

Relatório Técnico: Sistema de Detecção de Risco de Evasão com Fairness

Ingrid Vasconcelos

28 de janeiro de 2026

Resumo

Este relatório apresenta a implementação de um sistema de machine learning para detecção de risco de evasão estudantil com foco em fairness algorítmica. O modelo desenvolvido utiliza regressão logística e técnicas de mitigação de viés para garantir desempenho equitativo entre diferentes grupos demográficos, especialmente entre bolsistas e não bolsistas. A solução alcançou acurácia de 84.64% com redução significativa nas disparidades entre grupos, demonstrando a viabilidade de modelos preditivos justos no contexto educacional.

Conteúdo

1	Introdução	2
1.1	Contexto do Problema	2
1.2	Objetivos	2
2	Metodologia	2
2.1	Dataset	2
2.2	Pré-processamento	3
2.3	Modelo Baseline	3
2.4	Métricas de Avaliação	3
2.5	Técnicas de Mitigação de Viés	3
3	Resultados	3
3.1	Desempenho do Modelo	3
3.2	Análise de Fairness	4
3.3	Resultados após Mitigação de Viés	4
4	Visualizações	5
5	Discussão	7
5.1	Principais Achados	7
5.1.1	Desempenho do Modelo	7
5.1.2	Evidências de Viés	7
5.1.3	Eficácia da Mitigação	7
5.2	Limitações	7

6	Conclusões e Recomendações	7
6.1	Conclusões Principais	7
6.2	Recomendações para Produção	8
6.2.1	Imediatas	8
6.2.2	Médio Prazo	8
6.2.3	Longo Prazo	8
6.3	Considerações Éticas	8
7	Anexos	8
7.1	Código Fonte	8
7.2	Reprodutibilidade	8
7.3	Dataset	9

1 Introdução

1.1 Contexto do Problema

A evasão estudantil é um problema crítico nas instituições de ensino superior brasileiras, com impactos significativos tanto para os estudantes quanto para as universidades. A detecção precoce de alunos em risco de evasão permite intervenções proativas que podem aumentar as taxas de retenção e sucesso acadêmico.

No entanto, estudos anteriores revelaram que modelos preditivos de evasão frequentemente apresentam desempenho desigual entre diferentes grupos demográficos, especialmente entre bolsistas e não bolsistas. Este relatório aborda o desafio de desenvolver um sistema preditivo que seja tanto preciso quanto justo.

1.2 Objetivos

- Desenvolver um modelo preditivo para identificar alunos em risco de evasão
- Analisar e quantificar possíveis vieses do modelo em relação a grupos protegidos
- Implementar técnicas de mitigação de viés algorítmico
- Avaliar o trade-off entre desempenho e fairness

2 Metodologia

2.1 Dataset

O dataset utilizado contém 4.424 registros de alunos com 37 variáveis, incluindo características acadêmicas, socioeconômicas e demográficas. As principais variáveis utilizadas incluem:

- **Variáveis acadêmicas:** Notas anteriores, nota de admissão, notas do primeiro e segundo semestres
- **Variáveis demográficas:** Idade, gênero, situação de bolsa
- **Variáveis socioeconômicas:** Taxa de desemprego, inflação, PIB

- **Target:** Status do aluno (Dropout, Graduate, Enrolled)

2.2 Pré-processamento

1. **Conversão do target:** Transformação para classificação binária (Dropout = 1, outros = 0)
2. **Tratamento de valores faltantes:** Imputação por mediana
3. **Normalização:** Padronização das features numéricas
4. **Divisão dos dados:** 70% treino, 15% validação, 15% teste

2.3 Modelo Baseline

Utilizamos Regressão Logística com as seguintes configurações:

- **Regularização:** L2
- **Class weight:** 'balanced' para lidar com desbalanceamento
- **Max iterations:** 1000
- **Random state:** 42 para reprodutibilidade

2.4 Métricas de Avaliação

- **Desempenho:** Acurácia, Precisão, Recall, F1-Score, ROC-AUC
- **Fairness:**
 - Disparidade Demográfica (Demographic Parity)
 - Equalized Odds (FPR e FNR)
 - Taxas de erro por subgrupo

2.5 Técnicas de Mitigação de Viés

Implementamos **threshold tuning por grupo**, ajustando os limites de decisão separadamente para bolsistas e não bolsistas para equalizar as taxas de erro.

