Fairness em modelos preditivos de evasão no ensino superior: análise de viés e estratégias de mitigação

Nicole Petrica Araujo[1], Maria Clara Luchiari[1], João Silva[1], Victor Lima[1], Levy Boccato[1], Romis Attux[1]

[1]Faculdade de Engenharia Elétrica e Computação – Universidade Estadual de Campinas (FEEC)
Caixa Postal 6101 – 13083-852 – Campinas – SP – Brazil

n247300@dac.unicamp.br, m205674@dac.unicamp.br, j237409@dac.unicamp.br, victorcl@unicamp.br, lboccato@unicamp.br, attux@unicamp.br

Abstract. *This paper evaluates algorithmic fairness in a predictive model for student dropout in higher education, based on academic and socioeconomic data from UNICAMP students. Using a LightGBM model inspired by Farias (2025), we analyzed observed differences in dropout across sensitive groups, fairness metrics, and bias mitigation techniques. Three sensitive attributes were considered: sex, prior attendance at a public high school, and race/color. The results show that the evaluation of machine learning models, especially in sensitive applications, should consider not only predictive performance, but also the fairness of their predictions across sensitive groups.*

Resumo. *Este trabalho avalia a equidade algorítmica (fairness) em um modelo preditivo de evasão escolar no ensino superior, baseado em dados acadêmicos e socioeconômicos de estudantes da UNICAMP. A partir de um modelo LightGBM inspirado em Farias (2025), foram analisadas diferenças observadas na evasão entre grupos sensíveis, métricas de equidade e técnicas de mitigação de viés. Foram considerados três atributos sensíveis: sexo, ensino médio em escola pública e cor/raça. Os resultados evidenciam que a avaliação de modelos de aprendizado de máquina, principalmente em aplicações sensíveis, deve considerar não apenas o desempenho preditivo, mas também a equidade de suas predições entre grupos sensíveis.*

1. Introdução

A evasão escolar no ensino superior representa um grande desafio enfrentado por instituições universitárias, afetando diretamente a trajetória acadêmica dos estudantes, a eficiência da alocação de recursos e o planejamento de políticas de permanência. Em universidades públicas, esse problema assume uma dimensão ainda mais sensível, visto que a interrupção da trajetória formativa pode estar associada não apenas a dificuldades acadêmicas, mas também a fatores socioeconômicos, desigualdades educacionais anteriores e diferenças de acesso a recursos institucionais [Silva Filho et al. 2007].

Nos últimos anos, técnicas de aprendizado de máquina têm sido aplicadas ao contexto educacional ara identificar estudantes em risco de evasão e apoiar ações preventivas [Silva et al. 2020]. No entanto, quando esses modelos são treinados com dados históricos, eles também podem reproduzir desigualdades já presentes tanto na sociedade quanto nas próprias instituições, apresentando comportamentos distintos entre grupos específicos [Carvalho et al. 2024].

Essa preocupação é discutida na área de *fairness* em aprendizado de máquina. Esse campo de pesquisa busca identificar e reduzir vieses nos resultados de modelos preditivos entre grupos populacionais, especialmente quando tais modelos são aplicados em contextos de alto impacto social [Barocas et al. 2023].

Este trabalho toma como referência o estudo de [Farias et al. 2025], voltado à predição de evasão no ensino superior. A partir desse modelo, investiga-se se o bom desempenho global também corresponde a um comportamento equitativo entre grupos sensíveis. Para isso, são analisadas diferenças observadas nos dados, métricas de *fairness* e estratégias de mitigação de viés, considerando sexo, tipo de escola no ensino médio e cor/raça.

2. Materiais e Métodos

2.1. Base de dados e definição do problema preditivo

A base utilizada contempla registros anonimizados de 4.068 estudantes dos cursos de Engenharia Elétrica e Engenharia de Computação da UNICAMP, com anos de ingresso entre 2005 e 2024. A partir desses dados, foi construída uma base no formato estudante-semestre, na qual cada linha representa a situação de um estudante no semestre K, utilizando apenas informações disponíveis até esse momento, incluindo desempenho acadêmico, progressão curricular e variáveis socioeconômicas.

