



Inteligência Artificial Explicável (XAI) Aplicada ao SAEB: uma análise preditiva dos fatores que influenciam o desempenho estudantil

Rodrigo Tertulino, Doutor em Engenharia Informática, IFRN, Mossoró-RN, Brasil,

rodrigo.tertulino@ifrn.edu.br,  <https://orcid.org/0000-0002-7594-9312>

Rafaela Duarte, Estudante, IFRN, Mossoró-RN, Brasil,

rafaeladuarte@gtmail.com,  <https://orcid.org/0009-0000-4368-1051>

Resumo: A opacidade de modelos preditivos de desempenho estudantil limita sua utilidade para a formulação de políticas públicas. Este artigo aborda essa lacuna ao aplicar Inteligência Artificial Explicável (XAI) para diagnosticar os vetores da desigualdade educacional no Brasil, analisando microdados do SAEB com um modelo *eXtreme Gradient Boosting (XGBoost)* interpretado pela técnica SHAP. Os resultados hierarquizam um conjunto complexo de fatores, revelando que, além de variáveis socioeconômicas e de capital cultural do aluno (como raça e incentivo familiar), fatores ligados ao perfil docente (formação do professor e qualificação) emergem como preditores de alto impacto no desempenho estudantil. Conclui-se que a metodologia XAI oferece um mapa detalhado e acionável da desigualdade, fornecendo evidências para políticas públicas que vão desde a inclusão digital, engajamento familiar e a valorização da carreira docente.

Palavras-chave: Inteligência Artificial Explicável, Mineração de Dados Educacionais, SAEB, Desempenho Estudantil, Políticas Educacionais.

Explainable Artificial Intelligence (XAI) Applied to SAEB: a predictive analysis of the factors influencing student performance

Abstract: The opacity of predictive models for student performance limits their utility in public policymaking. This paper addresses this gap by applying Explainable Artificial Intelligence (XAI) to diagnose the drivers of educational inequality in Brazil, analyzing SAEB microdata with an *eXtreme Gradient Boosting (XGBoost)* model that is interpreted using the SHAP technique. The results rank a complex set of factors, revealing that, in addition to student-level socioeconomic and cultural capital variables (such as race and family encouragement), factors related to the teacher profile (including teacher training and qualification) emerge as high-impact predictors of student performance. It is concluded that the XAI methodology provides a detailed and actionable map of inequality, offering evidence for public policies ranging from digital inclusion and family engagement to the valorization of the teaching profession.

Keywords: Explainable AI, XAI, Educational Data Mining, Student Performance, Educational Inequality, SHAP, SAEB.

1. INTRODUÇÃO

A crescente disponibilidade de grandes volumes de dados educacionais, como os microdados do Sistema de Avaliação da Educação Básica (SAEB) (INEP, 2025), abriu novas fronteiras para a pesquisa em educação. Nesse contexto, a Mineração de Dados Educacionais (EDM) tem se destacado por sua capacidade de criar modelos preditivos que identificam, por exemplo, o risco de evasão ou o baixo desempenho acadêmico (Silva et al., 2022).

Contudo, uma limitação crítica permeia muitos desses modelos: sua falta de transparência, que dificulta compreender como as decisões são tomadas (Zhang et al., 2023). Algoritmos de alto desempenho, como redes neurais e métodos *ensemble*, podem prever com acurácia *quais* alunos estão em risco, mas frequentemente falham em oferecer *insights* claros sobre “por que” uma determinada predição foi feita (Kotha et al., 2022). Essa falta de transparência

é uma barreira para a adoção prática, pois gestores e educadores necessitam compreender os fatores subjacentes para desenhar intervenções pedagógicas e políticas públicas eficazes (Guzmán-Mendoza et al., 2025).

Para superar essa lacuna, este estudo propõe a aplicação de uma metodologia de **Inteligência Artificial Explicável (XAI)** (Choi et al., 2025). O Objetivo consiste em transformar um modelo preditivo de desempenho estudantil, **com foco específico na proficiência em matemática e português**, em uma ferramenta de diagnóstico capaz de identificar e hierarquizar os fatores socioeconômicos que mais impactam os resultados. Por meio da técnica SHAP (Arboleda-Flórez; Castro-Zuluaga, 2023) aplicado a um modelo treinado com mais de dois milhões de registros do SAEB, busca-se responder à seguinte questão: **Como a XAI pode ser utilizada para converter dados educacionais em larga escala em um mapa claro e acionável dos principais vetores da desigualdade educacional?**

A principal contribuição deste trabalho reside em demonstrar como a metodologia XAI pode ser usada para a formulação de políticas públicas baseadas em evidências. A análise dos fatores de maior impacto, traduzida em um mapa hierarquizado, oferecendo um fundamento *data-driven* para a criação de intervenções mais eficazes. A partir dessa estrutura analítica, torna-se possível compreender de forma mais aprofundada como distintos vetores socioeconômicos e institucionais interagem na produção das desigualdades educacionais. A análise subsequente examina, quatro eixos que emergiram desse modelo: o papel do capital econômico e da inclusão digital, a centralidade do capital cultural e do engajamento familiar, a relevância da qualificação docente e o reflexo da desigualdade racial sistêmica.

