Check for updates

DOI: https://doi.org/10.46502/issn.1856-7576/2025.19.02.1

Cómo citar:
Batista da Silva, J., Assis da Costa, L.C., dos Santos, J.N., & Oliveira da Silva Batista, R. (2025). Using machine learning techniques to predict academic performance in mathematics. *Revista Eduweb*, *19*(2), 9-19. https://doi.org/10.46502/issn.1856-7576/2025.19.02.1

Using machine learning techniques to predict academic performance in mathematics

Utilización de técnicas de aprendizaje automático para predecir el rendimiento escolar en matemáticas

João Batista da Silva
PhD in Education, State University of Ceará, Fortaleza, Brazil.
https://orcid.org/0000-0003-3556-9881
joaobatista.silva@aluno.uece.br
Luis Carlos Assis da Costa
Master's student in Science and Mathematics Teaching, Federal University of Ceará, Fortaleza, Brazil.
https://orcid.org/0009-0002-3000-9770
luiscarlosassisdacosta@alu.ufc.br
José Nilson dos Santos
Graduate degree in Pharmacy, Federal University of Ceará, Fortaleza, Brazil.
https://orcid.org/0009-0000-1822-3788
jsantostar@alu.ufc.br
Raquel Oliveira da Silva Batista
Graduate degree in Human Resources, Fametro University Center, Fortaleza, Brazil.
https://orcid.org/0009-0005-8825-2329
raquel.batista01@aluno.unifametro.edu.br

Recibido: 24/04/25
Aceptado: 19/05/25

Abstract

The purpose of this study was to investigate the predictive power of the SAEF exams in estimating the schools' performance in Mathematics on the SPAECE exam. To achieve this, we developed a predictive machine learning model. The model was trained using data from 133 schools that participated in the exams in 2022, and subsequently tested with data from 140 schools that took part in the exams in 2023. The results showed that the random forest (RF) model demonstrated moderate predictive power ($R^2 = 0.397$), which was superior to the linear model ($R^2 = 0.384$). This means that approximately 39.7% of the variance in schools' Mathematics performance on the SPAECE can be explained by the results of the SAEF exams. The first SAEF exam, administered at the beginning of the academic year, demonstrated the highest predictive power among the three, indicating that students' initial performance in Mathematics is a strong indicator of their future performance. These findings underscore the importance of early identification of learning difficulties to enable strategic pedagogical interventions throughout the year. Although this study was conducted within the Brazilian educational context, other countries can also utilize machine learning techniques to monitor students' academic trajectories and predict their outcomes in standardized assessments.

Keywords: artificial intelligence, assessment, large-scale, machine learning, random forest.

Creative Commons Attribution 4.0 International (CC BY 4.0)

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

Resumen

El propósito de este estudio fue investigar el poder predictivo de los exámenes del SAEF para estimar el desempeño en Matemáticas de las escuelas en el examen del SPAECE. Para ello, desarrollamos un modelo predictivo de Machine Learning. El modelo fue entrenado con datos de 133 escuelas que participaron en los exámenes en 2022 y posteriormente probado con datos de 140 escuelas que participaron en los exámenes en 2023. Los resultados mostraron que el modelo Random Forest (RF) presentó una capacidad predictiva moderada (R^2 = 0.397), superior al modelo lineal (R^2 = 0.384). Esto significa que aproximadamente el 39.7% de la variación en el desempeño de las escuelas en Matemáticas en el SPAECE puede ser explicada por los resultados de los exámenes del SAEF. El primer examen del SAEF, aplicado al inicio del año lectivo, presentó el mayor poder predictivo entre los tres, indicando que el desempeño inicial de los estudiantes en Matemáticas es un fuerte indicador de su desempeño futuro. Estos hallazgos destacan la relevancia de la identificación temprana de dificultades de aprendizaje para viabilizar intervenciones pedagógicas estratégicas a lo largo del año. Aunque este estudio fue realizado en el contexto educativo brasileño, otros países también pueden utilizar técnicas de Machine Learning para monitorear la trayectoria académica de los estudiantes y predecir sus resultados en evaluaciones estandarizadas.

Palabras clave: Inteligencia artificial, Evaluación, Gran escala, machine learning, random forest.

Introduction

Brazil's integration into the global economy has reinforced the implementation of educational policies by incorporating management models derived from a conception of an evaluative state that gradually strengthened and broadened the scope of large-scale assessments in the country, leading Brazilian states to follow suit (Afonso, 2009; Ball, 1998; Barroso, 2003; Gentili & Silva, 2015; Lessard & Carpentier, 2016; Costa et al., 2019).