A variável alvo foi definida considerando uma janela futura de quatro semestres. Assim, para um estudante observado no semestre K, o rótulo indica se houve evasão até $K + 4$. Essa formulação permite antecipar a evasão a partir das informações disponíveis em K, em vez de classificar todos os semestres anteriores à saída como casos de evasão.

Após a construção da variável alvo, foram obtidos 40.735 registros estudante-semestre, dos quais 5.809 foram rotulados como evasão, correspondendo a 14,26% das amostras. Essa distribuição evidencia o desbalanceamento entre as classes, justificando o uso de métricas como *recall*, acurácia balanceada e AUC, em vez da acurácia simples [Owusu-Adjei et al. 2023].

A divisão dos dados em treinamento e teste foi realizada por estudante, evitando que registros de um mesmo aluno aparecessem simultaneamente nos dois conjuntos. O conjunto de treino reuniu 2.847 estudantes e 28.654 amostras, das quais 14,29% pertencem à classe evasão, enquanto o conjunto de teste reuniu 1.221 estudantes e 12.081 amostras, com 14,18% da classe evasão.

2.2. Definição dos atributos sensíveis

Atributos sensíveis correspondem a variáveis associadas a grupos sociais sujeitos a desigualdades históricas e/ou socioeconômicas, exigindo avaliação cuidadosa de possíveis vieses em modelos preditivos [Mehrabi et al. 2021]. Considerando a disponibilidade da base de dados e a relevância para a permanência estudantil, foram analisados três atributos sensíveis: sexo registrado na base, cor/raça e origem em escola pública no ensino médio.

Cabe ressaltar que a variável sexo está restrita às categorias feminino e masculino. Dessa forma, embora o estudo dialogue com desigualdades de gênero em áreas de ciência, tecnologia e engenharia, a análise não contempla identidade de gênero ou outras dimensões socioculturais associadas ao gênero [Heidari et al. 2016].

2.3. Modelagem preditiva

A modelagem foi realizada com o algoritmo LightGBM, selecionado pelo desempenho no estudo de referência e pela adequação a dados tabulares não lineares [Ke et al. 2017]. Foram utilizadas 21 variáveis relacionadas ao ingresso, trajetória acadêmica e perfil socioeconômico, incluindo desempenho, progressão curricular, reprovações, trancamentos, renda, atividade remunerada e atributos sensíveis.

As variáveis numéricas foram padronizadas por transformação z-score, com parâmetros estimados no conjunto de treinamento e aplicados ao conjunto de teste. As variáveis categóricas foram codificadas por *label encoding*, mantendo valores ausentes como dados faltantes.

A variável alvo foi codificada de forma binária, com valor 0 para não evasão e 1 para evasão no horizonte futuro de quatro semestres. O limiar de decisão foi definido exclusivamente no conjunto de treino, por validação cruzada interna com 10 dobras e agrupamento por estudante, evitando vazamento de informação para o conjunto de teste [Kakkar et al. 2024]. Foram avaliados limiares entre 0,01 e 0,80, priorizando maior acurácia balanceada e, em caso de empate, maior *recall*. O procedimento resultou no limiar final de 0,116.

Com esse limiar, o modelo final foi treinado no conjunto de treino completo e avaliado no conjunto de teste, apresentando AUC de 0,9057, *recall* de 0,8383 e acurácia balanceada de 0,8195, valores próximos aos reportados no modelo base: 0,9266, 0,8201 e 0,8443, respectivamente [Farias et al. 2025].

2.4. Avaliação de *fairness* e técnicas de mitigação de viés

A avaliação de *fairness* foi conduzida em três etapas, considerando o desempenho global do modelo e seu comportamento entre grupos sensíveis. Primeiro, foram analisadas as taxas reais de evasão entre os grupos, antes da aplicação do modelo, para identificar diferenças já presentes nos dados. Em seguida, foram calculadas métricas de equidade para o modelo base. Por fim, foram aplicadas técnicas de mitigação de viés, avaliando se as diferenças entre grupos poderiam ser reduzidas sem perdas expressivas de desempenho.