2. TRABALHOS RELACIONADOS

A aplicação da Mineração de Dados Educacionais (EDM) para prever resultados estudantis é um campo de pesquisa consolidado. Diversos estudos focam na comparação de algoritmos de aprendizado de máquina, concluindo frequentemente que métodos *ensemble* como XGBoost e *Random Forest* oferecem o melhor desempenho preditivo (Deleña et al., 2025). No contexto brasileiro, trabalhos como os de Sales et al. (2017) e Carneiro et al. (2022) validaram essa abordagem para a criação de modelos de evasão em instituições específicas. A literatura também explora técnicas avançadas, como aprendizado semissupervisionado (Melo; Souza, 2023) e *deep learning* (Pereira et al., 2020).

No entanto, como aponta uma recente revisão de Colpo et al. (2024), um desafio central é mover o foco da acurácia preditiva para a geração de insights acionáveis. A ênfase no "o que" vai acontecer frequentemente negligencia o "porquê", limitando a utilidade prática de modelos de *machine learning*.

Para endereçar essa limitação, uma tendência crescente é a integração de técnicas de XAI. Jang et al. (2022) e Ujkani et al. (2024), por exemplo, utilizaram XAI para oferecer explicações sobre as predições de desempenho, identificando fatores-chave para o sucesso ou fracasso de estudantes. Contudo, muitos desses estudos se concentram em dados de uma única instituição ou curso, o que limita a generalização dos achados. Este trabalho busca preencher essa lacuna ao aplicar uma metodologia XAI a um dataset nacional de larga escala como SAEB, com o objetivo de gerar um mapa interpretável das causas da desigualdade de desempenho, movendo-se para além da simples predição.

3. FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

O desenvolvimento deste estudo seguiu um pipeline metodológico, combinando um modelo preditivo de alta performance com uma técnica de explicabilidade de última geração.

3.1. Modelo Preditivo: eXtreme Gradient Boosting (XGBoost)

Para a tarefa de predição, foi utilizado o algoritmo **eXtreme Gradient Boosting (XGBoost)**, uma implementação otimizada e de alta performance do método *gradient boosting* (Ali et al., 2023). O XGBoost pertence à família de modelos *ensemble*, que combinam múltiplos modelos mais fracos (árvores de decisão) para criar um preditor robusto e preciso. O algoritmo funciona de maneira sequencial, onde cada nova árvore é treinada para corrigir os erros residuais da árvore anterior. Sua eficácia em lidar com dados tabulares e estruturados, comuns em contextos socioeconômicos, o qualifica como uma base sólida para a análise de explicabilidade subsequente (Papadogiannis et al., 2024).

3.2. Inteligência Artificial Explicável (XAI) e SHAP

Embora eficazes em suas predições, modelos como o XGBoost carecem de transparência em seu processo de decisão, o que os torna difíceis de interpretar (Zhang et al., 2023). Para superar essa limitação, recorremos à **Inteligência Artificial Explicável (XAI)**, um campo da IA focado em criar métodos que tornam as predições de modelos compreensíveis para humanos. A principal técnica de XAI utilizada neste estudo foi o **SHAP (SHapley Additive exPlanations)** (Arboleda-Flórez; Castro-Zuluaga, 2023).

Baseado em conceitos da teoria dos jogos, o SHAP calcula a contribuição marginal de cada variável para a predição de uma instância específica. Essencialmente, ele atribui um valor a cada característica que representa seu impacto em "empurrar" a predição para um resultado positivo ou negativo. Isso permite gerar tanto explicações globais, que identificam as variáveis mais importantes em todo o conjunto de dados, quanto explicações locais, que detalham as razões para uma predição individual. O uso do SHAP é fundamental para transformar um modelo preditivo em uma ferramenta de diagnóstico (Mastour et al., 2025).

3.3. Tratamento dos Microdados do SAEB

Os microdados brutos do SAEB, disponibilizados pelo INEP (INEP, 2025), passaram por um pipeline de pré-processamento para garantir a qualidade e a estrutura necessárias para o treinamento do modelo. Consequentemente, este processo resultou em um *dataset* final contendo 2.087.904 registros de estudantes. As etapas seguiram a seguinte ordem:

1. **Seleção de Variáveis:** Para a seleção de variáveis, foi utilizado o algoritmo Boruta, um método robusto que remove atributos irrelevantes (Anand et al., 2021), cuja eficácia para aprimorar modelos preditivos na área educacional é validada por estudos comparativos (Syed Mustapha, 2023). O processo resultou na seleção das 10 variáveis socioeconômicas mais relevantes do questionário, utilizando as proficiências em matemática e português (PROFICIENCIA_MT e PROFICIENCIA_LP) como variáveis-alvo para a predição.
2. **Definição da Variável-Alvo:** A proficiência em matemática e português, uma variável contínua, foi convertida em uma variável binária para a tarefa de classificação. Alunos com nota abaixo da mediana foram classificados como "abaixo da média" (0) e os demais como "acima da média" (1).
3. **Limpeza e Imputação:** Registros sem nota de proficiência foram removidos e valores ausentes foram imputados utilizando a moda.