Brazilian states implement standardized tests to monitor student performance and promote pedagogical interventions aimed at improving educational indicators. The state of Ceará is a pioneering example in the implementation of specific evaluations using the Ceará Permanent System for the Evaluation of Basic Education (Sistema Permanente de Avaliação da Educação Básica do Ceará – SPAECE, in Portuguese). In Brazil, basic education refers to the entire spectrum of primary and secondary education, from kindergarten to high school (Vidal et al., 2024).

The government of Ceará created SPAECE in 1992 to adapt national guidelines to local specificities and assess student competencies (Lima et al., 2021). The Ceará Department of Education annually administers the standardized SPAECE exam in public schools, focusing on reading and mathematics skills.

In the city of Fortaleza, the capital of Ceará, the Municipal Department of Education (SME) developed the Elementary Education Assessment System (Sistema de Avaliação do Ensino Fundamental – SAEF) to monitor the performance of students in municipal public schools and prepare them for SPAECE (Mendes et al., 2023). The SME administers the SAEF tests three times a year to students in elementary school (1st–5th grade) and middle school (6th–8th grade).

SAEF data enable the performance diagnosis of students for administrators and teachers (Lopes et al., 2017). Moreover, it facilitates the implementation of pedagogical interventions not only to enhance learning but also to prepare and train students for state assessments, such as the SPAECE exam (Oliveira et al., 2021; Vidal & Costa, 2021). Educational administrators refer to this approach as performance management (Camarão et al., 2015).

The Ceará government has been strengthening the performance management policy as a strategic approach in public education, based on the logic of accountability (Costa & Vidal, 2020). This approach

Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

aims to hold schools accountable for their results, encouraging them to meet clear and measurable performance targets. In this context, the state government has implemented school bonuses tied to the results of large-scale assessments to stimulate improvements in performance indices, based on the scores that students achieve on the SPAECE.

Considering that one of the main purposes of the SAEF is to prepare students to take the large-scale SPAECE examination (Oliveira et al., 2021; Holanda, 2024), the following research question arises: What is the predictive power of the SAEF exams in estimating students' performance on the SPAECE?

In this context, the purpose of this study was to investigate the predictive power of the SAEF exams in estimating schools' Mathematics performance on the SPAECE examination. To achieve this, we applied machine learning techniques (random forest) to identify statistical correlations between the exam results and the performance on the SPAECE.

This study formulated two research hypotheses. Hypothesis 1 is that students' results in the SAEF have predictive power to estimate school performance in the SPAECE exam, particularly in Mathematics. Hypothesis 2 proposes that the random forest machine learning model has greater predictive capability for student performance in SPAECE compared to traditional statistical models, such as linear regression.

Random Forest: concepts, construction, and applications in machine learning

Random Forest (RF) is a supervised machine learning (ML) algorithm widely used for classification and regression tasks, notable for its effectiveness and versatility across various applications. In ML, a supervised algorithm is one that learns from labeled data, where each example in the dataset has an associated input and output.

The RF analysis method is a non-parametric technique aimed at improving the predictive performance of models (Breiman, 2001). This algorithm is extensively utilized in the field of Data Science, combining multiple decision trees to achieve results of high precision and stability (Queiroga et al., 2022).

The RF model is built through the combination of decision tree predictors, with each tree using a different bootstrap sample of the data, and successive trees not depending on previous ones (Liaw & Wiener, 2002). The statistical technique called bootstrap creates several data samples from the original sample of the same size, allowing some data to repeat while others may not be present in a given sample (Figure 1). This introduces variation among the samples, which can be beneficial for the robustness of the model.

Creative Commons Attribution 4.0 International (CC BY 4.0)

