

Árvore de Decisão Aplicada ao Estudo do Desempenho SAEB em Matemática de Estudantes do Amapá

Carla D. S. Vasconcelos^{1†}, Pedro dos S. Costa Junior¹, Edcarlos V. da Silva¹

¹ Universidade Federal do Amapá (UNIFAP); Macapá – Amapá, Brasil.

Resumo: *O SAEB é uma avaliação nacional que mede o desempenho dos alunos em Português e Matemática na Educação Básica. Os resultados do SAEB fornecem dados para o monitoramento e aprimoramento da qualidade da educação no país. O Grupo de Pesquisa em Educação Matemática do Amapá – GEMAP, vinculado a Universidade Federal do Amapá iniciou uma série de estudos sobre o SAEB no que tange ao rendimento dos alunos em Matemática e em Português com foco no Amapá e na região Norte do Brasil. O objetivo do presente recorte de estudo é o de estimar perfis de estudantes do Amapá abaixo da média nacional que participaram da prova do SAEB. Como metodologia foi utilizado o algoritmo de árvore de classificação, método CHAID, com base amostral de 9462 alunos do Amapá, ano de 2019, o banco de dados foi baixado do site INEP. Observou-se que 68,5% dos alunos estão abaixo da média nacional (277,3 pontos em Matemática no ano de 2019), quanto aos perfis, a árvore de classificação indicou que os perfis com maiores probabilidades de estarem abaixo da média são alunos de “escolas públicas que moram em área rural” com 83,7%. O estudo contribuiu para mostrar quais os perfis com maior probabilidade de terem notas abaixo da média em matemática e assim contribuir na formulação de políticas públicas educacionais.*

Palavras-chave: SAEB; Árvore de Classificação; Amapá; Ensino Médio.

Decision Tree Applied to the Study of SAEB Performance in Mathematics of Students from Amapá

Abstract: *SAEB is a national assessment that measures student performance in Portuguese and Mathematics in Basic Education. SAEB results provide data for monitoring and improving the quality of education in the country. The Amapá Mathematics Education Research Group – GEMAP, linked to the Federal University of Amapá, began a series of studies on SAEB regarding student performance in Mathematics and Portuguese with a focus on Amapá and the Norte region of Brazil. The objective of this study is to estimate profiles of students from Amapá below the national average who participated in the SAEB test. As a methodology, the classification tree algorithm, CHAID method, was used, with a sample base of 9462 students from Amapá, year 2019, the database was downloaded from the INEP website. It was observed that 68.5% of students are below the national average (277.3 points in Mathematics in 2019), regarding profiles, the classification tree indicated that the profiles with the highest probability of being below average are students from “public schools that live in an area rural” with 83.7%. The study helped to show which profiles are most likely to have below average grades in mathematics and thus contribute to the formulation of public educational policies.*

Keywords: SAEB; Classification Tree; Amapá; High school.

[†] Autor correspondente: carladominique4@gmail.com

Manuscrito recebido em: 08/06/2024

Manuscrito revisado em: 09/10/2024

Manuscrito aceito em: 09/10/2024

Introdução

O Sistema de Avaliação da Educação Básica (SAEB), administrado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), avalia a qualidade do ensino nas escolas públicas e privadas do Brasil por meio de testes padronizados de língua portuguesa e matemática. Esses testes são aplicados regularmente aos alunos do 5º e 9º anos do Ensino Fundamental e do 3º ano do Ensino Médio. Além das provas, o SAEB coleta informações sobre o contexto socioeconômico e educacional dos estudantes, bem como sobre a infraestrutura das escolas, permitindo uma análise abrangente dos fatores que influenciam o desempenho dos alunos e auxiliando na formulação de políticas públicas. Em 2019, a média nacional na prova de Matemática foi de 277,3 pontos (INEP, 2024).

Uma das análises que se pode fazer com os resultados disponíveis da prova do SAEB é o estudo do rendimento dos alunos. Com a introdução do Sistema de Avaliação da Educação Básica (SAEB), a ideia de que, além dos indicadores de rendimento, uma medida de aprendizagem dos alunos deveria ser usada para o monitoramento do sistema foi ganhando espaço na gestão pública educacional (Soares; Xavier, 2013). Considerando os dados do SAEB, ao cruzar os dados da prova de Matemática ou Língua Portuguesa com variáveis sociodemográficas dos estudantes, é possível conhecer quais perfis possuem as maiores e as menores notas nas respectivas disciplinas. Desta forma, disparidades podem ser identificadas, revelando os grupos mais vulneráveis ao insucesso escolar. Nesta direção, o estudo proposto objetivou uma análise no desempenho escolar de alunos do Ensino Médio do Amapá com foco na prova de Matemática, com o intuito de verificar quais perfis de estudantes possuem maior probabilidade de se situarem abaixo da nota média nacional do SAEB referente ao ano de 2019.