A seleção das métricas de *fairness* considerou o contexto de classificação binária da evasão escolar e foi fundamentada em [Caton and Haas 2024]. Foram priorizadas métricas de paridade e métricas baseadas na matriz de confusão, por permitirem avaliar, respectivamente, diferenças na atribuição do rótulo de risco e diferenças nos padrões de erro e acerto entre grupos sensíveis.

Dessa forma, foi selecionada uma métrica de paridade, a *Demographic Parity Difference* (DPD), e três métricas baseadas na matriz de confusão: *Equalized Odds Difference* (EOD), *True Positive Rate Difference* (TPRD) e *False Positive Rate Difference* (FPRD). A DPD mede diferenças na taxa de seleção entre grupos, isto é, na proporção de estudantes classificados como em risco de evasão. A EOD avalia diferenças conjuntas nas taxas de verdadeiros positivos e falsos positivos; a TPRD mede diferenças na capacidade de identificar corretamente estudantes que evadiram; e a FPRD mede diferenças na proporção de estudantes incorretamente classificados como em risco. Em todas essas métricas, valores menores indicam menor desigualdade entre grupos.

Além disso, foram avaliadas técnicas de mitigação de viés em três momentos da

modelagem. No pré-processamento, utilizaram-se o *Correlation Remover*, que reduz a associação entre atributos sensíveis e variáveis preditoras, e o *SMOTENC*, voltado ao tratamento do desbalanceamento da classe positiva em bases com variáveis numéricas e categóricas [Feldman et al. 2015, Chawla et al. 2002]. No processamento, a mitigação foi incorporada ao treinamento por meio do *Exponentiated Gradient*, sob restrições de *Demographic Parity* e *Equalized Odds* [Agarwal et al. 2018]. No pós-processamento, foram ajustados os limiares de decisão com o *Threshold Optimizer*, considerando restrições de *Demographic Parity*, *Equalized Odds* e *True Positive Rate Parity* [Hardt et al. 2016]. As abordagens foram comparadas segundo o equilíbrio entre desempenho preditivo e métricas de *fairness*.

3. Resultados e Discussões

3.1. Análise exploratória

A análise considera ingressantes da UNICAMP entre 2005 e 2024 dos cursos de Engenharia Elétrica e Engenharia de Computação. Nesse período, a universidade passou por mudanças nas políticas de acesso e permanência, como o PAAIS [Universidade Estadual de Campinas 2004] e a reformulação dos sistemas de ingresso [Universidade Estadual de Campinas 2017]. Esse contexto, somado à persistência de desigualdades de gênero em áreas de ciência, tecnologia e engenharia [Barata 2019], orienta a interpretação dos resultados como associações restritas ao período, aos cursos e à base analisada, e não como efeitos causais dos atributos sobre a evasão.

Diante desse recorte, a análise exploratória avaliou diferenças nas taxas de evasão entre grupos sensíveis antes da aplicação do modelo. Os resultados são sintetizados na Figura 1.

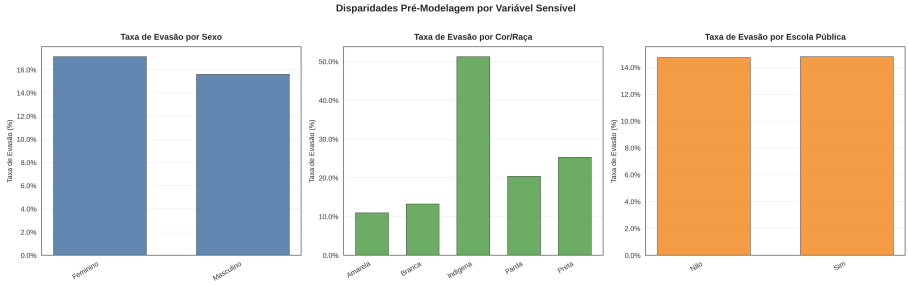

Figura 1. Taxa de evasão dentro de cada categoria dos atributos sensíveis.