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

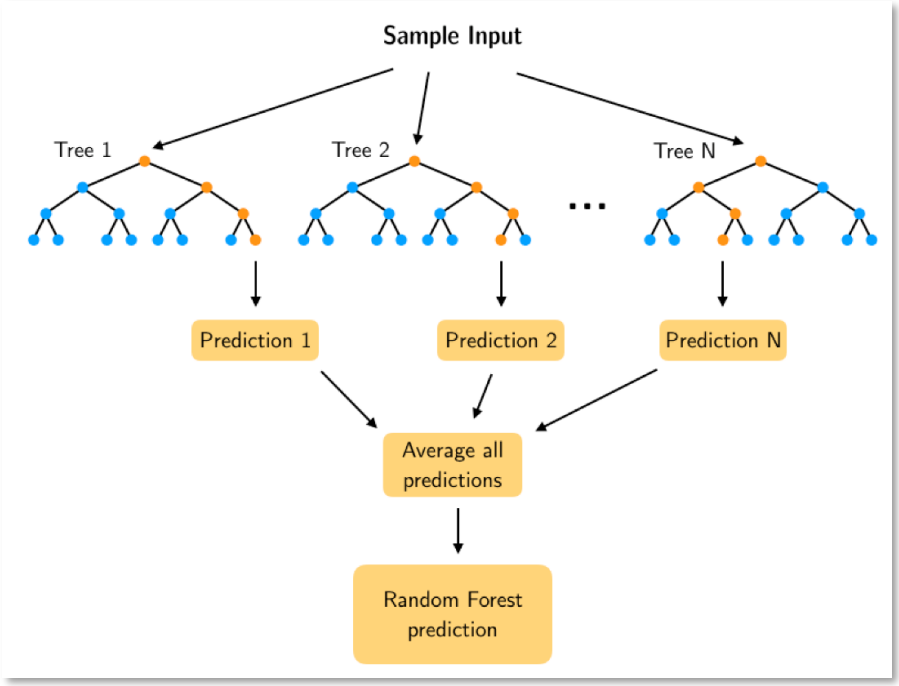


Figure 1. Schematic diagram of random forest model (Yang et al., 2022)

The training of each tree takes place on a random subset of the training data. The splitting of data at each tree node is conducted using a random subset of features (Queiroga et al., 2022). This introduces randomness and diversity into the learning process, making the trees less prone to overfitting. Overfitting is an excessive tailoring to the training data in more complex models (Dewi & Chen, 2019), which results in excellent performance on the training set but likely leads to poorer performance on test data (Probst et al., 2019).

After the trees are trained, the RF algorithm produces the final prediction by computing the average of the predictions from the individual trees that make up the forest (Dewi & Chen, 2019). The principle of the random forest algorithm can be expressed as follows:

$$Y = \frac{1}{N}\sum_{t=1}^{n} F_t(X)$$

where X is the input feature vector, Y is the prediction result, and N is the number of regression tree models built (Yang et al., 2022).

One of the advantages of RF is its ability to handle multicollinearity by accommodating highly correlated variables without introducing bias (Liaw & Wiener, 2002). Additionally, RF excels in feature selection, allowing for the assessment of the relative importance of each variable, thereby assisting in the selection of the most relevant features.

Methodology

Participants

This study analyzes information on the mathematics performance of 18,389 students from 133 municipal

Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

public schools in Fortaleza, in the state of Ceará, Brazil. The performance of each school is determined by the average scores of eighth-grade students on the standardized SPAECE mathematics exams. The students' average ages range from 14 to 15 years.

Data collection

This study utilized data from two official government sources: SAEF (SME, n.d.) and SPAECE (SEDUC, n.d.). The dataset covers the period from 2022 to 2023. The SAEF administers three mathematics exams throughout the year: the first (Saef_1) at the beginning of the academic year in February, the second (Saef_2) in June, and the third (Saef_3) in October. Each exam contains 26 mathematics items. The detailed analysis of student performance is conducted by the Learning Assessment Cell (Célula de Avaliação da Aprendizagem - CEAVA), a division of the Municipal Education Secretariat (Secretaria Municipal de Educação - SME) of Fortaleza, which plays a crucial role in planning, developing, coordinating, and monitoring the SAEF. Furthermore, CEAVA collects and analyzes data on student learning, providing valuable information for the development of public educational policies (Fortaleza, 2019).

In November, SPAECE conducts a large-scale assessment consisting of an exam with 26 mathematics items. The Center for Public Policies and Evaluation of Education (Centro de Políticas Públicas e Avaliação da Educação - CAEd, in Portuguese) is responsible for designing and administering the SPAECE exam. CAEd is an institution affiliated with the Federal University of Juiz de Fora (UFJF), specializing in educational assessment. Its goal is to support public managers and educational institutions in the formulation, implementation, and monitoring of educational policies, based on data collected through large-scale assessments and other methodologies.

Estimation of Model

To achieve the purpose of this research, we developed a predictive model using advanced machine learning techniques (Edwards et al., 2021), with an emphasis on the random forest algorithm. We chose this algorithm for its robust ability to handle complex data and its functionalities in prediction processes involving multiple variables. The application of the model aimed to clarify the impact of SAEF on school performance in SPAECE.