Para implementar este estudo, foram baixados do site do INEP o banco de dados (microdados) referente ao ano de 2019, de onde foram filtradas as informações dos estudantes do Amapá incluindo a nota numérica da prova de Matemática e suas variáveis sociodemográficas para análises. Foram realizadas estatísticas descritivas, bem como o uso do algoritmo computacional de árvore de decisão (Pestana; Gageiro, 2009) para calcular as probabilidades de os estudantes terem um perfil com nota abaixo da média nacional no referido ano.

Materiais e Métodos

Este estudo consistiu em uma pesquisa de abordagem quantitativa (Silva; Menezes, 2005) com uso de banco de dados secundários (LAPEI, 2024). Os dados foram baixados diretamente do site do INEP, e corresponderam ao banco de dados do SAEB do ano de 2019, disponibilizados gratuitamente pela plataforma para que pesquisadores possam ter acesso e desenvolver estudos.

Optou-se pela modalidade de dados conhecida como microdados, que reúnem um conjunto de informações detalhadas relacionadas às pesquisas, aos exames e avaliações do INEP, filtrado para as provas de Matemática do SAEB, com foco nos estudantes do Amapá, esses microdados são informações individuais dos alunos do Ensino Médio que fizeram a prova no ano de 2019 (INEP, 2024).

O objetivo deste estudo foi analisar as notas de Matemática dos alunos, estimando a probabilidade de um estudante situar-se abaixo da média nacional em Matemática do SAEB em 2019. Nesse ano, a média nacional na prova de Matemática foi de 277,3 pontos, enquanto na prova de Língua Portuguesa foi de 278,4 pontos (INEP, 2024). Utilizando essas referências, classificou-se

os estudantes do Amapá, do Ensino Médio, em relação à sua nota pessoal e à média nacional, conforme a seguinte categorização:

- **Aluno Abaixo da Média Nacional:** Nota de Matemática abaixo de 277,3 pontos.
- **Aluno Igual ou Acima da Média Nacional:** Nota de Matemática igual ou acima de 277,3 pontos.

Esses dados foram cruzados com variáveis sociodemográficas dos alunos como: zona de residência (capital e interior), área de residência (urbana e rural), sexo (masculino e feminino), acesso a rede Wi-fi (sim e não), raça/cor (branco, pardo/preto e indígena).

Descrição do Método Estatístico

Os dados foram organizados e tratados em planilha Excel e em seguida foram exportados para o pacote estatístico SPSS para fins de implementação do algoritmo de Árvore de Decisão. A árvore é uma rotina pré-implementada no programa, cabendo apenas a identificação da variável resposta (nota de matemática categorizada) e as variáveis explicativas (tipo de escola, local de residência, uso de notebook, acesso a wifi em casa, raça, acesso a livros, acesso a notícias).

O cruzamento de dados foi feito usando o algoritmo Árvore de Decisão, modelo Chi-Square Automatic Interaction Detector – CHAID, que analisa interações de uma variável resposta Y (Notas dos alunos em Matemática) com prováveis variáveis explicativas X (sexo, raça/cor, local de residência, etc) na forma de probabilidades (Pestana; Gageiro, 2009). O modelo identifica um perfil com base em todas as variáveis incluídas e estabelece uma probabilidade de ocorrência daquele perfil com base na variável resposta.

Assim, método CHAID foi programado para analisar qual perfil de aluno do Ensino Médio do Amapá tem maior probabilidade de estar abaixo da média nacional na prova de Matemática do SAEB, revelando assim os perfis de maior vulnerabilidade para a prova. O algoritmo foi implementado via pacote estatístico SPSS versão 26.

A população em estudo consistiu no censo total de estudantes do Amapá do Ensino Médio que participaram da prova em 2019, totalizando 9.462 estudantes. Desse total, o algoritmo CHAID seleciona entre 60% a 70% da amostra para treinamento do algoritmo e 40% a 30% da amostra para efetivamente estimar as árvores de decisão com os perfis de probabilidades.