A associação entre atributos sensíveis e evasão foi avaliada pelo teste qui-quadrado, complementado pelo V de Cramér para análise da magnitude dos efeitos [Agresti 2013, Cramér 1946]. Como complemento ao p-valor, calculou-se o V de Cramér, medida que permite distinguir significância estatística de relevância prática [Cramér 1946]. Para cor/raça, os resíduos padronizados ajustados auxiliaram a interpretação das categorias com maior contribuição [Sharpe 2015].

Para sexo, a informação estava disponível em 76,2% dos registros analisados. Embora o teste qui-quadrado tenha indicado associação estatisticamente significativa com a evasão ($\chi^2 = 5{,}83$; $gl = 1$; $p = 0{,}0157$), a disparidade máxima entre os grupos foi de apenas 1,5 ponto percentual, com tamanho de efeito muito baixo (V de Cramér = 0,0137). Assim, a magnitude da associação sugere baixa relevância substantiva,

pois o p-valor não expressa, isoladamente, a importância prática do efeito observado [Wasserstein and Lazar 2016].

Para escola pública, a informação estava disponível em 85,6% dos registros analisados. As taxas de evasão foram praticamente equivalentes entre estudantes com ensino médio em escola pública e os demais, e o teste qui-quadrado não indicou associação estatisticamente significativa com a evasão ($\chi^2 = 0{,}02$; $gl = 1$; $p = 0{,}8883$), com tamanho de efeito praticamente nulo (V de Cramér = 0,0008). Esse resultado é restrito ao recorte analisado e não exclui relações entre trajetória escolar e evasão associadas a outros fatores, especialmente socioeconômicos e acadêmicos, como renda, desempenho e condições de permanência.

Para cor/raça, disponível em 85,6% dos registros, observaram-se as maiores diferenças, com maiores taxas de evasão entre estudantes indígenas, pretos e pardos. O teste qui-quadrado indicou associação estatisticamente significativa com a evasão ($\chi^2 = 471{,}08$; $gl = 4$; $p < 0{,}001$), com disparidade máxima de 40,3 pontos percentuais e tamanho de efeito pequeno (V de Cramér = 0,1181). Os resíduos padronizados ajustados indicaram maior contribuição das categorias indígena, parda e preta para a associação global. Esse resultado sugere que desigualdades prévias associadas ao acesso, à trajetória acadêmica e às condições de permanência podem estar refletidas nas taxas observadas, devendo-se ponderar, em particular, o menor número de registros da categoria indígena.

Em síntese, os resultados indicam que as diferenças observadas não foram homogêneas entre os atributos sensíveis: sexo apresentou associação estatisticamente significativa, porém de baixa magnitude prática; escola pública não apresentou associação relevante com a evasão; e cor/raça concentrou as maiores variações. Essas diferenças descrevem padrões presentes na base de dados antes da aplicação do modelo e, isoladamente, não caracterizam viés produzido pelo algoritmo, podendo refletir desigualdades históricas, institucionais e socioeconômicas, além de fatores não registrados nos dados. Dessa forma, a avaliação de *fairness* torna-se essencial para verificar se o modelo preditivo mantém, reduz ou amplia essas desigualdades nas suas predições, especialmente em relação a grupos historicamente sub-representados no ensino superior.

3.2. *Fairness* do modelo base

Após a análise exploratória da evasão, avaliou-se o comportamento do modelo base entre os grupos sensíveis no conjunto de teste. Embora o modelo tenha apresentado bom desempenho geral, com AUC de 0,9057, *recall* de 0,8383 e acurácia balanceada de 0,8195, as métricas de *fairness* indicaram que esse desempenho não foi homogêneo entre os grupos.