In constructing the predictive model, we selected three explanatory variables, named Saef_1, Saef_2, and Saef_3. These variables correspond to the schools' performance in the SAEF mathematics exam throughout the academic year. Additionally, we used a response variable, named Spaece_Mathematics, which represents the schools' performance in the SPAECE mathematics exam. Table 1 provides a detailed description of each variable.

Table 1.
Description of the predictor variables

Code	Variables	Type	Description	Month of Application	Developers	Level
Saef_1	Explanatory	Continuous	Performance in the first SAEF Mathematics exam	February	CEAVA/SME	District level
Saef_2	Explanatory	Continuous	Performance in the second SAEF Mathematics exam	June	CEAVA/SME	District level
Saef_3	Explanatory	Continuous	Performance in the third SAEF Mathematics exam	Octuber	CEAVA/SME	District level
Spaece_Mathematics	Response	Continuous	Performance in the SPAECE Mathematics exam	November	CAEd	State Level

Creative Commons Attribution 4.0 International (CC BY 4.0)

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

Based on the Hold-out method (Kumar, 2022), we trained the model using performance data from 133 schools that participated in the mathematics exams in 2022. After training, we tested the model with data from 140 schools that participated in the exams in 2023.

We analyzed the data using the random forest regression technique, a robust machine learning method based on decision trees. This approach allows for capturing complex relationships between explanatory variables (Saef_1, Saef_2, and Saef_3) and the response variable (Spaece_Mathematics). Additionally, the RF analysis method efficiently handles nonlinear and multivariate data. The generated model identifies patterns in the data and evaluates the relative contribution of each explanatory variable in predicting the final performance in the SPAECE exam.

Additionally, we examined variable importance and the distribution of minimum depth through tree modeling to identify which are most influential in predicting the outcome (Spaece_Mathematics). We conducted the RF model estimation using RStudio software and the 'randomForest' and 'randomForestExplainer' packages.

Data analysis

We conducted an evaluation of the RF model's performance using three specific metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and the Coefficient of Determination (R^2). These metrics provide crucial insights into the accuracy and generalization capability of the developed model.

The MAE quantifies the average absolute differences between the predicted and observed values, serving as an intuitive measure of the model's accuracy. Since it does not account for the direction of errors (positive or negative), a lower value indicates greater accuracy in predictions.

The MSE, in turn, measures the discrepancy between the predicted and actual values by averaging the squares of the prediction errors. The lower the MSE, the more precise the model's predictions; higher values of this metric indicate a larger error in estimating the outcomes (Yang et al., 2018).

Meanwhile, the Coefficient of Determination (R^2) expresses the proportion of variability in the dependent variable that is explained by the model. Its value ranges between 0 and 1, with values closer to 1 indicating a higher explanatory power of the model, while values nearing 0 suggest low adherence to the observed data (Hair et al., 2009).

By applying these metrics together, we comprehensively evaluated the model's performance, ensuring its accuracy and reliability in analyzing the factors that influence academic performance in mathematics. Furthermore, we used these metrics to fine-tune the model, fostering a continuous process of validation and refinement of the results.

Results and Discussion

This research aimed to investigate the predictive power of the SAEF examinations for estimating students' performance in Mathematics on the large-scale standardized SPAECE exam. To achieve this, we developed a predictive model using the random forest algorithm.

The results indicated that the random forest model (Model_rf) exhibited moderate predictive ability (R^2 = 0.397), suggesting that approximately 40% of the variance in schools' mathematics performance on SPAECE can be explained by the variables included in the model. Furthermore, the performance of the random forest-based model (Model_rf) was slightly superior to that of the Multiple Linear Regression model (Model_lm) across all the metrics analyzed. Table 2 presents the performance metrics of these models, which we evaluated using the test data.

Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

Table 2.
Comparison of model performance metrics with test data

	Testing		
	R^2	MAE	MSE
Model_rf	0.397	10.582	186.924
Model_lm	0.384	11.928	277.345

In terms of Mean Absolute Error (MAE), the Model_rf showed a lower value (10.582) compared to Model_lm (11.928), demonstrating greater prediction accuracy. The analysis of the Mean Squared Error (MSE) also revealed an advantage of Model_rf, with a value of 186.924 compared to 277.345 for Model_lm, suggesting that the predictions of the RF-based model were more consistent. Figure 2 displays the relative importance ranking of the explanatory variables (Saef_1, Saef_2, Saef_3), according to the Percentage Increase in Mean Squared Error (%IncMSE) and the Increase in Node Purity (IncNodePurity). These results confirm Hypothesis 2.

