*** (0.003)	0.021*** (0.005)	0.021*** (0.003)
culture_activ	0.003 (0.003)	0.003 (0.004)	0.003 (0.004)
family_income	0.081*** (0.003)	0.081*** (0.008)	0.081*** (0.003)
Age	-0.050*** (0.002)	-0.050*** (0.009)	-0.050*** (0.005)
study_hours	0.050*** (0.002)	0.050*** (0.006)	0.050*** (0.002)
books_read	0.028*** (0.002)	0.028*** (0.004)	0.028*** (0.003)
gender (1=if male, 0=female)	0.111***	0.111***	0.111***

	(0.005)	(0.024)	(0.009)
num_people	-0.032***	-0.032***	-0.032***
	(0.002)	(0.003)	(0.003)
start_grad	0.062***	0.062***	0.062***
	(0.003)	(0.004)	(0.005)
Fixed effect			
Top 10 variables	Yes	Yes	Yes
Fit statistics			
Observations	765,899	765,899	765,899
R^2	0.595	0.595	0.595
Adjusted R^2	0.154	0.154	0.154
Robust error (clustering)	student	group	course_state

Note: Significance levels used: ***1%; **5%; *10%. The 10 attributes selected by the machine learning algorithm using the feature selection procedure and used as fixed effects are: “group”, “financial_aid”, “high_school”, “acad_scholarship”, “mother_edu”, “acad_orga”, “father_edu”, “curric_activ”, “course_state” and “household_higher_edu”.

We used the method of robust ordinary least squares (OLS) to address heteroscedasticity issues by clusters (student, group, state), as presented in columns 1, 2, and 3. In all three estimated specifications, we observed that household income, study hours, and the year the student started his/her undergraduate studies have a significant positive impact on the grade, with a 5% level of significance. This means that students from financially better-off households tend to have higher grades. Additionally, the more hours a student dedicates to studying, the higher their grade. Lastly, the less time a student remains enrolled in higher education, generally the better their grade tends to be.

We also found that age has a negative effect on the grade, indicating that older students tend to have lower grades. On the other hand, the binary variable “gender” (1 = male; 0 = female) positively impacts students’ grades, suggesting that males have higher grades than females. Finally, the variable “num_people”, representing the number of people living with the student, has a negative effect on the grade. This may be explained by factors such as limited physical space or an inadequate environment that could potentially hinder the academic performance of the student.

Regarding the variables of interest in this study, we observed, according to students' perceptions, that four attributes have a statistically significant positive effect on grades, with 5% significance level. These attributes include the use of information and communication technologies as a teaching strategy (ICT), the ability of the student to work in a team (team_work), the availability of teachers to assist students outside of class hours (outside_class), and the conditions of the course that allow students to participate in events both internal and external to the institution (course_events).

On the other hand, we did not find evidence of the effects of the variables “practical_activ”, “ext_projects”, and “culture_activ” on the grade. Finally, we also found that the variables “innov_learning”, “sci_init_prog”, and “teaching_assistant” negatively affected the student's grade. In other words, as practices related to these variables are implemented, the student's grade decreases. This shows that not all non-traditional practices can bring benefits to educational performance. A plausible explanation is that the way these practices are being implemented is not delivering the expected effects, and instead of being beneficial, they end up being ineffective.

3.3 Robustness analysis

To demonstrate the robustness of the results, we re-estimated the models, incorporating two additional sets of variables, identified through machine learning (ML). Table 6 presents the results of the estimations, where the first three columns display the results considering the top five variables selected by ML, while the last three columns correspond to the results considering the top 15 variables identified by ML. Overall, the results remain robust to those presented in Table 5. In particular, we note that the variable “ext_projects” becomes statistically significant at the 5% level when considering the top five variables chosen by ML (Models 1, 2 and 3) as well as in the case of the top 15 variables (Model 5). Conversely, in other instances, this variable remains statistically insignificant.