As árvores foram apresentadas conforme a saída gráfica do modelo matemático, na forma de caixas e nós. Como o estudo foi feito com base em dados secundários, disponibilizados de forma pública na internet, não foi necessário protocolar registro da pesquisa em Comitê de Ética em Pesquisa, uma vez que os dados secundários, embora apresentem informações individuais, não há identificação pessoal de cada aluno e a pesquisa também não trabalhou diretamente com pessoas, sendo dispensado o registro do projeto de pesquisa em um órgão de fiscalização.

Resultados e Discussões

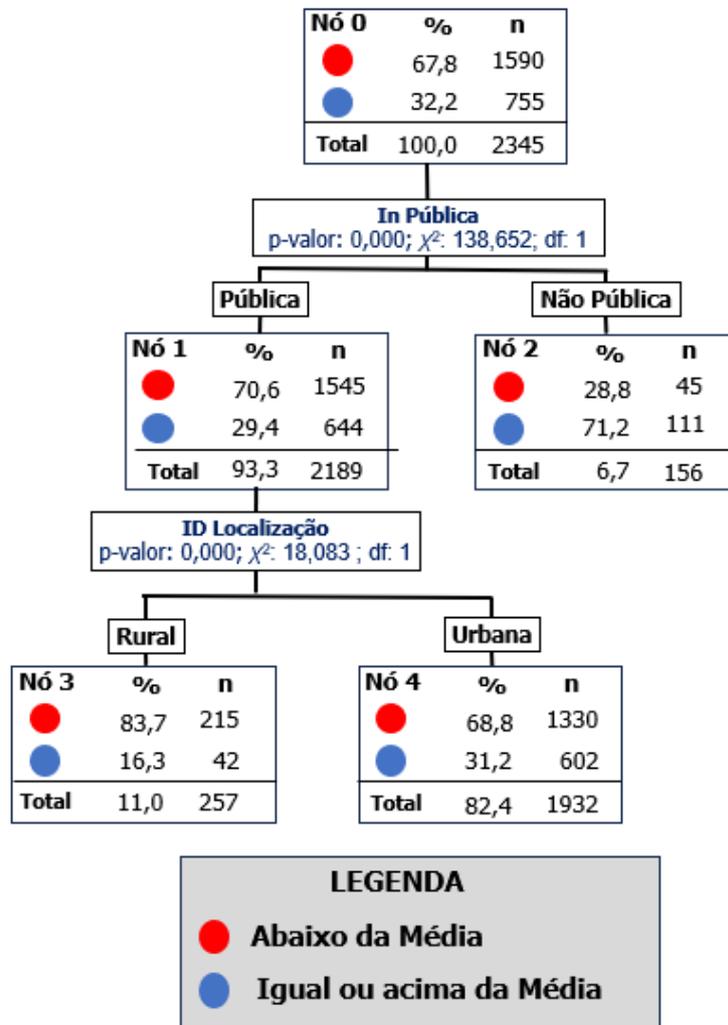
Neste estudo, aplicamos uma árvore de decisão para analisar variáveis sociodemográficas e seu impacto nas notas de Matemática dos alunos. Ao explorar essas variáveis, buscamos identificar padrões e relações que possam influenciar o desempenho dos alunos em Matemática, comparando suas notas com a média nacional do SAEB de 2019. A estrutura da árvore de decisão resultante permitirá uma análise clara dos fatores sociodemográficos mais determinantes para o desempenho acadêmico.

A primeira árvore de decisão apresenta 5 nós, sendo 3 deles considerados nós terminais (2, 3 e 4) a amostra incluída na árvore teve tamanho $n = 2345$ e analisou a proficiência dos estudantes (abaixo da média e igual ou acima da média) com significância para as seguintes variáveis: tipo de escola (pública ou privada) e localização de residência (zona urbana e zona rural). A Figura 1 apresenta a árvore com os perfis de probabilidades para identificar estudantes do Amapá com mais chances de estarem abaixo da média nacional em Matemática na prova do SAEB.

Os perfis identificados foram:

- a) Estudantes da rede pública rural: possuem 83,7% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- b) Estudantes da rede pública urbana: possuem 68,8% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- c) Estudantes de escola não pública: possuem 28,8% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;

Figure 1: Probability tree for sociodemographic variables of students who participated in the 2019 SAEB Mathematics exam.



Source: from the authors (2024).