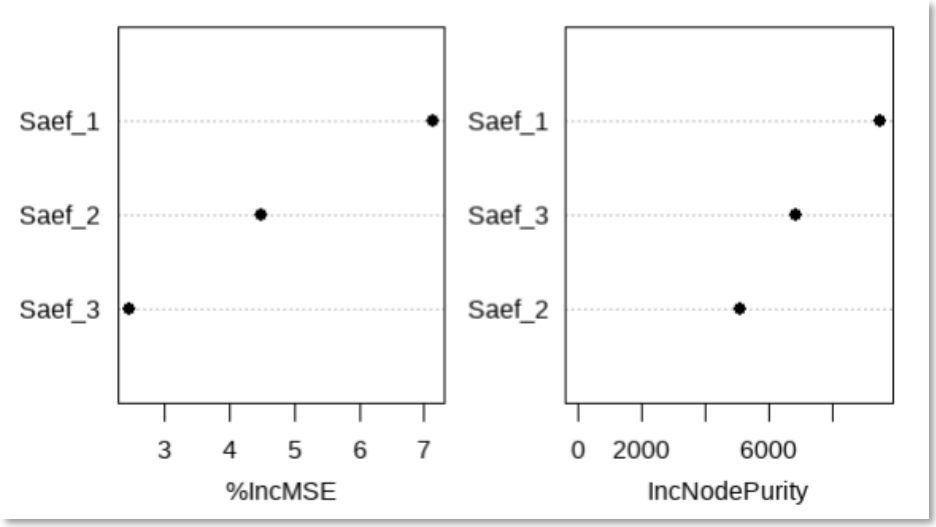


Figure 2. Relative importance of variables in the random forest model.

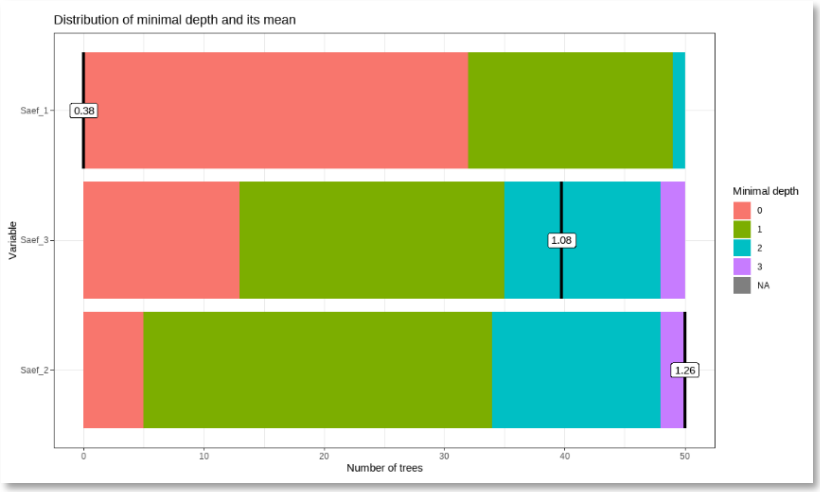


Figure 3. Distribution of minimal depth for variables according to the average minimum depth calculated.

Creative Commons Attribution 4.0 International (CC BY 4.0)

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

The first exam (Saef_1), administered at the beginning of the academic year, represents the most significant variable in the predictive model. This exam ranks first in terms of the percentage increase in average mean squared error (%IncMSE) (Figure 2) as well as in node impurity reduction in the tree (IncNodePurity). Furthermore, the Saef_1 variable exhibits a low minimum depth in the tree (Figure 3), underscoring its high importance to the model. This is justified by the significant impact of removing this variable, which results in a substantial increase in mean squared error (MSE). Thus, the analysis reinforces the central role of Saef_1 in predicting student performance. These results support Hypothesis 1.

Variables with the highest %IncMSE values have the greatest influence on the predictive model. However, it is essential to emphasize that %IncMSE is a relative measure, as the importance of a variable is determined in relation to the others present in the model (Liaw & Wiener, 2002). Consequently, including or removing a variable can directly impact the %IncMSE values of others, altering their importance hierarchy within the model.