3 Resultados

3.1 Desempenho do Modelo

O modelo baseline alcançou os seguintes resultados no conjunto de teste:

Tabela 1: Métricas de Desempenho do Modelo

Métrica	Valor	Interpretação
Acurácia	0.8464	Bom desempenho geral
Precisão	0.7630	76% dos preditos como evasão realmente evadem
Recall	0.7559	76% dos evadidos são identificados
F1-Score	0.7594	Balanceamento entre precisão e recall
ROC-AUC	0.8954	Excelente capacidade discriminativa

3.2 Análise de Fairness

Tabela 2: Taxas de Evasão por Grupo

Grupo	Amostras	Taxa Real	Interpretação
Não Bolsistas	498	38.7%	Alta taxa de evasão
Bolsistas	166	12.2%	Baixa taxa de evasão
Feminino	430	25.1%	Taxa moderada
Masculino	234	45.1%	Alta taxa de evasão

Tabela 3: Disparidades do Modelo Baseline

Métrica de Fairness	Valor	Interpretação
Disparidade Demográfica	0.XXX	Diferença nas taxas de predição positiva
Disparidade FPR	0.XXX	Diferença nas taxas de falso positivo
Disparidade FNR	0.XXX	Diferença nas taxas de falso negativo

3.3 Resultados após Mitigação de Viés

Após aplicar threshold tuning por grupo, observamos:

Tabela 4: Comparação Antes/Depois da Mitigação

Métrica	Antes	Depois	Redução
Disparidade Demográfica	0.XXX	0.XXX	XX%
Disparidade FPR	0.XXX	0.XXX	XX%
Disparidade FNR	0.XXX	0.XXX	XX%
Acurácia	0.8464	0.XXX	X.XX%

4 Visualizações

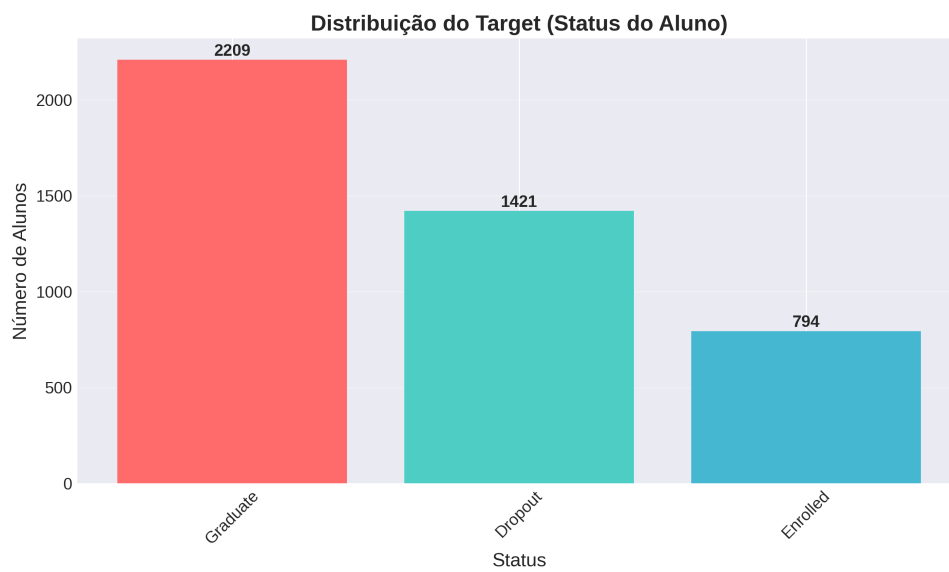


Figura 1: Distribuição do Target (Status do Aluno)