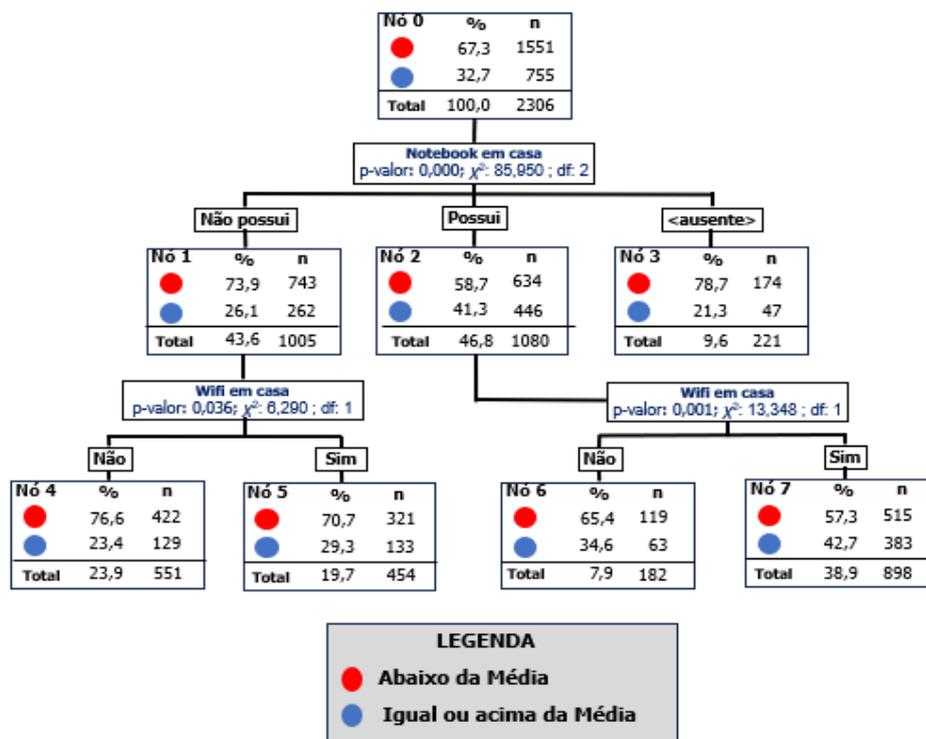
Na Figura 1, observa-se que estudantes que estudam e residem em escolas públicas localizadas em áreas rurais possuem uma probabilidade de 83,7% de obterem uma nota abaixo da média nacional em Matemática na prova do SAEB. Em contraste, aqueles que estudam em escolas privadas na área urbana apresentam uma maior probabilidade de estarem igual ou acima da média nacional. Esta análise evidencia a influência significativa das variáveis sociodemográficas no desempenho acadêmico dos estudantes, destacando a disparidade entre diferentes contextos educacionais e geográficos.

A segunda árvore de decisão apresenta 8 nós sendo 4 deles considerados nós terminais (4, 5,

6 e 7) a amostra incluída na árvore teve tamanho $n = 2306$ e analisou a proficiência dos estudantes (abaixo da média e igual ou acima da média) com significância para as seguintes variáveis: Notebook em casa (possui e não possui) Wi-Fi em casa (sim e não). A Figura 2 apresenta os perfis de probabilidades para identificar estudantes do Amapá com maior probabilidade de estarem abaixo da média nacional em Matemática na prova do SAEB. Os perfis identificados foram:

- a) Estudantes que não possuem Notebook e não tem Wi-Fi em casa: possuem 76,6% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- b) Estudantes que não possuem Notebook e tem Wi-Fi em casa: possuem 70,7% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- c) Estudantes que possuem Notebook e não tem Wi-Fi em casa: possuem 65,4% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- d) Estudantes que possuem Notebook e tem Wi-Fi em casa: possuem 57,3% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;

Figure 2: Probability tree for technology-related variables of students who participated in the 2019 SAEB Mathematics exam.



Source: from the authors (2024).