Table 6 - *Robustness Analysis*

Dependent variable	**final_score**					
Model	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables						
innov_learning	-0.048***	-0.048***	-0.048***	-0.045***	-0.045***	-0.045***
	(0.002)	(0.004)	(0.003)	(0.010)	(0.007)	(0.006)
practical_activ	-0.006***	-0.006	-0.006*	-0.007	-0.007	-0.007*
	(0.002)	(0.004)	(0.003)	(0.011)	(0.007)	(0.004)
outside_class	0.015***	0.015***	0.015***	0.018*	0.018***	0.018***
	(0.002)	(0.003)	(0.002)	(0.010)	(0.005)	(0.004)
course_events	0.038***	0.038***	0.038***	0.037***	0.037***	0.037***
	(0.002)	(0.005)	(0.005)	(0.011)	(0.009)	(0.007)
sci_init_prog	-0.035***	-0.035***	-0.035***	-0.037***	-0.037***	-0.037***
	(0.002)	(0.003)	(0.004)	(0.012)	(0.007)	(0.007)
ext_projects	0.013***	0.013***	0.013**	0.012	0.012**	0.012
	(0.002)	(0.002)	(0.006)	(0.012)	(0.006)	(0.011)
teaching assistant	-0.030***	-0.030***	-0.030***	-0.032***	-0.032***	-0.032***
	(0.002)	(0.004)	(0.004)	(0.010)	(0.006)	(0.005)
ICT	0.042***	0.042***	0.042***	0.040***	0.040***	0.040***
	(0.002)	(0.008)	(0.003)	(0.010)	(0.013)	(0.006)
team_work	0.024***	0.024***	0.024***	0.023**	0.023***	0.023***
	(0.002)	(0.005)	(0.002)	(0.010)	(0.007)	(0.006)
culture_activ	-0.001	-0.001	-0.001	0.002	0.002	0.002
	(0.002)	(0.003)	(0.004)	(0.010)	(0.008)	(0.007)
family_income	0.101***	0.101***	0.101***	0.060***	0.060***	0.060***
	(0.001)	(0.007)	(0.003)	(0.008)	(0.008)	(0.003)
age	-0.060***	-0.060***	-0.060***	-0.058***	-0.058***	-0.058***
	(0.001)	(0.008)	(0.005)	(0.011)	(0.010)	(0.009)
study_hours	0.054***	0.054***	0.054***	0.049***	0.049***	0.049***
	(0.001)	(0.004)	(0.002)	(0.007)	(0.005)	(0.002)
books_read	0.028***	0.028***	0.028***	0.034***	0.034***	0.034***
	(0.001)	(0.003)	(0.003)	(0.007)	(0.004)	(0.003)
gender (1=if male, 0=female)	0.088***	0.088***	0.088***	0.111***	0.111***	0.111***
	(0.003)	(0.022)	(0.007)	(0.014)	(0.024)	(0.010)
num_people	-0.042***	-0.042***	-0.042***	-0.034***	-0.034***	-0.034***
	(0.001)	(0.002)	(0.003)	(0.007)	(0.003)	(0.002)
start_grad	0.049***	0.049***	0.049***	0.070***	0.070***	0.070***
	(0.002)	(0.003)	(0.007)	(0.010)	(0.005)	(0.005)
Fixed effect						
Top 5 variables	Yes	Yes	Yes	-	-	-
Top 15 variables	-	-	-	Yes	Yes	Yes
Fit statistics						
Observations	765,899	765,899	765,899	765,899	765,899	765,899
R^2	0.160	0.160	0.160	0.880	0.880	0.880
R^2-adjusted	0.119	0.119	0.119	0.162	0.162	0.162
Robust error (clustering)	student	group	course_state	student	group	course_state

Note: Significance levels used: ***1%; **5%; *10%. The machine learning algorithm, through the feature selection procedure, identified 5 relevant attributes: group, financial_aid, high_school, acad_scholarship, and mother_edu. Considering now the top 15 attributes selected through feature selection, we have: group, financial_aid, high_school, acad_scholarship, mother_edu, acad_orga, father_edu, curric_activ, course_state, fam_higher_edu, fed_unit, race, marital_status, degree_mode, and empl_status. For a detailed description of these variables, please refer to Table 4.

3.4 Potential mechanisms

The positive impact of information and communication technologies (ICTs) on students' perceptions in higher education can be elucidated by multiple factors. Primarily, technology enhances the classroom environment by bringing students closer to a diverse array of information sources, providing teachers with opportunities to explore various resources to improve learning. Additionally, it fosters interactive and collaborative learning, more effectively engaging students. The internet offers an extensive range of educational materials, including videos and e-books, enabling students to independently supplement their classroom learning. ICTs also facilitate immediate assessments and feedback through online tests and interactive exercises, assisting students in identifying areas of weakness.