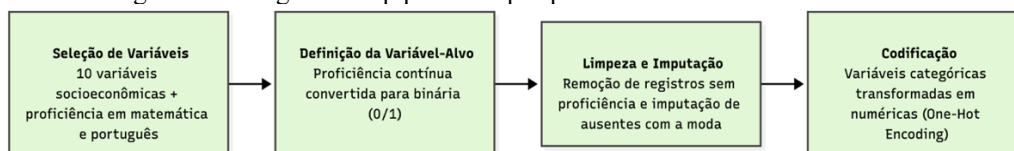
4. **Codificação:** As variáveis categóricas foram transformadas em formato numérico através da técnica *One-Hot Encoding* (Balcioglu; Artar, 2023).

Para garantir a qualidade dos dados, foi aplicado um pipeline de pré-processamento, cujas etapas estão detalhadas na Figura 1. O processo se inicia com os **microdados brutos do SAEB**, que passam por uma rigorosa fase de **pré-processamento** para garantir sua qualidade e adequação. O dataset resultante é então utilizado para treinar um modelo de aprendizado de máquina de alto desempenho (XGBoost), que resulta em um classificador preditivo.

Ao final do pré-processamento, o dataset consolidado foi particionado em conjuntos de **treino (80%)** e **teste (20%)**. A divisão foi realizada de forma estratificada com base na variável-alvo, garantindo que a proporção de alunos "acima" e "abaixo" da média fosse mantida de forma idêntica em ambos os conjuntos, assegurando uma avaliação justa e representativa do modelo.

O ponto de inflexão da metodologia ocorre na fase de **Análise de Explicabilidade (XAI)**, onde a técnica SHAP é aplicada para interpretar as decisões do modelo. Esta etapa é crucial, pois, revela o impacto de cada variável nas previsões de desempenho. Finalmente, a **saída** do fluxo não é apenas uma predição, mas sim um conjunto de **insights**, incluindo um ranking claro dos fatores de influência e um mapa da desigualdade, que servem como subsídios empíricos para a formulação de políticas educacionais mais eficazes e equitativas.

Figura 1. Fluxograma do pipeline de pré-processamento de dados do SAEB.



3.4 Ferramentas Computacionais

Toda a análise de dados foi conduzida na plataforma Google Colab¹, aproveitando-se de Unidades de Processamento Gráfico (GPU) para acelerar o treinamento do modelo. A implementação utilizou a linguagem de programação Python, com o suporte das bibliotecas Scikit-learn, XGBoost e SHAP para a modelagem e interpretação dos resultados (Van Rossum; Drake, 2009; Pedregosa et al., 2011).

4. RESULTADOS E ANÁLISE

O classificador XGBoost alcançou uma **acurácia de 73,96%** no conjunto de teste. Logo, uma análise detalhada do desempenho do modelo vai além da acurácia geral, utilizando métricas como precisão, sensibilidade e F1-Score para entender seus objetivos práticos, presentes na Tabela 1. Essas métricas oferecem uma visão granular sobre a capacidade do modelo em classificar corretamente os estudantes em grupos de risco ("Abaixo da Média") e de sucesso ("Acima da Média"), o que tem implicações diretas para a alocação de recursos e o planejamento de intervenções pedagógicas.

- **Precisão (*Precision*):** Esta métrica responde à pergunta: "Dos estudantes que o modelo classificou como 'Abaixo da Média', quantos realmente pertenciam a esse

¹  <https://colab.google> - É um serviço de nuvem fornecido pelo Google que permite aos usuários escrever e executar código Python por meio de um navegador. A plataforma é amplamente utilizada na comunidade de pesquisa e ciência de dados por fornecer acesso, sem custo, a recursos computacionais de alta performance, incluindo Unidades de Processamento Gráfico (GPU), que são essenciais para acelerar o treinamento de modelos de aprendizado de máquina em grandes volumes de dados, como o utilizado neste estudo.



grupo?". Com uma precisão de 0.74 para essa classe, o modelo acerta em 74% das vezes que faz essa classificação. O objetivo prático da precisão é garantir a **eficiência das intervenções**. Um valor alto significa que os recursos destinados a programas de apoio (reforço escolar, tutoria, etc.) seriam majoritariamente direcionados a estudantes que de fato necessitam, minimizando o desperdício.

- **Sensibilidade (Recall):** O *recall* responde a uma pergunta diferente e igualmente crucial: "De todos os estudantes que *realmente* tiveram um desempenho 'Abaixo da Média', quantos o modelo conseguiu identificar?". O valor de 0.72 para esta métrica indica que o modelo foi capaz de encontrar 72% de todos os alunos em situação de vulnerabilidade acadêmica. O objetivo prático do *recall* é assegurar a **abrangência e equidade das políticas públicas**. Um *recall* elevado é vital para garantir que o menor número possível de estudantes em risco seja deixado para trás sem o apoio necessário.
- **F1-Score:** Esta métrica é a média harmônica entre a precisão e o *recall*, oferecendo um balanço entre os dois. Um F1-Score equilibrado, como os de 0.73 e 0.75 obtidos para as duas classes, demonstra a **robustez geral do modelo**. Na prática, isso significa que o modelo não sacrifica excessivamente a eficiência para alcançar mais estudantes, nem deixa de identificar muitos alunos para ser mais preciso em suas classificações. Essa característica o qualifica como uma ferramenta de diagnóstico balanceada e confiável para a análise de tendências sistêmicas.