No contexto da evasão escolar, falsos negativos representam estudantes que evadiram, mas não foram classificados como em risco, podendo deixar de ser priorizados em ações de apoio e permanência. Falsos positivos correspondem a estudantes sinalizados sem evasão posterior, o que pode gerar alocação desnecessária de recursos e reduzir a confiança nos alertas do modelo [Kemper et al. 2019, Del Bonifro et al. 2020]. Assim, a análise de *fairness* deve considerar a distribuição dos alertas, a identificação dos casos reais de evasão e os erros de classificação.

Para sexo, a DPD de 0,0084 indica taxas de seleção muito semelhantes entre estudantes do sexo feminino e masculino. A principal diferença ocorreu na sensibilidade:

entre os estudantes que evadiram, o modelo identificou corretamente 86,00% dos homens e 78,61% das mulheres. Ainda assim, essa diferença deve ser ponderada pela baixa discrepância na taxa de seleção, pela FPRD reduzida (0,0182) e pela distribuição desigual de registros entre os grupos.

Para escola pública, as assimetrias foram intermediárias. O modelo classificou estudantes com ensino médio em escola pública como em risco em proporção um pouco maior que os demais e apresentou maior sensibilidade para esse grupo, identificando 90,76% dos casos reais de evasão, contra 80,49% entre os demais. Do ponto de vista preventivo, esse resultado reduz falsos negativos, mas a FPRD de 0,0377 indica maior taxa de falsos positivos.

As maiores diferenças foram observadas para cor/raça. A DPD de 0,5877 indica que os grupos raciais foram classificados como em risco de evasão em proporções bastante distintas, afetando a distribuição dos alertas de acompanhamento. A EOD de 0,4807 mostra que a desigualdade também ocorreu nos padrões de erro e acerto do modelo. Nos grupos indígena e preto, a maior sensibilidade veio acompanhada de maior taxa de falsos positivos, indicando maior identificação de casos reais de evasão, mas também mais falsos alertas. Por outro lado, o grupo amarelo apresentou menor taxa de seleção e menor *recall*, sugerindo maior risco de não identificação de estudantes que efetivamente evadiriam. Assim, para cor/raça, a preocupação central envolve tanto a distribuição desigual dos alertas quanto a distribuição dos erros entre grupos.

De modo geral, o modelo base apresentou baixa disparidade para sexo, disparidade moderada para escola pública e disparidades expressivas para cor/raça. Esses resultados mostram que métricas agregadas de desempenho não são suficientes em contextos sensíveis, pois diferentes grupos podem ser afetados de modo desigual pela distribuição dos alertas e pelos padrões de erro do modelo. A Tabela 1 resume os principais indicadores de *fairness* obtidos.

Tabela 1. Métricas de *fairness* do modelo base por atributo sensível.

Variável sensível	DPD	EOD	TPRD	FPRD
Sexo	0,0084	0,0739	0,0739	0,0182
Cor/raça	0,5877	0,4807	0,2473	0,4807
Escola pública	0,0475	0,1027	0,1027	0,0377

3.3. Comparação das técnicas de mitigação de viés

As técnicas de pré-processamento, processamento e pós-processamento foram comparadas pelo equilíbrio entre desempenho preditivo e redução das diferenças entre grupos. A Figura 2 apresenta a relação entre *recall* e as métricas EOD e DPD. No painel superior, a região superior esquerda representa melhor equilíbrio entre maior *recall* e menor diferença nos padrões de erro e acerto entre grupos; no painel inferior, representa melhor equilíbrio entre maior *recall* e menor diferença na taxa de seleção entre grupos.

Os resultados indicaram que não houve uma técnica superior para todos os atributos sensíveis. Para sexo, o modelo base permaneceu como a configuração mais equilibrada, pois já apresentava baixa DPD (0,0084), EOD reduzida (0,0739) e elevado *recall* (0,8498), sem ganhos relevantes com as técnicas de mitigação.