Concerning the effects of teamwork, the process of bringing together individuals with diverse backgrounds and experiences enriches discussions and deepens the understanding of the subjects studied. It promotes collaborative learning, aiding students to grasp complex concepts by making them more accessible through mutual explanation. Moreover, team collaboration serves as a source of motivation, as students feel part of a collective effort. This sense of belonging can enhance student engagement and commitment to academic activities. Thus, teamwork during undergraduate studies not only improves educational performance but also prepares students for a more successful and rewarding academic and professional life. This result aligns with the "good practices in higher education" proposed by Chickering and Gamson (1987, 1991), who emphasized the development of reciprocity and cooperation among students.

Regarding the influence of individualized teacher support outside the classroom, it can be elucidated by multiple factors. For instance, each student has unique needs and learning styles, and the individualized approach allows teachers to adjust their teaching to meet these specific needs, providing personalized support that is not feasible during regular classes. This support can range from clarifying doubts to motivation, progress monitoring, socio-emotional skill development, and academic guidance. This includes identifying areas where students may be struggling and implementing targeted strategies to improve their academic performance. This result is also aligned with the "good practices in higher education" proposed by Chickering and Gamson (1987, 1991), who stressed the importance of fostering contact between teachers and students as a crucial element for educational success.

Finally, the mechanism of disseminating pedagogical practice involving student participation in academic activities and events, both internal and external to the institution, can be driven by the following fundamental reasons. In the academic environment, there is often a tendency toward knowledge isolation, often motivated by concerns related to inadequate sharing. However, active participation in academic events promotes a more dynamic interaction between theory and practice. These events provide students with a unique opportunity to understand how theoretical knowledge can be applied in real situations, consolidating and enriching their understanding. This prepares them not only for a stronger education but also for the challenges of the job market, enhancing their success in both educational and professional realms.

4. CONCLUSION

This study investigated students' perceptions of 10 unconventional educational practices in higher education institutions (HEIs) in Brazil with the aim of identifying variables that positively impact academic performance. Using machine learning techniques and econometric analysis, we analyzed microdata from the National Student Performance Examination for the pre-COVID-19 pandemic triennium, encompassing all courses (major programs) of HEIs.

Our empirical analyses revealed four variables with a positive impact on academic performance: the integration of information and communication technologies (ICTs) as a teaching strategy; the development of teamwork skills among students; the availability of teachers to assist students outside of class hours; and the promotion of conditions that enable students to participate in internal and external events at institutions.

This study goes beyond highlighting the vital importance of these educational strategies in Brazilian HEIs; it also emphasizes the prominent utility of artificial intelligence in applied economic analysis of educational issues. The automation and refinement of the process of selecting qualifying attributes of HEIs, coupled with their integration in econometric analysis, resulted in substantial and significant evidence. Furthermore, the comprehensive exploration of extensive datasets from the National Student Performance Examination provided unprecedented and in-depth insights into the determinants of educational performance in HEIs.

In summary, these substantial and comprehensive findings underscore the undeniable relevance of these educational strategies in Brazilian HEIs, providing robust support for the development of public policies aimed at optimizing the pedagogical use of these innovative practices. Moreover, our results convincingly demonstrate how the use of advanced methods, such as artificial intelligence, can not only contribute to more accurate analysis but also to efficient automation in identifying the determinants of educational performance in HEIs, thus promoting excellence in higher education.

Finally, we acknowledge the inherent limitations of this study, particularly concerning the distribution of observations by gender and course, resulting from database sampling. While we do not believe this issue significantly impacted the results, refining the approach before econometric estimation could enhance the comprehensiveness of the analysis, contributing to future studies and reinforcing the robustness of the conclusions. Another limitation is associated with the machine learning approach, where we opted for a single technique. Exploring alternative approaches for attribute selection, such as decision trees, among others, could be addressed in future research.

REFERENCES

Aithal, P. S. and Kumar, P. M., How Innovations and Best Practices Can Transform Higher Education Institutions: A Case Study of SIMS (February 17, 2015). *International Journal of Management (IJM)*, Volume 6, Issue 2, February (2015), pp. 83-98.

Alencar, E. M. L. S.; Fleith, D. S. Inventário de Práticas Docentes que Favorecem a Criatividade no Ensino Superior. *Psicologia: Reflexão e Crítica*, 2004, 17(1), pp.105-110.

Arias, A. V., Lozano, A.B., Cabanach, R.G., Pérez, J.C.N. (1999). Las estratégias de aprendizaje: revisión teorica y conceptual. *Revista Latinoamericana de Psicologia*, 31 (3), pp. 425-461.