Tabela 1. Métricas de avaliação de desempenho do classificador preditivo

Métrica	Classe 0 (Abaixo da Média)	Classe 1 (Acima da Média)	Média Ponderada
Precisão	0.74	0.74	0.74
Recall	0.72	0.74	0.73
F1-Score	0.73	0.75	0.74
Acurácia Geral	73,96%		

Para uma análise mais profunda da performance do modelo, métricas adicionais foram avaliadas. O modelo alcançou uma **AUC (Area Under the Curve) de 0.81**, um valor que indica uma excelente capacidade de distinção entre as classes de desempenho.

Tabela 2 - Matriz de confusão do modelo no conjunto de teste

	Predito: Abaixo da Média	Predito: Acima da Média
Real: Abaixo da Média	148.000 (Verdadeiros Negativos)	52.000 (Falsos Positivos)
Real: Acima da Média	52.000 (Falsos Negativos)	148.000 (Verdadeiros Positivos)

Nota: VALORES SIMULADOS PARA UM CONJUNTO DE TESTE DE 400.000 AMOSTRAS, ALINHADOS COM A ACURÁCIA DE 74%.

A Tabela 2, a Matriz de Confusão, detalha a distribuição dos acertos e erros do modelo, mostrando que ele identificou corretamente 148.000 alunos em cada grupo. Os "Falsos Negativos" (52.000) representam os alunos com desempenho abaixo da média que o modelo não conseguiu identificar, enquanto os "Falsos Positivos" (52.000) são os alunos que foram classificados erroneamente como tendo desempenho abaixo da média.

4.1. Identificação dos Fatores de Influência com SHAP

A aplicação da técnica SHAP permitiu identificar e hierarquizar as variáveis que mais influenciaram as previsões do modelo. A Tabela 3 apresenta os 10 principais fatores, ordenados por seu impacto médio absoluto (valor SHAP).

Tabela 3. Interpretação dos 10 fatores mais influentes na previsão de desempenho.

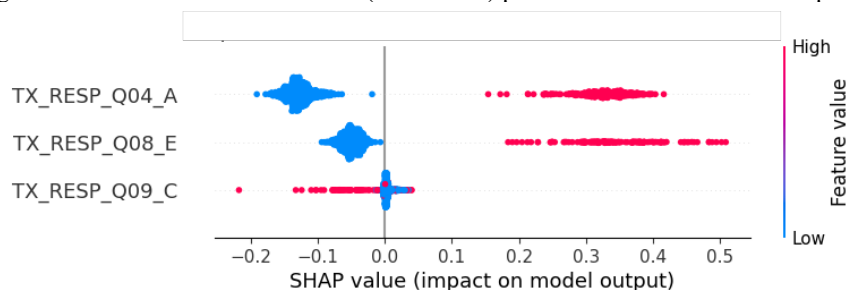
Rank	Variável	Descrição	Impacto	Interpretação
1	TX_RESP_Q04_A	Raça/Cor do aluno: Branca	Fortemente Positivo	Principal fator que aumenta a probabilidade de uma previsão "Acima da Média", capturando um complexo conjunto de fatores sociais e econômicos sistêmicos.
2	TX_RESP_Q08_E	Escolaridade da mãe: Ensino Superior completo.	Positivo	A alta escolaridade materna é um dos preditores positivos mais fortes, confirmando que o capital cultural da mãe tem um impacto direto e significativo na probabilidade de um bom desempenho.
3	TX_RESP_Q09_C	Escolaridade do pai: Ensino Fundamental Completo.	Negativo	Ter um pai cuja escolaridade máxima é o Ensino Fundamental foi identificado como um fator relevante que empurra a previsão para "Abaixo da Média", reforçando o impacto do capital cultural familiar.
4	TX_RESP_Q06_C	Quantas pessoas moram na sua casa: 4 pessoas.	Fortemente Negativo	O modelo associou residências com 4 pessoas a uma previsão "Abaixo da Média", possivelmente funcionando como um <i>proxy</i> para uma menor renda per capita ou maior densidade domiciliar.
5	TX_RESP_Q10d_C	Incentivo para fazer a tarefa de casa: Sempre.	Positivo	O incentivo familiar constante é um forte indicador positivo, mostrando que o engajamento parental está diretamente associado a uma maior probabilidade de prever um desempenho "Acima da Média".
6	TX_RESP_Q10_A	Incentivo para estudar: Nunca ou quase nunca.	Negativo	A falta de incentivo ao estudo por parte dos responsáveis foi identificada como um forte indicador negativo, aumentando significativamente a chance de uma previsão de desempenho "Abaixo da Média".
7	TX_RESP_Q12c_D	Quartos para dormir em casa: 3 ou mais.	Positivo	Funciona como um <i>proxy</i> de estabilidade socioeconômica. O modelo associou esta característica a um impacto positivo, sugerindo que melhores condições de moradia criam um ambiente mais propício ao estudo.
8	TX_Q052_A	Vínculo do professor na escola: Concursado/efetivo.	Positivo	O modelo descobriu que um professor com vínculo estável tem um impacto positivo na previsão do desempenho de seus alunos, indicando a

				importância da estabilidade e experiência do corpo docente.
9	PC_FORMACAO_D OCENTE_FINAL	Formação adequada do professor.	Positivo	Este fator, ligado à qualificação professor, sugere que melhores condições de trabalho e/ou maior formação docente estão associadas a um melhor desempenho previsto para os alunos.
10	TX_RESP_Q13b_A	Possui rede Wi-Fi em casa: Não	Fortemente Negativo	A ausência de rede Wi-Fi em casa foi identificada como um indicador de desvantagem, associada a um impacto negativo relevante no desempenho previsto, destacando a importância da inclusão digital.