The results of this study evidence that the first SAEF exam possesses a high predictive power for students' final performance in SPAECE. This finding suggests that data collected at the beginning of the school year play a crucial role in anticipating academic performance throughout the academic period, reinforcing the relevance of diagnostic assessments in the early identification of difficulties and the planning of pedagogical interventions.

The results indicate that the initial exam plays a fundamental role in measuring the prior knowledge students bring at the start of the academic year, thereby exerting a significant influence on their academic performance throughout the school period. As argued by Lima et al. (2021), upon entering school, students already possess empirical knowledge that needs to be restructured for the construction and consolidation of new learning.

In this context, students' initial performance emerges as a key indicator for predicting academic progress, enabling the early identification of learning gaps and guiding the implementation of more effective pedagogical interventions. Therefore, this investigation underscores the importance of the first exam not merely as a diagnostic instrument but also as a strategic tool for educational planning. By offering an overview of students' knowledge level, this assessment allows educators to develop teaching plans that are more aligned with the specific needs of each group, contributing to the reduction of learning gaps and the optimization of students' performance in external assessments, such as SPAECE.

The findings of this research are consistent with the existing literature, particularly with the contributions of Ausubel et al. (1968), which emphasize the importance of subsumers – prior concepts stored in the cognitive structure – as fundamental elements for meaningful learning. More recent studies support this theory (Lopes et al., 2017; Silva, 2020; Queiroga et al., 2022), reinforcing the direct relationship between the student's initial knowledge and its implications for subsequent learning processes.

Lopes et al. (2017) highlight that when diagnostic assessments are conducted in alignment with the student's level of knowledge, they not only provide crucial information for the teacher's pedagogical support in daily practice but also allow for more accurate predictions of student outcomes in external evaluations, such as SPAECE. This predictive capability underscores the relevance of diagnostic assessment as a tool not only for diagnosing learning gaps but also as a strategic instrument for improving academic performance in large-scale assessments.

Queiroga et al. (2022) highlight that students' prior academic performance in Uruguay is one of the main predictors of their future performance, underscoring the importance of considering the knowledge base already acquired by students throughout their educational journey. This perspective is widely supported by David Ausubel's Theory of Meaningful Learning, which emphasizes the crucial role of prior knowledge in the learning process, considering it one of the most determining factors in students' cognitive development (Silva, 2020).

Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

According to Ausubel et al. (1968), among the various variables influencing learning, prior knowledge holds a central position and serves as a foundation for building new understandings. In this context, it is imperative that the starting point of the educational process takes into account what the student already knows, as this pre-existing knowledge facilitates the assimilation of new content. In line with this theory, the results of this study highlight the importance of using initial assessments as strategic tools to guide pedagogical decisions and plan instruction (Lima et al., 2021). By effectively exploring these initial assessments, it is possible to enhance student learning throughout the academic year, promoting more effective development aligned with their cognitive and educational needs.

Despite the significance of the study's findings, we acknowledge that it presents certain limitations that must be considered. First, the model explains approximately 40% of the variance in SPAECE scores, suggesting that other unmeasured factors also influence student outcomes. These factors may include socioeconomic status, teacher quality, school resources, student motivation, among others. Secondly, the study focuses exclusively on schools in Fortaleza, Ceará, Brazil. This regional focus restricts the generalizability of the results to other contexts, particularly those with different educational systems or demographic characteristics.

Conclusions

The purpose of this study was to investigate the predictive power of the SAEF exams in estimating Mathematics performance of public schools in the municipality of Fortaleza on the large-scale SPAECE exam. To this end, we employed a predictive model based on regression with the application of machine learning techniques, specifically using the random forest (RF) algorithm.

The results indicated that the RF model exhibited a moderate predictive capability ($R^2 = 0.397$), suggesting that approximately 39.7% of the variance in Mathematics performance on the SPAECE can be explained by the SAEF exams. The study's findings demonstrate the potential of SAEF as a diagnostic assessment tool. We identified that the first SAEF exam has the greatest predictive power for SPAECE performance, which suggests the importance of early pedagogical interventions, right at the beginning of the academic year.

However, it is important to highlight that the predictive capability is not absolute, indicating that there are other factors not captured by the model that also influence student performance. Furthermore, diagnostic assessment is merely a complementary tool, constructed from a formative perspective, aimed at indicating proficiency levels in Mathematics (Lima et al., 2021).

In this regard, for the diagnostic purpose of the SAEF to be fully effective, it is crucial that the results are interpreted with caution, avoiding inappropriate uses, such as the mere classification or ranking of schools. Additionally, continuous support for administrators and teachers is essential to use the data strategically in educational planning.