Barbel, N. A. N. As metodologias ativas e a promoção da autonomia de estudantes. *Semina: Ciências Sociais e Humanas, Londrina*, v. 32, n. 1, p. 25-40, jan./jun. 2011.

Batts, D., Colaric, S. M., & McFadden, C. (2006). Online courses demonstrate use of seven principles. *International Journal of Instructional Technology & Distance Learning*, 3(12), 15-26.

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *Clearing House*, 83(2), 39-43.doi.org/10.1080/00098650903505415.

Bergmann, J., & Sams, A. (2009). Remixing chemistry class: Two Colorado teachers make vodcasts of their lectures to free up class time for hands-on activities. *Learning & Leading with Technology*, 36(4), 22–27.

Biffi, Maríndia, et al. "Active learning methodologies: challenges for professors of two medical schools in Rio Grande do Sul, Brazil." *Revista Brasileira de Educação Médica* 44 (2020).

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3/4), 369. doi.org/10.1080/00461520.1991.9653139.

Bonwell, C. C. & Eison, J. A. (1991). Active learning: creating excitement in the classroom. ASHE-ERIC higher education reports. *ERIC Clearinghouse on Higher Education*, The George Washington University, Washington, DC.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *American Association for Higher Education Bulletin*, 39(7), 3-7.

Chickering, A. W., & Gamson, Z. F. (1991). Applying the seven principles for good practice in higher education. San Francisco, CA: Jossey-Bass.

Chickering, A. W., & Ehrmann, S. C. (1996). Implementing the seven principles: Technology as a lever. *American Association for Higher Education Bulletin*, 49(2), 3-6.

Cornachione Junior, E. B.; Cunha, J. V. A.; de Luca, M. M. M.; Ott, E. O bom é meu, o ruim é seu: perspectivas da teoria da atribuição sobre o desempenho acadêmico de alunos da graduação em Ciências Contábeis. *Rev. contab. finanç*. 2010.

Corrêa, Matheus, et al. "An analysis of the insertion of sustainability elements in undergraduate design courses offered by Brazilian higher education institutions: An exploratory study." *Journal of Cleaner Production* 272 (2020): 122733.

Cualheta, Luciana Padovez, and Gardenia da Silva Abbad. "What does entrepreneurship education look like in Brazil? An analysis of undergraduate teaching plans." *Education+ Training* 63.7/8 (2021): 1043-1057.

Cyrino, L. G.; Pereira, M. L. T. Trabalhando com estratégias de ensino-aprendizado por descoberta na área da saúde: a problematização e a aprendizagem baseada em problemas. *Cad. Saúde Pública*, Rio de Janeiro, 20(3):780-788. 2004.

de Oliveira Neto, José Dutra, and Gilvania de Sousa Gomes. "Technology Enhanced Learning for Higher Education in Brazil." *eLearn* 2016.8 (2016).

Dean, T.; Lee-Post, A.; Hapke, H. Universal Design for Learning in Teaching Large Lecture Classes. *Journal of Marketing Education*. 2016.

Duggan, S., & Gott, R. (2002). What sort of science education do we really need? *International Journal of Science Education*, 24(7), 661-679. doi.org/10.1080/09500690110110133

Estévez-Ayres, Iria; Alario-Hoyos, Carlos; Pérez-Sanagustín, Mar; Pardo, Abelardo; Crespo-García, Raquel M.; Leony, Derick; Parada G., Hugo A.; Delgado-Kloos, Carlos. A methodology for improving active learning engineering courses with a large number of students and teachers through feedback gathering and iterative refinement. ERIC - *International Journal of Technology and Design Education*. 2015.

Gaižiūnienė, Lina. Innovative Study Methods Adaptation: The Power of An Effective Teacher. *The European Journal of Social and Behavioural Sciences* EJSBS Volume XXIV. 2018.

Garcia, R; Falkner, K; Vivian, R. Systematic literature review: Self-Regulated Learning strategies using e-learning tools for Computer Science, Computers & Education, Volume 123, 2018, Pages 150-163, ISSN 0360-1315. https://doi.org/10.1016/j.compedu.2018.05.006.

Guerra, C. J. O., & Teixeira, A. J. C. Os impactos da adoção de metodologias ativas no desempenho dos discentes do curso de ciências contábeis de instituição de ensino superior mineira. *Revista De Educação E Pesquisa Em Contabilidade* (REPeC), 10(4). 2016.

Heringer, Markoni Ramires, et al. "Innovation in Brazilian private higher education: A proposal for the application of active methodologies based on the flipped classroom." *International Journal of Innovation* 7.2 (2019): 221-340.