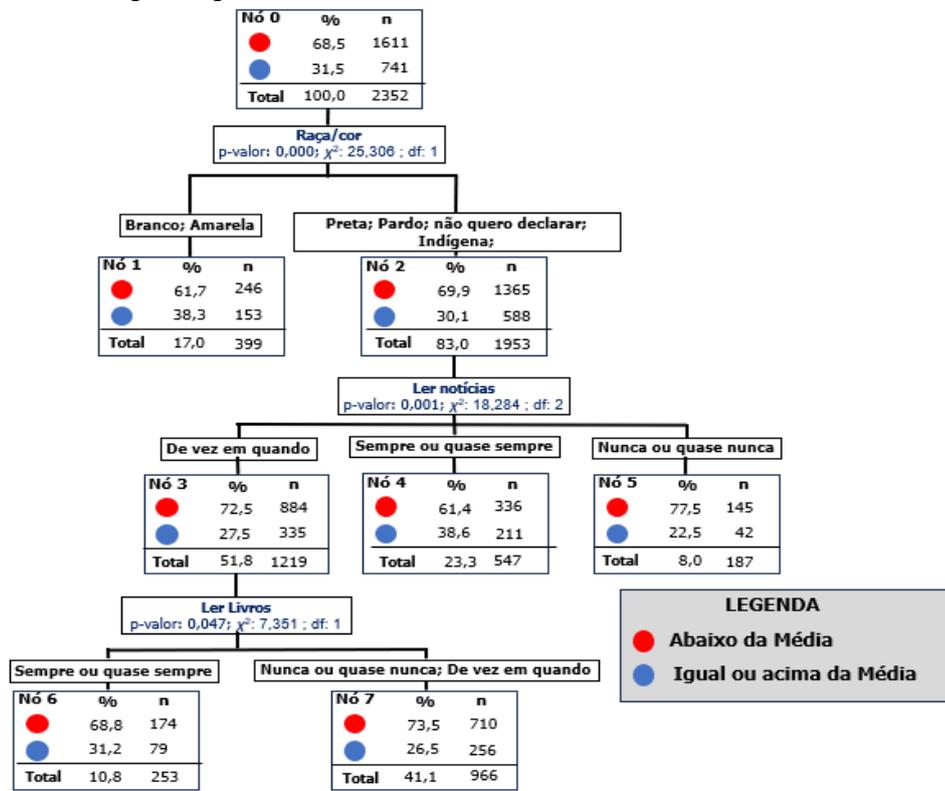
Na Figura 2, observa-se que estudantes que não possuem um notebook e não têm acesso a Wi-Fi em casa apresentam uma probabilidade de 76,6% de obterem uma nota abaixo da média nacional em Matemática na prova do SAEB. Esta análise ressalta a importância do acesso à tecnologia como um fator crucial para o desempenho acadêmico. A falta de recursos tecnológicos adequados pode limitar significativamente as oportunidades de aprendizado e o acesso a materiais didáticos de qualidade, resultando em um desempenho inferior nas avaliações educacionais.

A terceira árvore de decisão apresenta 8 nós sendo 5 deles considerados nós terminais (1, 4, 5, 6 e 7) a amostra incluída na árvore teve tamanho $n = 2352$ e analisou a proficiência dos estudantes (abaixo da média e igual ou acima da média) com significância para as seguintes variáveis: raça/cor (branco/amarela e pardo/preta/indígena), ler notícias (nunca, de vez em quando ou sempre) e ler livro (nunca, de vez em quando ou sempre). A Figura 03 apresenta os perfis de probabilidades para identificar estudantes do Amapá com mais chances de estarem abaixo da média nacional em Matemática na prova do SAEB. Os perfis identificados foram:

- a) Estudantes de raça/cor (preta; pardo; não quero declarar; indígena) que nunca leem notícias: possuem 77,5% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- b) Estudantes de raça/cor (preta; pardo; não quero declarar; indígena) que de vez em quando leem notícias e de vez em quando ou nunca leem livros: possuem 73,5% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- c) Estudantes de raça/cor (preta; pardo; não quero declarar; indígena) que de vez em quando leem notícias e sempre leem livros: possuem 68,8% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- d) Estudantes de raça/cor (branco/amarela): possuem 61,7% de probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;
- e) Estudantes de raça/cor (preta; pardo; não quero declarar; indígena) que sempre leem notícias: possuem 61,4% probabilidade de estarem abaixo da média nacional na prova de Matemática SAEB;

Na Figura 3, observa-se que estudantes de raça/cor (preta; pardo; não quero declarar; indígena) que nunca leem notícias apresentam uma probabilidade de 77,5% de obterem uma nota abaixo da média nacional em Matemática na prova do SAEB. Esta análise evidencia a influência combinada de fatores culturais e comportamentais no desempenho acadêmico.