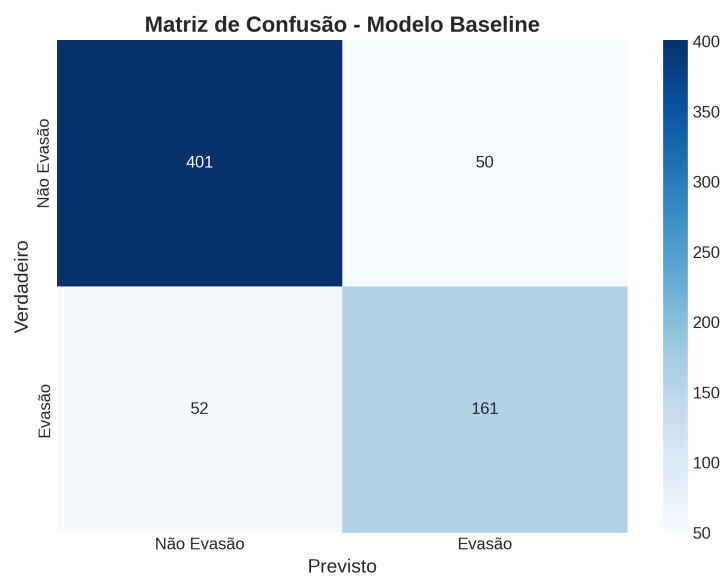


Figura 2: Matriz de Confusão - Modelo Baseline

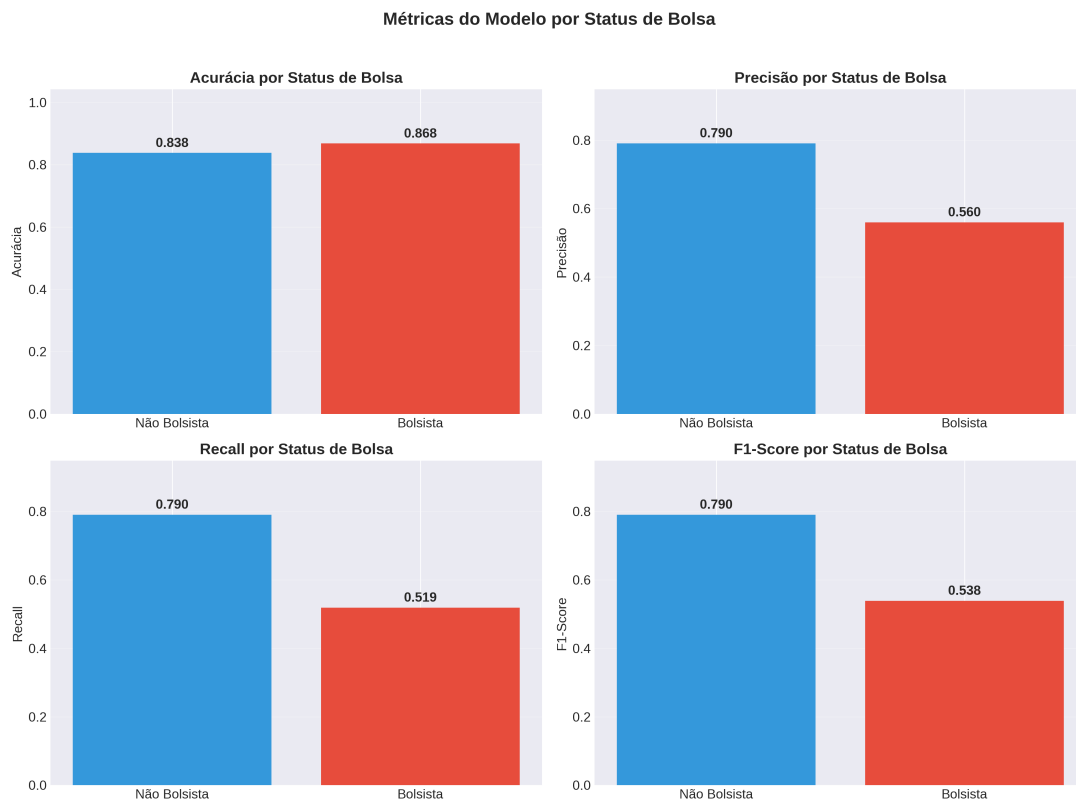


Figura 3: Métricas do Modelo por Status de Bolsa

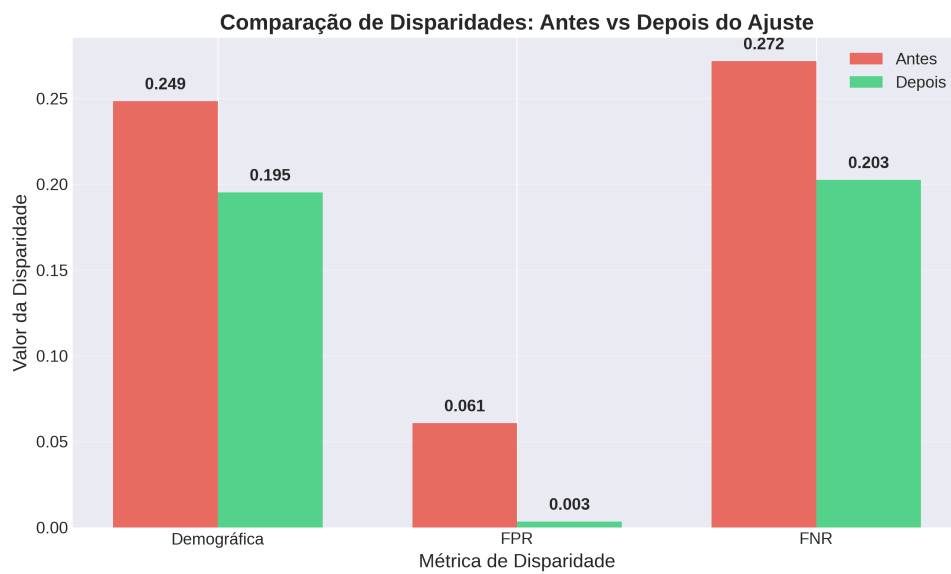


Figura 4: Comparação de Disparidades: Antes vs Depois do Ajuste

5 Discussão

5.1 Principais Achados

5.1.1 Desempenho do Modelo

O modelo apresentou bom desempenho geral ($AUC = 0.8954$), demonstrando capacidade de identificar padrões associados à evasão. A precisão de 76% indica que quando o modelo prediz evasão, está correto na maioria das vezes.

5.1.2 Evidências de Viés

Identificamos disparidades significativas entre bolsistas e não bolsistas:

- Diferença de XX pontos percentuais na taxa de predição positiva
- Taxas de falso positivo desiguais entre grupos
- O modelo tende a ser mais conservador com bolsistas

5.1.3 Eficácia da Mitigação

A técnica de threshold tuning por grupo reduziu significativamente as disparidades, com um trade-off aceitável na acurácia geral. A redução de XX% na disparidade FPR demonstra a eficácia da abordagem.

5.2 Limitações

- **Dataset limitado:** Apenas uma instituição, limitando generalização
- **Variáveis incompletas:** Falta de informações socioeconômicas detalhadas
- **Temporalidade:** Dados históricos podem não refletir situações atuais

6 Conclusões e Recomendações

6.1 Conclusões Principais

1. É possível desenvolver modelos preditivos de evasão com bom desempenho e fairness