Herreid, C. F., N. A. Schiller, K. F. Herreid, and C. Wright. 2011. In case you are interested: results of a survey of case study teachers. *J. Col. Sci. Teach.* 40(4):76–80.

Hoffman, B., & Ritchie, D. (1997). Using multimedia to overcome the problems with problem-based learning. *Instructional Science*, 25, 97–115. http://link.springer.com/journal/11251

Huffman, W. H.; Huffman, A. H. Beyond basic study skills: The use of technology for success in college. *Elsevier*. 2011.

Karabulut, U.S. (2002). Curricular Elements of Problem-Based Learning at Cause Developments of SelfDirected Learning Behaviors Among Students and Its Implications on Elementary Education. Master's thesis, *University of Tennessee*, Retrieved from http://trace.tennessee.edu/utk_gradthes/2078

Ketpichainarong, W., Panijpan, B., & Ruenwongsa, P. (2010). Enhanced learning of biotechnology students by an inquiry-based cellulose laboratory. *International Journal of Environmental and Science Education*, 5(2), 169-187. Retrieved from http://www.ijese.net

Kolb, D. A. (1981). Learning styles and disciplinary differences. In A. Chickering (Ed.), *The modern American college* (pp. 232-255). San Francisco, CA: Jossey-Bass.

Lage, M., & Platt, G. (2000). The internet and the inverted classroom. *Journal of Economic Education*, 31, 11.

Lima, Rui M., Diana Mesquita, and Luciana Coelho. "Five Years of Project-Based Learning Training *Experiences in Higher Education Institutions in Brazil*. 2016".

Limberger, J. B. Metodologias ativas de ensino-aprendizagem para educação farmacêutica: um relato de experiência. *Interface*. 2013.

Machado, A.; Terra, R.; Tannuri-Pianto, M. Higher education responses to accountability. *Economics of Education Review*, Volume 98, 2024, 102493, ISSN 0272-7757. https://doi.org/10.1016/j.econedurev.2023.102493.

Malerva, I. S.; Escorza, Y. H. Estrategias de aprendizaje y desempeño académico en estudiantes de Medicina. *Educación Médica*. 2018.

McCabe, D. B.; Meuter, M. L. A Student View of Technology in the Classroom: Does It Enhance the Seven Principles of Good Practice in Undergraduate Education? *Journal of Marketing Education*, 2011.

Mello, M. B. J. B.; Leme, M. I. S. Motivação de alunos dos cursos superiores de tecnologia. *Psicol. Esc. Educ.* 2016.

Ngeow, K., & Kong, Y. S. (2001). Learning To Learn: Preparing Teachers and Students for Problem-Based Learning. *ERIC Digest*. Retrieved from http://www.ericdigests.org/2002-2/problem.htm

Tavares, P.A. The impact of school management practices on educational performance: Evidence from public schools in São Paulo. *Economics of Education Review*, Volume 48, 2015, Pages 1-15, ISSN 0272-7757.

https://doi.org/10.1016/j.econedurev.2015.05.002.

Riccomini, Fernanda Edileuza, et al. "Educational innovation: Trends for higher education in Brazil." *International Journal of Educational Management* 35.3 (2021): 564-578.

Santana, A. L. A.; Araújo, A. M. P. Aspectos do perfil do professor de Ciências Contábeis e seu reflexo no Exame Nacional de Desempenho dos estudantes (ENADE) – um estudo nas universidades federais do Brasil. *Universidade Federal de Minas Gerais*, Belo Horizonte, 2011.

Signori, Gláuber Guilherme, et al. "Gamification as an innovative method in the processes of learning in higher education institutions." *International Journal of Innovation and Learning* 24.2 (2018): 115-137.

Silva, T. C.; Tabak, B. M.; Ferreira, I. M. Modeling Investor Behavior Using Machine Learning: Mean-Reversion and Momentum Trading Strategies. *Hindawi*. 2019.

Souza, C. S.; Iglesias, A.G; Pazin Filho, A. Estratégias inovadoras para métodos de ensino tradicionais – aspectos gerais. *Revista FMRP USP*. 2014. https://www.revistas.usp.br/rmrp/article/view/86617/89547.

KOLB, D. A. Learning Style Inventary Technical Manual. *Boston: Hay Mcber*, 1976.

KOLB, D. A. Experimental learning: Experience as the source of learning and development. *New Jersey: Prentice-Hall*, Englewood Cliffs, 1984.