Based on the results obtained, we conclude that SAEF mathematics exams have significant potential to predict performance in large-scale assessments, especially when applied alongside algorithms that exploit the data structure more effectively, such as the random forest. This type of approach allows for greater precision in anticipating results, grounding strategic decisions in concrete data, which is essential for supporting decision-making in educational policy development.

SPAECE exemplifies a widely adopted practice of large-scale assessment used in various countries to monitor the quality of education and guide public policy formulation. In this context, the predictive capability of diagnostic exams, as demonstrated by the SAEF study, plays a fundamental role in supporting informed decision-making by educational managers.

Although this study is situated within the Brazilian educational context, the concern with school performance in large-scale assessments is a global issue. The research conducted in the municipality of Fortaleza, Ceará, demonstrates that the administration of periodic exams throughout the school year can significantly

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

Creative Commons Attribution 4.0 International (CC BY 4.0)

contribute to predicting student performance in large-scale assessments. This finding suggests that educational systems utilizing regular diagnostic assessments may adopt similar approaches to monitor and anticipate students' academic performance.

The incorporation of machine learning techniques in the analysis of educational data, such as the regression used in this study, reflects an international trend in the field of education. These techniques enable the identification of complex patterns and enhance the accuracy of predictions, aiding in the early detection of at-risk students and facilitating the implementation of more effective pedagogical interventions.

Countries that adopt periodic exams can benefit from the application of predictive models based on machine learning to estimate student performance in standardized assessments. The increasing use of these methodologies underscores their potential in education. Ultimately, this study contributes to the field by demonstrating the feasibility and applicability of predictive models in the analysis and management of academic performance.

Bibliographic references

Afonso, A. J. (2009). Nem tudo o que conta em educação é mensurável ou comparável. Crítica à accountability baseada em testes estandartizados e rankings escolares. *Revista Lusófona de Educação*, *13*(13), 13-29. https://revistas.ulusofona.pt/index.php/rleducacao/article/view/545

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1968). *Educational psychology: A cognitive view*. Nova Iorque: Holt, Rinehart & Winston.

Ball, S. J. (1998). Cidadania global, consumo e política educacional. In L. H. Silva (Org.), *A escola cidadã no contexto da globalização* (pp. 121–137). Petrópolis, RJ: Vozes.

Barroso, J. (Org.) (2003). *A escola pública: regulação, desregulação e privatização*. Porto: Asa.

Breiman, L. (2001). *Random forests*. Berkeley: University of California. https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

Camarão, V. C., Ramos, J. F. P., & Albuquerque, F. C. A. (2015). Política da gestão por resultados na educação cearense (1995-2014). *Revista Praia Vermelha*, *25*(2), 369–391. https://revistas.ufrj.br/index.php/praiavermelha/article/view/10160

Costa, A. G., Vidal, E. M., & Vieira, S. L. (2019). Avaliação em larga escala no Brasil. *Revista Educação Em Questão*, *57*(51). https://doi.org/10.21680/1981-1802.2019v57n51ID15806

Costa, A. G., & Vidal, E. M. (2020). Accountability e regulação da educação básica municipal no estado do Ceará–Brasil. *Revista Iberoamericana de Educación*, *83*(1), 121-141. https://doi.org/10.35362/rie8313852

Dewi, C., & Chen, R. C. (2019). Random Forest and Support Vector Machine on features selection for regression analysis. *International Journal of Innovative Computing*, *15*(6). https://doi.org/10.24507/ijicic.15.06.2027

Edwards, A. S., Kaplan, B., & Jie, T. (2021). A primer on machine learning. *Transplantation*, *105*(4), 699-703. https://doi.org/10.1097/tp.0000000000003316

Fortaleza. (2019). *Decreto nº 14.405, de 22 de abril de 2019*. Aprova o regulamento da Secretaria Municipal da Educação (SME). *Diário Oficial do Município*, *LXIV*(16.492), 1–17. https://acervo.fortaleza.ce.gov.br/download-file/documentById?id=21a8fc0b-bee3-4df2-a194-10716b6a0c0c

Gentili, P., & Silva, T. T. da. (orgs.) (2015). *Neoliberalismo, qualidade total e educação: visões críticas* (15ª ed). Petrópolis, RJ: Vozes.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados* (6ª ed.). Porto Alegre: Bookman.