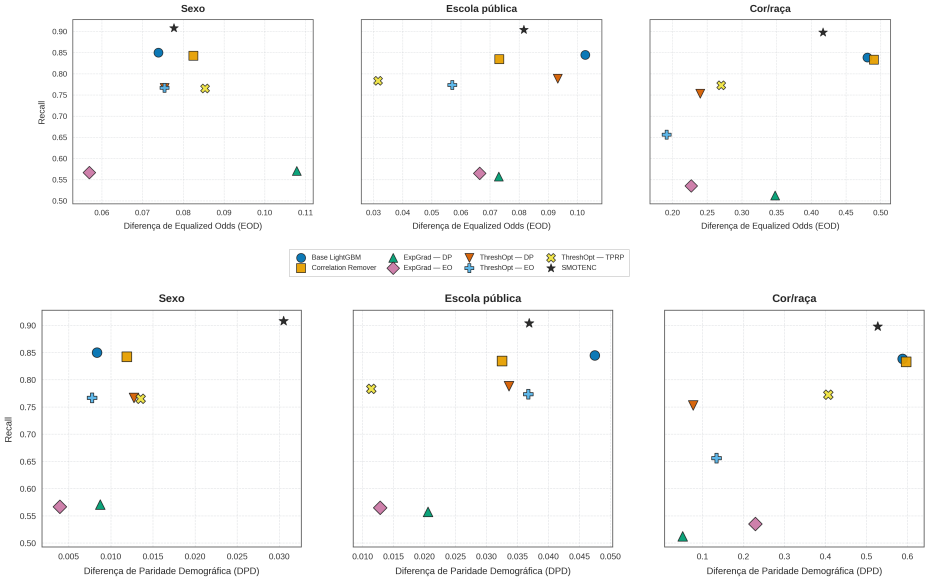

Figura 2. Relação entre *recall* e métricas de equidade por atributo sensível.

Para escola pública, os melhores resultados foram obtidos com métodos de pós-processamento. O *Threshold Optimizer* com restrição de *True Positive Rate Parity* reduziu a EOD e a TPRD de 0,1027 para 0,0235, além de reduzir a DPD de 0,0475 para 0,0120, mantendo a acurácia balanceada em patamar semelhante ao modelo base. Assim, houve maior aproximação da sensibilidade entre os grupos, embora com redução no *recall* geral.

Para cor/raça, as assimetrias entre grupos foram mais expressivas e difíceis de mitigar. O *Threshold Optimizer* com restrição de *Demographic Parity* reduziu fortemente a DPD, de 0,5877 para 0,0286, indicando maior equilíbrio na taxa de seleção entre grupos raciais. Contudo, a EOD permaneceu elevada (0,3521), mostrando que a redução da disparidade de seleção não equalizou os padrões de erro e acerto. A restrição de *Equalized Odds* reduziu mais a EOD (0,2006), mas com queda substancial de *recall* e acurácia balanceada. Assim, a mitigação por *Demographic Parity* foi considerada a alternativa mais equilibrada para cor/raça, embora limitada quanto à redução dos erros entre grupos.

A Tabela 2 sintetiza o desempenho e as métricas de *fairness* do modelo base e das configurações selecionadas após a mitigação de viés.

Tabela 2. Comparação entre o modelo base e as configurações selecionadas após mitigação por atributo sensível.

Atributo sensível	Configuração	*Recall*	Bal. Acc	AUC	DPD	EOD	TPRD
Sexo	Base	0,8498	0,8197	0,9109	0,0084	0,0739	0,0739
Cor/raça	Base	0,8384	0,8264	0,9152	0,5877	0,4807	0,2473
Cor/raça	*Threshold Optimizer* – DP	0,7472	0,8165	0,9161	0,0286	0,3521	0,3521
Escola pública	Base	0,8447	0,8293	0,9169	0,0475	0,1027	0,1027
Escola pública	*Threshold Optimizer* – TPRP	0,7849	0,8297	0,9175	0,0120	0,0235	0,0235

As métricas preditivas do modelo base foram calculadas separadamente para cada grupo dos atributos sensíveis, utilizando apenas as amostras do conjunto de teste disponíveis em cada subconjunto. Por esse motivo, os valores podem apresentar diferenças em relação às métricas globais, obtidas considerando todo o conjunto de teste.