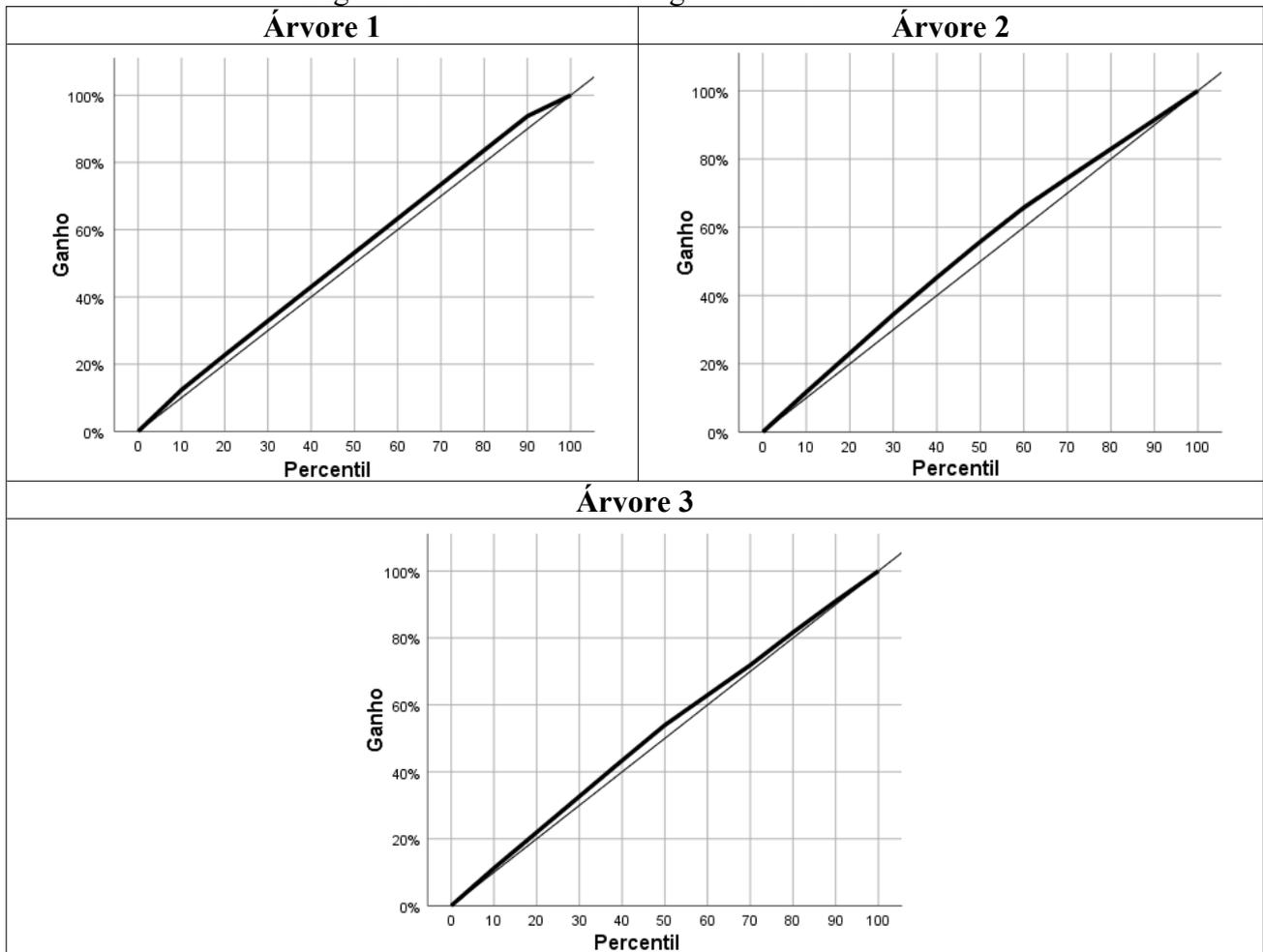
Figure 3: Probability tree for the race/color variable and cultural variables of students who participated in the 2019 SAEB Mathematics exam.



Source: from the authors (2024).

A ausência do hábito de ler notícias pode refletir um menor engajamento com atividades que promovem o pensamento crítico e a leitura, habilidades essenciais para o sucesso em diversas áreas do conhecimento, incluindo Matemática. Além disso, o fator racial pode estar associado a desigualdades estruturais mais amplas que afetam o acesso a recursos educacionais e oportunidades de aprendizado. A Figura 4 a seguir apresenta o conjunto de gráficos de percentis e ganhos proposta por Pestana e Gageiro (2009). Cada árvore gerada pelo modelo é representada em um gráfico individual, permitindo uma avaliação visual da capacidade de previsão das probabilidades de desempenho dos alunos. Conforme sugerido pelos autores, a posição da curva em relação à linha diagonal indica o grau de precisão do modelo: quanto mais acima da linha, maior a precisão. Todos os modelos propostos para o estudo apresentam capacidade preditiva nos dados.