Holanda, R. H. R. (2024). *O sistema de avaliação do ensino fundamental/SAEF e sua contribuição na gestão escolar pública municipal de Fortaleza* (Dissertação de mestrado). Universidade Federal do Ceará, Centro de Ciências Agrárias, Mestrado Profissional em Avaliação de Políticas Públicas, Fortaleza. https://repositorio.ufc.br/handle/riufc/77264

Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.

Kumar, A. (2022). *Hold-out Method for Training Machine Learning Models*. Vital Flux. https://vitalflux.com/hold-out-method-for-training-machine-learning-model/

Lessard, C., & Carpentier, A. (2016). *Políticas Educativas a aplicação na prática.* Petrópolis: Vozes.

Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, *2*(3), 18–22. https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf

Lima, M. A. M., Vasconcelos, S. O. S., Oliveira, S. de A. G., & Costa, F. C. C. da. (2021). Aspecto formativo dos dados das avaliações diagnósticas para os trabalhadores da gestão escolar nas escolas públicas de ensino médio do estado do Ceará. *Devir Educação*, *5*(2), 224–248. https://doi.org/10.30905/rde.v5i2.490

Lopes, A. F. N., Vieira, L. M. N., & Ramos, J. F. P. (2017). O SAEF como instrumento de avaliação educacional. *Revista Brasileira de Educação Básica*, *1*, 1–15. https://acortar.link/GdxuWm

Mendes, D. D. B., Maciel, A. de O., Araújo, A. A. C. de, & Amâncio, G. da S. P. (2023). Saef: avaliação e monitoramento da aprendizagem nas escolas municipais de Fortaleza, conhecer para intervir. *Revista De Instrumentos, Modelos E Políticas Em Avaliação Educacional*, *4*, e023008. https://doi.org/10.51281/impa.e023008

Oliveira, A., Costa, A., & Vidal, E. (2021). Avaliações municipais no Ceará: características e usos dos resultados. *Revista Meta: Avaliação*, *13*(39), 274-299. http://dx.doi.org/10.22347/2175-2753v13i39.3333

Probst, P., Wright, M. N., & Boulesteix, A. (2019). Hyperparameters and tuning strategies for random forest. *WIRES Data Mining and Knowledge Discovery*, *9*(3). https://doi.org/10.1002/widm.1301

Queiroga, E. M., Batista Machado, M. F., Paragarino, V. R., Primo, T. T., & Cechinel, C. (2022). Early Prediction of At-Risk Students in Secondary Education: A Countrywide K-12 Learning Analytics Initiative in Uruguay. *Information*, *13*(9), 401. https://doi.org/10.3390/info13090401

SEDUC. (n.d.). *Sistema Permanente de Avaliação da Educação Básica do Ceará – SPAECE*. https://www.seduc.ce.gov.br/spaece/

Silva, J. B. (2020). David Ausubel's Theory of Meaningful Learning: an analysis of the necessary conditions. *Research, Society and Development*, *9*(4), e09932803. https://doi.org/10.33448/rsd-v9i4.2803

SME. (n.d.). *Sistema de Avaliação do Ensino Fundamental. SAEF*. https://saef.sme.fortaleza.ce.gov.br/saef/pagina/alterar-unidade-trabalho.jsf

Vidal, E. M., & Costa, A. G. (Org.) (2021). *Responsabilização educacional no Ceará: trajetórias e evidências Organizadores*. 1ª Edição, Brasília, DF: ANPAE.

Vidal, E. M., Silva, J. B., Marinho, I. C., & Nogueira, J. F. F. (2024). Municipal assessments and the relationship with Ideb, according to Saeb contextual questionnaire 2019. *Práxis Educacional, 20*(51). https://doi.org/10.22481/praxisedu.v20i51.13559

Yang, S. J. H., Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Ogata, H., & Lin, A. J. Q. (2018). Predicting students' academic performance using multiple linear regression and principal component analysis. *Journal of Information Processing*, *26.* https://doi.org/10.2197/ipsjjip.26.170

Yang, Z., Wu, Y., Zhou, Y., Tang, H., & Fu, S. (2022). Assessment of Machine Learning Models for the Prediction of Rate-Dependent Compressive Strength of Rocks. *Minerals*, *12*(6), 731. https://doi.org/10.3390/min12060731

Creative Commons Attribution 4.0 International (CC BY 4.0)

Este artículo no presenta ningún conflicto de intereses. Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0). Se permite la reproducción, distribución y comunicación pública de la obra, así como la creación de obras derivadas, siempre que se cite la fuente original.