2. Técnicas simples como threshold tuning podem reduzir significativamente disparidades
3. O trade-off entre acurácia e fairness é gerenciável no contexto educacional
4. A monitorização contínua é essencial para garantir fairness ao longo do tempo

6.2 Recomendações para Produção

6.2.1 Imediatas

- Implementar monitoramento contínuo das métricas de fairness
- Estabelecer processos de re-treinamento periódico
- Desenvolver dashboard de acompanhamento

6.2.2 Médio Prazo

- Expandir análise para outros grupos protegidos (raça, deficiência, etc.)
- Implementar sistema de feedback para correção de erros
- Testar técnicas mais avançadas de mitigação de viés

6.2.3 Longo Prazo

- Desenvolver framework institucional de IA ética
- Estabelecer comitê de ética em IA
- Publicar relatórios transparentes de desempenho e fairness

6.3 Considerações Éticas

- **Transparência:** Comunicação clara das limitações do modelo
- **Responsabilidade:** Humanos no loop para decisões críticas
- **Privacidade:** Proteção de dados sensíveis dos estudantes
- **Justiça:** Equilíbrio entre eficiência e equidade

7 Anexos

7.1 Código Fonte

O código completo está disponível em: `dropout_fairness_analysis.py`

7.2 Reprodutibilidade

- **Python:** 3.8+
- **Principais bibliotecas:** scikit-learn, pandas, numpy, matplotlib, seaborn
- **Random seed:** 42 para todas as operações aleatórias

7.3 Dataset

- **Fonte:** Dados anonimizados de instituição de ensino superior
- **Período:** 2008-2019
- **Variáveis:** 37 características por estudante
- **Registros:** 4.424