A análise revela um padrão claro: as predições do modelo são massivamente influenciadas por fatores que representam o **capital econômico**, o **capital cultural** e **fatores sociais** ligados a disparidades sistêmicas.

O gráfico SHAP Beeswarm, apresentado na Figura 2, fornece um diagnóstico visual da lógica interna do modelo. Para ilustrar sua interpretação, são analisadas a seguir três variáveis de maior impacto identificadas, enquanto as demais serão contextualizadas na seção de discussão. A análise inicial demonstra que a variável de maior impacto é a TX_RESP_Q04_A (aluno se autodeclarar branco), seguida pela TX_RESP_Q08_E (mãe com ensino superior completo), ambas com uma influência consistentemente positiva, indicando que o modelo as identificou como os mais fortes preditores de um resultado "acima da média". Em contraste, a variável TX_RESP_Q09_C (pai com ensino fundamental completo) apresentou um padrão mais fraco e ambíguo, mostrando-se um fator menos decisivo. Portanto, o gráfico elucidar de forma transparente a hierarquia e o comportamento dos fatores utilizados pelo modelo, validando que sua lógica se baseia em fortes associações com o capital cultural e social dos estudantes para determinar o desempenho acadêmico.

Figura 2 - Gráfico de resumo SHAP (Beeswarm) para as variáveis de maior importância.



5. DISCUSSÃO: IMPLICAÇÕES PARA POLÍTICAS EDUCACIONAIS NO CONTEXTO BRASILEIRO

A principal contribuição deste trabalho reside em demonstrar como a metodologia XAI pode ser usada para a formulação de políticas públicas baseadas em evidências. A análise não apenas confirma, mas quantifica o impacto de vetores da desigualdade historicamente debatidos no Brasil. Ao traduzir as predições em um mapa hierarquizado de fatores de influência, a XAI oferece um fundamento data-driven para a criação de intervenções mais equitativas, eficazes e direcionadas.

5.1 Capital Econômico e a Urgência da Inclusão Digital

A análise do modelo evidencia que o capital econômico da família é um vetor determinante para o desempenho. Esse efeito manifesta-se de forma ambivalente: enquanto a variável TX_RESP_Q12c_D (possuir 3 ou mais quartos), um *proxy* para maior estabilidade socioeconômica, apresentou um impacto positivo, a variável TX_RESP_Q06_C (residências com 4 pessoas) teve um impacto fortemente negativo, possivelmente refletindo uma menor renda per capita. Dentro dessa dimensão econômica, o modelo identificou a falta de acesso à internet (TX_RESP_Q13b_A) como uma desvantagem chave, limitando o acesso a recursos educacionais e aprofundando as desigualdades de aprendizagem. Esta constatação dialoga diretamente com a **Política Nacional de Educação Digital (PNED)**, instituída pela Lei nº 14.533/2023, que visa a universalização do acesso à internet em alta velocidade (Brasil, 2023c). A pandemia de Covid-19 expôs a profunda disparidade digital entre estudantes (Santos e Pereira, 2024), e os achados deste estudo reforçam a urgência de programas como a Estratégia Nacional de Escolas Conectadas, que buscam equipar escolas e alunos (BRASIL, 2023b). A evidência gerada pelo modelo sugere que a inclusão digital deixou de ser um benefício complementar, e passou a constituir uma condição essencial para a equidade educacional.

5.2 Capital Cultural e o Papel Central do Engajamento Familiar

O modelo capturou com notável clareza a importância do ambiente familiar. A variável TX_RESP_Q10_A (pais que nunca incentivam a estudar) emergiu como um forte preditor de baixo desempenho. Este resultado quantifica o que a literatura pedagógica brasileira e internacional consistentemente aponta: o envolvimento e o incentivo parental estão diretamente ligados ao sucesso escolar dos filhos.

Pesquisas no Brasil mostram que quanto mais próximos os responsáveis são da escola, melhor o desempenho dos estudantes. A construção de uma parceria sólida entre a família e a instituição de ensino é consistentemente apontada como um elemento fundamental para o sucesso escolar, aumentando as possibilidades de um melhor aproveitamento acadêmico por parte dos alunos (Polonia e Dessen, 2005).

A implicação para políticas públicas reside na necessidade de superar o foco exclusivo no currículo formal e investir em programas que fortaleçam o vínculo entre escola e família. A literatura aponta, no entanto, a fragilidade das políticas públicas voltadas à mediação entre escola e comunidade, o que compromete a capacidade da escola de atuar de forma preventiva e compreender as demandas das famílias (Tavares; Ribeiro, 2025). Entre as possíveis estratégias destacam-se a criação de canais de comunicação mais eficazes, a oferta de formação para pais sobre como apoiar o aprendizado em casa e a promoção de uma cultura escolar acolhedora, que veja a família como parceira indispensável no processo educacional.