A Figura 3 sintetiza visualmente o equilíbrio entre desempenho preditivo e equidade nas configurações selecionadas, representando DPD, EOD e TPRD como métricas invertidas, de modo que valores maiores indiquem melhor resultado.

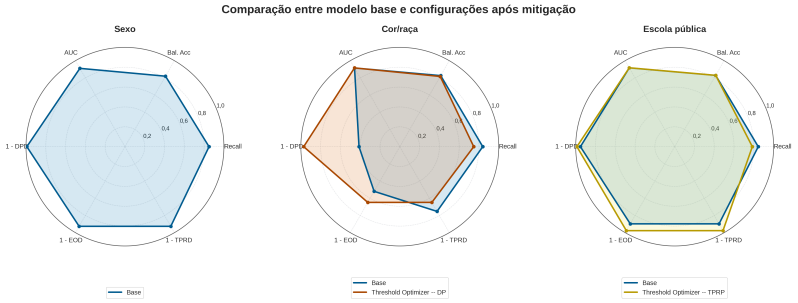

Figura 3. Comparação entre o modelo base e as configurações selecionadas após mitigação de viés.

Os resultados corroboram a literatura de *fairness*, segundo a qual diferentes critérios de equidade nem sempre podem ser satisfeitos simultaneamente [Kleinberg et al. 2017, Chouldechova 2017]. Isso ficou evidente em cor/raça: a redução da DPD pelo *Threshold Optimizer* com *Demographic Parity* não eliminou as diferenças nos padrões de erro medidos pela EOD. Assim, a escolha da técnica de mitigação deve ser tratada como uma decisão contextual e ética, orientada pelo objetivo do modelo e pelo tipo de erro que se pretende reduzir.

4. Conclusões

Este trabalho avaliou a equidade algorítmica em um modelo preditivo de evasão escolar no ensino superior, baseado em *LightGBM* e aplicado a dados acadêmicos e socioeconômicos de estudantes da UNICAMP. Embora o modelo reproduzido tenha apresentado desempenho próximo ao estudo de referência, a análise por grupos mostrou que métricas globais de desempenho não asseguram comportamento equitativo entre os atributos sensíveis avaliados.

As maiores diferenças foram observadas para cor/raça, tanto nas taxas de evasão anteriores à modelagem quanto nas predições do modelo base, especialmente na distribuição dos alertas e nos padrões de erro. Para sexo, as diferenças foram menos expressivas, enquanto escola pública apresentou comportamento intermediário. A comparação das técnicas de mitigação também mostrou que não há solução única para reduzir todas as dimensões de *fairness*, pois diferentes estratégias atuam sobre taxa de seleção, sensibilidade ou padrões de erro.

Assim, modelos preditivos podem apoiar ações de permanência estudantil quando utilizados como ferramentas auxiliares, com monitoramento contínuo de desempenho e equidade. Como limitação, destaca-se o recorte específico de cursos e a utilização de uma base histórica no formato estudante-semestre, restringindo a generalização para outros contextos. Trabalhos futuros podem ampliar a análise para outros cursos, períodos e unidades acadêmicas, incorporar variáveis relacionadas à permanência e avaliar estratégias de mitigação integradas a protocolos de acompanhamento e auditoria contínua.

5. Agradecimentos

Agradecemos ao DEAPE/UNICAMP e ao CNPq pelo apoio institucional e financeiro à iniciação científica que deu origem a este trabalho.

Referências

Agarwal, A., Beygelzimer, A., Dudík, M., Langford, J., and Wallach, H. (2018). A reductions approach to fair classification. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 60–69. PMLR.

Agresti, A. (2013). *Categorical Data Analysis*. Wiley, 3 edition.