Figure 4: Presentation of the gain charts for all trees.



Source: from the authors (2024).

De acordo com Pestana e Gageiro (2009), o modelo de árvore de decisão CHAID apresenta algum risco de classificação incorreta da amostra total de estudantes. Esse risco é medido em porcentagem e está indicado na coluna de Intervalo de Confiança IC da Tabela 1. Desta forma, para a árvore 1, o modelo tem o risco de classificar incorretamente cerca de 27% a 31% dos alunos que compõem a amostra total. A mesma interpretação deve-se fazer para os demais intervalos das demais árvores.

Table 1: Risk Estimates and Confidence Interval for Prediction Errors.

Árvore	Risco	Erro	%	IC
1	0,294	0,009	29,4%	[27%; 31%]
2	0,327	0,010	32,7%	[30%; 34%]
3	0,315	0,010	31,5%	[29%; 33%]

Source: from the authors (2024).

A análise de Estimativas de Riscos e Intervalo de Confiança para erros de previsão é um indicador importante para avaliar a precisão e a confiabilidade dos modelos estatísticos utilizados.

Ao calcular os intervalos de confiança, consegue-se determinar a margem de erro das previsões, enquanto as estimativas de risco destacaram possíveis áreas de vulnerabilidade nos modelos. Essas análises forneceram uma base sólida para validar as previsões e identificar oportunidades de melhorias, aprimorando significativamente a qualidade da análise preditiva no contexto educacional.

Na literatura pertinente para comparação não foram localizados estudos que utilizam estudos de probabilidades para estimar perfis abaixo da média ou acima da média na prova do SAEB, no entanto, existem pesquisas com objetivos semelhantes e utilização de métodos quantitativos variados para identificar tais perfis, estudos como o de Diaz (2012) que analisaram desigualdades entre estudantes de escolas públicas e privadas do Ensino Médio, com base nas notas de Português e Matemática do SAEB são uma referência no que tange aos estudos de rendimento escolar comparando grupos de variáveis.

No estudo de Diaz (2012) o autor empregou um índice de igualdade de oportunidades entre estudantes de escolas públicas e privadas e os resultados mostraram que as desigualdades de oportunidades entre estudantes de escolas públicas e privadas foram menores em Língua Portuguesa, cuja diferença se situou em torno de 16% da desigualdade total e, em Matemática essa diferença foi próxima de 24%, este estudo tem correlação com os resultados de nossa pesquisa em função de avaliar o rendimento escolar de estudantes com base em notas de Língua Portuguesa e Matemática

Nas pesquisas de Gusmão e Amorim (2022) os autores objetivaram examinar a relação entre o desempenho escolar e os indicadores de desigualdade social e econômica dos estados brasileiro utilizando as notas do SAEB em Língua Portuguesa e Matemática com foco nos alunos de Ensino Médio. Os autores concluíram que, existe correlação positiva entre os PIB estaduais e o rendimento escolar em Língua Portuguesa e em Matemática, quando analisado o recorte temporal de 2013 a 2015, inferindo que, estados com elevado PIB possuem as notas mais elevadas em Língua Portuguesa e em Matemática. Observou-se resultado semelhante para o Índice de Desenvolvimento Humano - IDH, pois estados com melhores IDH também são aqueles que possuem os melhores rendimento de Língua Portuguesa e Matemática no SAEB. Estes estudos também coadunam com os resultados de nosso estudo em função de reforçar a ideia de que região mais pobres possuem desempenhos piores na prova SAEB.

Assim, o estudo ora apresentado neste artigo traz grandes contribuições para a literatura, especialmente no contexto da Região Norte do Brasil e ressalta a importância do SAEB como um sistema de grande relevância para avaliação educacional nacional. No entanto, é fundamental transformar seus resultados em medidas políticas concretas, baseadas nos diagnósticos fornecidos por essa e por demais pesquisas que visem expor as desigualdades dos estudantes de Ensino Médio em relação ao desempenho em Língua Portuguesa e em Matemática.

Conclusões

O presente estudo teve como objetivo geral analisar o desempenho escolar dos alunos do Ensino Médio do Amapá na disciplina de Matemática, com foco na identificação dos perfis de estudantes com maior probabilidade de apresentar notas abaixo da média nacional do SAEB referente ao ano de 2019. Para atingir esse propósito, foi realizado o estudo dos microdados disponibilizados pelo INEP para o referido ano e selecionadas as informações pertinentes aos alunos do Amapá. Posteriormente, foram conduzidas análises estatísticas descritivas e aplicado um

algoritmo de árvore de decisão, possibilitando a identificação dos fatores associados ao desempenho acadêmico desses estudantes.