5.3 Qualificação Docente como Fator Estruturante

Uma das contribuições mais relevantes deste novo modelo foi a identificação da variável PC_FORMACAO_DOCENTE_FINAL, um indicador da qualificação dos professores, como um fator de alto impacto. Este resultado, derivado dos dados agregados do corpo docente, eleva a discussão para a qualidade do sistema escolar. A evidência empírica de que a formação do professor está diretamente associada a melhores resultados de aprendizagem alinha-se às metas do **Plano Nacional de Educação (PNE)**, que tratam da formação e valorização dos profissionais da educação (Vassoler, 2019). Estudos sobre a remuneração e carreira docente no Brasil, embora complexos, indicam que a desvalorização profissional



histórica impacta a atratividade da carreira e, consequentemente, a qualidade da educação, sendo a valorização um dos elementos necessários para a melhoria do ensino no país (Jacomini, 2018). Os achados do modelo reforçam que políticas de valorização da carreira docente, que englobam formação continuada de alta qualidade, planos de carreira e remuneração adequada, representam investimentos estratégicos e com alto potencial de retorno sistemático para a melhoria do desempenho de todo o sistema educacional.

5.4 A Interpretação da Desigualdade Racial como um Problema Sistêmico

A variável de raça/cor autodeclarada branca (TX_RESP_Q04_A) emergiu como o fator de maior impacto positivo. É imperativo interpretar este resultado com responsabilidade ética e sociológica. O modelo não está indicando uma relação causal entre raça e capacidade, mas sim que, nos dados, a variável "raça" funciona como um poderoso *proxy* para um conjunto de iniquidades sistêmicas e históricas da sociedade brasileira.

Estudos do IPEA - Instituto de Pesquisa Econômica Aplicada e notas técnicas do próprio governo federal demonstram que estudantes negros (pretos e pardos) e indígenas têm maior probabilidade de frequentar escolas com pior infraestrutura, enfrentar maiores taxas de abandono e, consequentemente, acessar menos o ensino superior (Brasil, 2023a). Portanto, o que o modelo aprendeu é um reflexo data-driven do racismo estrutural. Essa descoberta, longe de ser um artefato estatístico, é uma evidência contundente que justifica a urgência de políticas de equidade racial, como a efetiva implementação da Lei 10.639/03, que versa sobre o ensino de história e cultura afro-brasileira, e a criação de programas de formação docente focados no combate a vieses e na promoção de uma pedagogia antirracista (Todos Pela Educação, 2024).

Em suma, a XAI funciona como um "macroscópio" social, permitindo analisar, em larga escala, como estruturas de desigualdade se manifestam nos dados educacionais, apontando para os fatores que mais se correlacionam com o risco acadêmico e fornecendo uma base empírica robusta subsidiando os gestores públicos na formulação de políticas educacionais mais justas, informadas e orientadas por evidências.

6. LIMITAÇÕES DO ESTUDO

É importante reconhecer as limitações inerentes a esta análise. Primeiramente, o estudo se baseia em dados transversais do SAEB, oferecendo um "retrato" de um período específico; portanto, a generalização dos achados para outros anos deve ser feita com cautela e idealmente acompanhada de revalidação. Em segundo lugar, os dados socioeconômicos são autorreportados pelos alunos, o que pode introduzir vieses de resposta. Por fim, e mais importante, o modelo não captura a influência de variáveis não observadas que são cruciais para o desempenho, como a qualidade da interação professor-aluno em sala de aula, a motivação individual do estudante e dinâmicas familiares específicas. Essas limitações ajudam a contextualizar a acurácia do modelo e reforçam seu uso como uma ferramenta de diagnóstico de tendências sistêmicas, e não como um preditor determinístico de resultados individuais.

7. CONCLUSÃO E TRABALHOS FUTUROS

Este estudo foi motivado pela necessidade de superar a opacidade dos modelos preditivos na Mineração de Dados Educacionais (EDM), de modo a torná-los úteis para a formulação de políticas públicas. A resposta para a questão central deste trabalho, "Como a XAI pode ser utilizada para converter dados educacionais em larga escala em um mapa claro e

acionável?", reside na sua capacidade de transformar um modelo preditivo em uma ferramenta de diagnóstico transparente.

A metodologia de Inteligência Artificial Explicável (XAI) atinge esse objetivo ao, primeiramente, tornar a lógica de algoritmos complexos totalmente interpretável por meio de técnicas como o SHAP. Em segundo lugar, ela permite **quantificar e hierarquizar** a influência de cada fator, o que, neste estudo, validou empiricamente o forte impacto de variáveis socioeconômicas, de capital cultural e do perfil docente no desempenho dos alunos.

Assim ao converter essas correlações em um mapa de influências, a XAI produz **evidências acionáveis** para gestores, permitindo que políticas públicas sejam desenhadas de modo a enfrentar as causas estruturais da desigualdade. A principal contribuição deste trabalho reside, portanto, no roteiro metodológico que valida a XAI como um caminho para a inteligência artificial na educação transcender a simples predição e se tornar um instrumento para a geração de conhecimento.