Barata, G. (2019). Ainda há muito espaço para mulheres e meninas na ciência e tecnologia. Portal Unicamp. Acesso em: 18 maio 2026.

Barocas, S., Hardt, M., and Narayanan, A. (2023). *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press.

Carvalho, C. S., Mattos, J. C. B., and Aguiar, M. S. (2024). Interpretabilidade e justiça algorítmica: Avançando na transparência de modelos preditivos de evasão escolar. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 1658–1673. Sociedade Brasileira de Computação.

Caton, S. and Haas, C. (2024). Fairness in machine learning: A survey. *ACM Computing Surveys*, 56(7):1–38. Published online: 23 August 2023.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2):153–163.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton.

Del Bonifro, F., Gabbrielli, M., Lisanti, G., and Zingaro, S. P. (2020). Student dropout prediction. In *Artificial Intelligence in Education*, pages 129–140. Springer International Publishing.

Farias, J. A., Lima, V. C., Souza, M., and Lopes, R. d. R. (2025). Modelagem preditiva de evasão escolar no ensino superior. In *Anais do 53º Congresso Brasileiro de Educação em Engenharia*. Associação Brasileira de Educação em Engenharia.

Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., and Venkatasubramanian, S. (2015). Certifying and removing disparate impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 259–268. ACM.

Hardt, M., Price, E., and Srebro, N. (2016). Equality of opportunity in supervised learning. In *Advances in Neural Information Processing Systems*, volume 29, pages 3315–3323.

Heidari, S., Babor, T. F., De Castro, P., Tort, S., and Curno, M. (2016). Sex and gender equity in research: Rationale for the sager guidelines and recommended use. *Research Integrity and Peer Review*, 1(2).

Kakkar, S., Jaya Lakshmi, A., Durai C, A. D., Ahmad, W., Ayoobkhan, M. U. A., Veeramanikandan, P., Reddy, P. N., Dwivedi, V. K., and Rajaram, A. (2024). Enhancing energy efficiency and classification modeling through a combined approach of LightGBM and stratified KFold cross-validation. *Electric Power Components and Systems*, pages 1–19.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30.

Kemper, L., Vorhoff, G., and Wigger, B. U. (2019). Early identification of college dropouts using machine-learning: Conceptual considerations and an empirical example. IZA Research Report 89, IZA Institute of Labor Economics.

Kleinberg, J., Mullainathan, S., and Raghavan, M. (2017). Inherent trade-offs in the fair determination of risk scores. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference*, volume 67 of *LIPIcs*, pages 43:1–43:23. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., and Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6):1–35.

Owusu-Adjei, M., Ben Hayfron-Acquah, J., Frimpong, T., and Abdul-Salaam, G. (2023). Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems. *PLOS Digital Health*, 2(11):e0000290.

Sharpe, D. (2015). Your chi-square test is statistically significant: Now what? *Practical Assessment, Research, and Evaluation*, 20(8):1–10.

Silva, F. C. d., Cabral, T. L. d. O., and Pacheco, A. S. V. (2020). Evasão ou permanência? modelos preditivos para a gestão do ensino superior. *Education Policy Analysis Archives*, 28(149):1–32.

Silva Filho, R. L. L. e., Motejunas, P. R., Hipólito, O., and Lobo, M. B. d. C. M. (2007). A evasão no ensino superior brasileiro. *Cadernos de Pesquisa*, 37(132):641–659.

Universidade Estadual de Campinas (2004). Deliberação CONSU-A-012/2004, de 25/05/2004: Estabelece o Programa de Ação Afirmativa para Inclusão Social na UNICAMP. Conselho Universitário.

Universidade Estadual de Campinas (2017). Deliberação CONSU-A-032/2017, de 21/11/2017: Dispõe sobre os sistemas de ingresso aos cursos de Graduação da Unicamp. Conselho Universitário.

Wasserstein, R. L. and Lazar, N. A. (2016). The asa statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2):129–133.