Estudar o desempenho dos alunos do Ensino Médio, especialmente em disciplinas como Matemática, é de relevância para o Amapá. Com desafios como falta de recursos educacionais e desigualdades socioeconômicas, entender as causas do baixo desempenho escolar permite criar políticas educacionais mais eficazes. Isso não só ajuda os alunos a superar obstáculos específicos, mas também promove um desenvolvimento mais equitativo e fortalece as políticas públicas do estado ao melhor capacitar a população para os desafios socioeconômicos do futuro. Este estudo representou uma contribuição significativa para o contexto educacional do Amapá, uma vez que aborda uma lacuna de conhecimento pouco explorada na região.

O foco na análise do desempenho dos alunos do Ensino Médio em Matemática e a identificação de perfis com maior probabilidade de obter notas abaixo da média nacional do SAEB em 2019 são de suma importância para o estado. Ao preencher essa lacuna, o estudo não apenas fornece uma compreensão mais profunda dos desafios enfrentados pelos estudantes amapaenses, mas também contribui significativamente para a geração de conhecimento local. As informações e resultados gerados por esta pesquisa podem servir de base para a formulação de políticas educacionais mais direcionadas e eficazes, adaptadas às necessidades específicas da região.

A continuidade deste trabalho em estudos futuros é altamente relevante para a construção de um corpo de conhecimento mais abrangente e aprofundado sobre o desempenho escolar dos alunos do Ensino Médio no Amapá. Novas pesquisas podem expandir as análises realizadas, explorando diferentes variáveis sociodemográficas, contextos escolares e métodos estatísticos para uma compreensão mais completa dos determinantes do desempenho acadêmico.

Outros pesquisadores podem se concentrar em investigar os impactos específicos de políticas públicas e programas educacionais no desempenho dos alunos, bem como explorar abordagens qualitativas para entender as percepções dos estudantes, professores e gestores escolares. A continuidade deste estudo por meio de colaborações interdisciplinares e investigações multidimensionais pode enriquecer significativamente o entendimento e promover avanços concretos no campo da educação no Amapá.

Referências

DIAZ, M. D. M. (2012). (Des)igualdades de oportunidades no ensino médio brasileiro: **Escolas públicas e privadas**. *Economia*, 13(3a), 553-568. Disponível em: https://www.anpec.org.br/revista/vol13/vol13n3ap553_568.pdf.

GUSMÃO, F. A. F.; AMORIM, S. S. **Desigualdade educacional no ensino médio brasileiro**. *Educação*, v. 47, n. 1, p. e108, 2022. Disponível em: https://www.researchgate.net/profile/Fabio-Alexandre-Ferreira-Gusmao/publication/365703210_Desigualdade_educacional_no_ensino_medio_brasileiro/links/637f2a281766b34c544f182a/Desigualdade-educacional-no-ensino-medio-brasileiro.pdf.

INEP. Microdados do INEP, 2024. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>

INEP. Resultados do SAEB, 2019. Disponível em: <https://undime.org.br/noticia/16-09-2020-09-24-inep-divulga-resultados-do-saeb-2019>.

LAPEI UFG. **Uso de dados na pesquisa: qual a diferença entre dados primários e secundários?**. 2024. Disponível em; <https://lapei.face.ufg.br/p/43695-17-uso-de-dados-na-pesquisa-qual-a-diferenca-entre-dados-primarios-e-secundarios>

Pestana, M. H; Gageiro, J. N. (2009). *Análise categórica, árvores de decisão e análise de conteúdo em ciências sociais e da saúde com o SPSS*. Lisboa; Porto: Lidel, D.L. 2009. - XVIII, 551, [7] p. : il. ; 24 cm. - (Estatística). - Bibliografia, p. 543-546. - ISBN 978-972-757- 585-5.

Silva, E. L. da.; Menezes, E. M. **Metodologia de pesquisa e elaboração de teses e dissertações**, 2005. 4. ed. rev. atual. Universidade Federal de Santa Catarina – UFSC. Florianópolis 2005.

Soares, J. F; Xavier, F. P. **Pressupostos educacionais e estatísticos do IDEB**, 2013. Educ. Soc., Campinas, v. 34, n. 124, p. 903-923.