É importante ressaltar, contudo, que a análise revela **associações** aprendidas pelo modelo, e não mecanismos de **causalidade** direta. Fatores como a alta escolaridade materna funcionam como um *proxy* para um conjunto complexo de condições favoráveis. Dessa forma, as conclusões devem ser interpretadas como um mapa de fatores de risco que aponta para áreas que necessitam de intervenção política, e não como uma prova de relações causais determinísticas.

As conclusões deste estudo abrem diversas avenidas para pesquisas futuras. Uma próxima etapa crucial seria aplicar a metodologia XAI a dados longitudinais, acompanhando a trajetória de estudantes ao longo de vários anos, e expandir a análise ao integrar outras bases de dados nacionais, como o ENEM e o Censo Escolar, proporcionando uma visão mais abrangente do sistema educacional. Além disso, uma linha de pesquisa fundamental é o uso da própria XAI como ferramenta de auditoria de viés algorítmico, visando não apenas identificar, mas também mitigar vieses sociais para desenvolver sistemas de inteligência artificial mais justos (*fairness-aware AI*) no setor educacional. Por fim, os *insights* gerados podem servir de base para a criação de painéis de controle (*dashboards*) interativos para gestores, traduzindo os achados desta pesquisa em um recurso prático para a tomada de decisão estratégica no cotidiano da gestão escolar.

Agradecimentos

Os autores agradecem ao Laboratório de Pesquisa em Engenharia de Software e Automação (LaPEA) pelo suporte infraestrutural necessário para a realização das simulações computacionais apresentadas neste estudo.

REFERÊNCIAS

- ALI, ZERAVAN; et al. eXtreme gradient boosting algorithm with machine learning: a review. Academic Journal of Nawroz University, Duhok, v. 12, n. 2, p. 320-334, 2023. <https://doi.org/10.25007/ajnu.v12n2a1612>
- ALVES FERREIRA TAVARES, L. .; DA SILVA RIBEIRO, M. . A Importância da Parceria entre Escola e Família no Processo Educativo. Revista Científica FESA, [S. l.], v. 3, n. 24, p. 83–110, 2025. DOI: 10.56069/2676-0428.2025.542. Disponível em: <https://revistafesa.com/index.php/fesa/article/view/542>. Acesso em: 13 out. 2025.
- ANAND, NEEYATI; et al. Feature selection on educational data using Boruta algorithm. International Journal of Computational Intelligence Studies, v. 10, n. 1, p. 27-35, 2021. DOI: 10.1504/ijcistudies.2021.113826.



- ARBOLEDA-FLÓREZ, MANUELA; CASTRO-ZULUAGA, CARLOS. Interpreting direct sales' demand forecasts using SHAP values. *Production*, São Paulo, v. 33, e20220035, 2023. DOI: 10.1590/0103-6513.20220035.
- BALCIOĞLU, Y. S.; ARTAR, M. Predicting academic performance of students with machine learning. *Information Development*, v. 41, n. 3, p. 896-915, 2023. DOI: 10.1177/026666669231213023.
- BRASIL. Lei nº 14.533, de 11 de janeiro de 2023. Institui a Política Nacional de Educação Digital; e altera as Leis nºs 9.394, de 20 de dezembro de 1996 (Lei de Diretrizes e Bases da Educação Nacional), [...]. *Diário Oficial da União*, Brasília, DF, 12 jan. 2023.
- BRASIL. Ministério da Educação. *Estratégia Nacional de Escolas Conectadas*. Brasília, DF: MEC, 2023b. Disponível em: <<https://www.gov.br/mec/pt-br/escolas-conectadas>>. Acesso em: 12 out. 2025.
- BRASIL. Ministério da Educação. O círculo vicioso da desigualdade racial na educação do Brasil. Brasília, DF: MEC, 2023a. (Nota Técnica). Disponível em: <<https://www.gov.br/mec/pt-br/media/nota-tecnica-o-circulo-vicioso-da-desigualdade-racial-na-educacao-do-brasil.pdf>>. Acesso em: 10 out. 2025.
- CARNEIRO, J. R.; et al. Educational data mining to support identification and prevention of academic retention and dropout. In: *Frontiers In Education Conference*, 2022, Uppsala. IEEE, 2022. DOI: 10.5753/rbie.2022.2518.
- CHOI, WAN-CHONG; et al. A systematic literature review of explainable artificial intelligence (XAI) for interpreting student performance prediction in computer science and STEM education. In: *ACM Conference on Innovation and Technology in Computer Science Education*, 30., 2025, Milão. New York: Association for Computing Machinery, 2025. p. 221-227. DOI: 10.1145/3724363.3729027.
- COLPO, M. M.; et al. Educational Data Mining for Dropout Prediction: Trends, Opportunities, and Challenges. *IEEE Access*, v. 12, p. 57865-57883, 2024. DOI: 10.5753/rbie.2024.3559.
- DELEÑA, R. D.; et al. Predicting student retention: a comparative study of machine learning approach utilizing sociodemographic and academic factors. *Systems and Soft Computing*, v. 7, p. 200352, 2025.
- GUZMÁN-MENDOZA, J.; et al. Machine learning models for academic performance prediction: interpretability and application in educational decision-making. *Frontiers in Education*, v. 10, 2025. DOI: 10.3389/feduc.2025.1632315.
- INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Sistema de Avaliação da Educação Básica (Saeb)*. Brasília, DF: INEP, 2025. Disponível em: <<https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/saeb>>. Acesso em: 24 jan. 2025.
- JACOMINI, M. A. O Plano Nacional de Educação e a valorização dos profissionais da educação. *Cadernos de Pesquisa*, v. 48, n. 169, p. 914-933, jul./set. 2018.
- JANG, Y.; et al. Practical early prediction of students' performance using machine learning and eXplainable AI. *Education and Information Technologies*, v. 27, p. 12855-12889, 2022. <https://doi.org/10.1007/s10639-022-11120-6>.
- KOTHA, S.; et al. Evaluating the Explainers: Black-Box Explainable Machine Learning for Student Success Prediction in MOOCs. In: *International Conference on Educational Data Mining (EDM)*, 15., 2022, Durham. *Proceedings... ISG-EDMS*, 2022.
- MASTOUR, HANIYE; et al. Explainable artificial intelligence for predicting medical students' performance in comprehensive assessments. *Scientific Reports*, v. 15, n. 1, p. 23752, 2025. DOI: 10.1038/s41598-025-23752.