Rideout, E. (2001). Transforming nursing education through problem-based learning. Sudbury, MA: *Jones and Bartlett Publishers*

Sungur, S., & Tekkaya, C. (2006). Effects of problem-based learning and traditional instruction on self-regulated learning. *The journal of educational research*, 99(5), 307-320. doi.org/10.3200/JOER.99.5.307-320

Torp, L., & Sage, S. (2002). Problems as possibilities: Problem-based learning for K-12 education. *ASCD*.

Vieira, R.S; Arends-Kuenning, M,. Affirmative action in Brazilian universities: Effects on the enrollment of targeted groups. *Economics of Education Review*, Volume 73, 2019, 101931. ISSN 0272-7757, https://doi.org/10.1016/j.econedurev.2019.101931.

Xu, M., Benson, S. N. K., Mudrey-Camino, R., & Steiner, R. P. (2010). The relationship between parental involvement, self-regulated learning, and reading achievement of fifth graders: A path analysis using the ECLS-K database. *Social Psychology of Education*, 13(2), pp. 237-269. doi: 10.1007/ s11218-009-9104-4

Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a Structured Interview for Assessing Student Use of Self-Regulated Learning Strategies. *American Educational Research Journal*, 23, 614-628. https://doi.org/10.3102/00028312023004614

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). Self-regulated learning and academic achievement: *Theoretical perspectives*. Routledge.

APPENDIX

Table 7 - Composition of HEI courses in the "*group*" variable

Group area	Descriptions
Educations	702 = Mathematics (Bachelor's Degree) 904 = Portuguese Language and Literature (Bachelor's Degree) 905 = Portuguese and English Language and Literature (Bachelor's Degree) 906 = Portuguese and Spanish Language and Literature (Bachelor's Degree) 1402 = Physics (Bachelor's Degree) 1502 = Chemistry (Bachelor's Degree) 1602 = Biological Sciences (Bachelor's Degree) 2001 = Pedagogy (Bachelor's Degree) 2402 = History (Bachelor's Degree) 3002 = Geography (Bachelor's Degree) 3502 = Physical Education (Bachelor's Degree)
Humanities and Arts	26 = Design 83 = Fashion Design Technology 103 = Interior Design Technology 104 = Graphic Design Technology 903 = Portuguese Language and Literature (Bachelor's Degree) 2401 = History (Bachelor's Degree) 2501 = Visual Arts (Bachelor's Degree)
Social Sciences, Business, and Law	5401 = Social Sciences (Bachelor's Degree) 102 = International Trade Technology 100 = Public Administration 94 = Logistics Technology 93 = Commercial Management Technology 87 = Financial Management Technology 86 = Human Resource Management Technology 1 = Business Administration 2 = Law 13 = Economics 18 = Psychology 22 = Accounting 81 = International Relations 84 = Marketing Technology 85 = Management Processes Technology

	106 = Public Administration Technology 803 = Social Communication - Journalism 804 = Social Communication - Advertising and Propaganda 91 = Hospital Management Technology 5402 = Social Sciences (Teacher Education)
Science, Mathematics, and Computer Science	72 = Information Systems Analysis and Development Technology 79 = Computer Networks Technology 701 = Mathematics (Bachelor's Degree) 1601 = Biological Sciences (Bachelor's Degree) 3001 = Geography (Bachelor's Degree) 55 = Biomedicine 4006 = Information Systems 6409 = Information Technology Management Technology
Engineering, Production, and Construction	21 = Architecture and Urbanism 4003 = Computer Engineering 5710 = Civil Engineering 5806 = Electrical Engineering 5814 = Control and Automation Engineering 5902 = Mechanical Engineering 6002 = Food Engineering 6008 = Chemical Engineering 6208 = Production Engineering 6306 = Engineering 6307 = Environmental Engineering 6405 = Forestry Engineering
Agriculture and Veterinary Science	5 = Veterinary Medicine 90 = Agribusiness Technology
Health and Social Well-being	6 = Dentistry 12 = Medicine 19 = Pharmacy 23 = Nursing 27 = Speech Therapy 28 = Nutrition 36 = Physiotherapy 38 = Social Work 69 = Radiology Technology 3501 = Physical Education (Bachelor's Degree)
Services	29 = Tourism 88 = Gastronomy Technology 92 = Environmental Management Technology 95 = Aesthetics and Cosmetics Technology

Note: Based on the ENADE questionnaires for the years 2016, 2017, and 2018.