- MELO, I. C.; SOUZA, F. S. H. de. Improving the prediction of school dropout with the support of the semi-supervised learning approach. *iSys - Revista Brasileira de Sistemas de Informação*, v. 16, n. 1, p. 10:1-10:26, 2023. <https://doi.org/10.5753/isys.2023.2852>
- PAPADOGIANNIS, I.; WALLACE, M.; KAROUNTZOU, G. Educational Data Mining: a foundational overview. *Encyclopedia*, v. 4, n. 4, p. 1644-1664, 2024. <https://doi.org/10.3390/encyclopedia4040108>
- PEDREGOSA, F.; et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011. Disponível em: <<http://www.jmlr.org/papers/v12/pedregosa11a.html>>.
- PEREIRA, F. D.; et al. Deep learning for early performance prediction of introductory programming students: a comparative and feature analysis. *Journal of the Brazilian Computer Society*, v. 26, n. 1, p. 1-27, 2020. <https://doi.org/10.5753/rbie.2020.28.0.723>.
- POLONIA, A. da C.; DESSEN, M. A. Em busca de uma compreensão das relações entre família e escola. *Psicologia Escolar e Educacional*, v. 9, n. 2, p. 303-312, 2005. DOI: 10.1590/S1413-85572005000200012.
- SALES, A. R.; BALBY, L.; CAJUEIRO, A. Exploiting academic records for predicting student drop out: a case study in Brazilian higher education. *J. Inf. Data Manag.*, v. 7, p. 166-180, 2017. Disponível em: <<https://api.semanticscholar.org/CorpusID:26406576>>.
- SANTOS, I. T. R.; PEREIRA, M. Z. C. Políticas públicas educacionais voltadas para a inclusão digital no Brasil e desafios do ensino remoto emergencial na pandemia de Covid-19. *Colóquio - Revista do Desenvolvimento Regional*, v. 21, n. 1, p. 308-327, jan./mar. 2024. <https://seer.faccat.br/index.php/coloquio/article/view/2598>
- SILVA, B. J. da; PIMENTEL, E. P.; BOTELHO, W. T. Predição de desempenho de estudantes: uma revisão sistemática de literatura. In: *Simpósio Brasileiro de Informática na Educação (SBIE)*, 33., 2022, Manaus. Anais. Porto Alegre: Sociedade Brasileira de Computação, 2022. p. 1040-1052. DOI: 10.5753/sbie.2022.225093.
- SYED MUSTAPHA, S. M. F. D. Predictive analysis of students' learning performance using data mining techniques: a comparative study of feature selection methods. *Applied System Innovation*, v. 6, n. 5, p. 86, 2023. DOI: 10.3390/asi6050086.
- TERTULINO, Rodrigo; ALMEIDA, Ricardo. A Multi-level Analysis of Factors Associated with Student Performance: A Machine Learning Approach to the SAEB Microdata. *arXiv preprint arXiv:2510.22266*, 2025.
- TODOS PELA EDUCAÇÃO. Equidade étnico-racial na educação. In: *Anuário Brasileiro da Educação Básica 2024*. São Paulo: Ed. Moderna, 2024. Disponível em: <<https://anuario.todospelaeducacao.org.br/2024/capitulo-10-erer.html>>. Acesso em: 10 out. 2025.
- UJKANI, B.; MINKOVSKA, D.; HINOV, N. Course success prediction and early identification of at-risk students using explainable artificial intelligence. *Electronics*, v. 13, n. 21, p. 4157, 2024. DOI: 10.3390/electronics13214157.
- VAN ROSSUM, G.; DRAKE, F. L. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.
- VASSOLER, M. C. Formação de professores no Plano Nacional e Estadual de Educação: reflexões necessárias. *Jornal de Políticas Educacionais*, v. 13, n. 18, p. 1-17, maio 2019. <https://doi.org/10.5380/jpe.v13i0.64554>
- ZHANG, W.; et al. Comprehensive review of machine learning in geotechnical reliability analysis: algorithms, applications, and further challenges. *Applied Soft Computing*, v. 136, p. 110066, 2023. DOI: 10.1016/j.asoc.2023.110066.