



Protótipo de IA para Redução da Evasão Escolar

Desafio ED_TECH Senac-GO

Aldrei Marucci Veiga e Diego Manuel - 1ª Turma Técnico em IA

Agosto 2025

CONHEÇA O DESAFIO

Contexto e Foco da Solução

O desafio proposto pelo gerente Gabriel Schimchak visa a criação de um protótipo com Inteligência Artificial para solucionar um problema educacional premente.

Nossa escolha foi a evasão escolar, um problema transversal que afeta todos os níveis de ensino no Brasil.



Por que evasão escolar?

É um fenômeno complexo com impactos sociais e econômicos profundos, que exige soluções inovadoras e assertivas.

A REALIDADE POR TRÁS DOS NÚMEROS

A História de Lucas



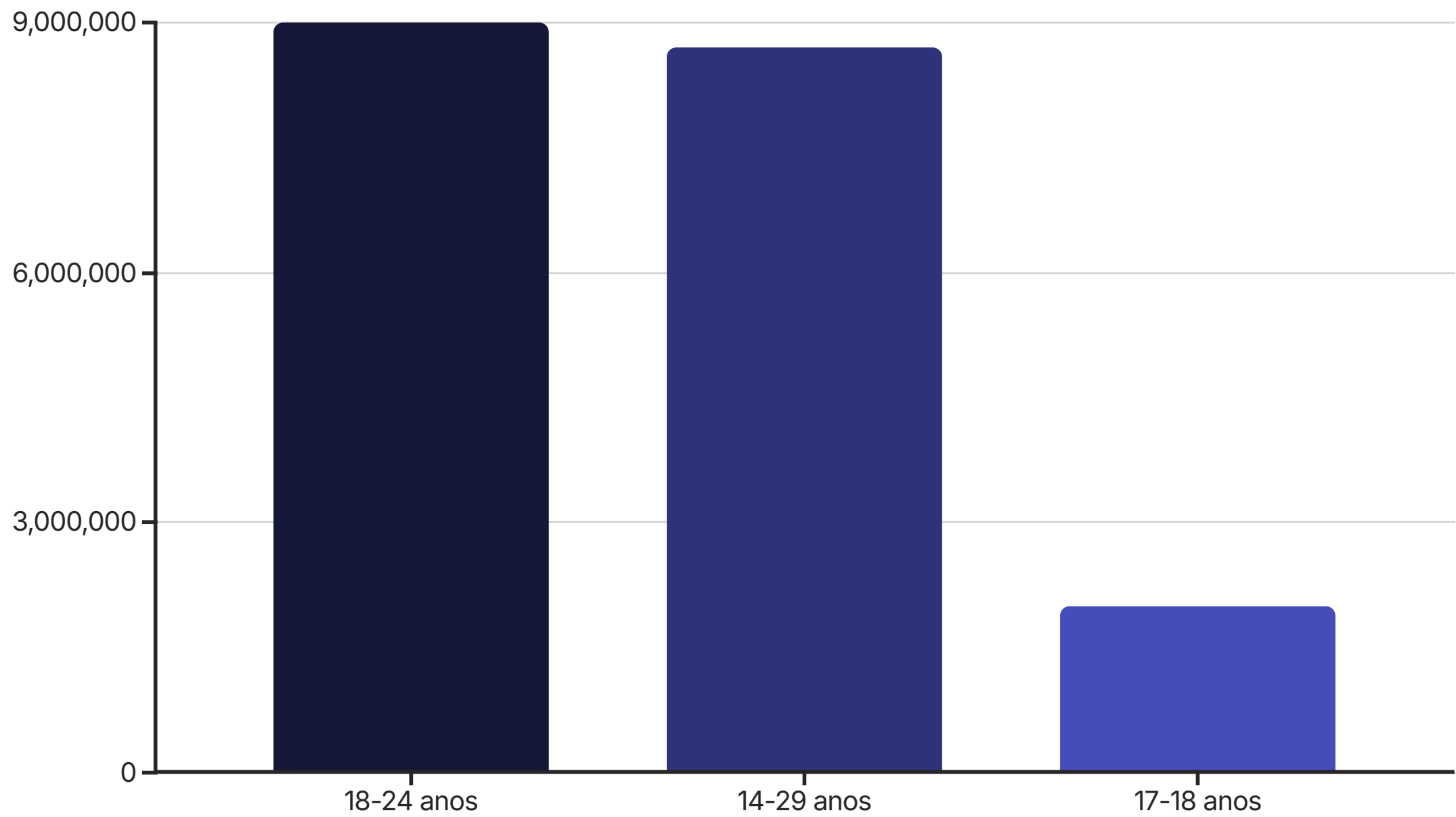
Lucas, 16 anos, morador do interior de Goiás, representa milhões de jovens. Seus pais não concluíram o ensino médio, e ele enfrenta desafios diários:

- Dificuldades financeiras
- Falta de apoio familiar e escolar
- Ambiente escolar desmotivador
- Escassez de recursos didáticos
- Transporte escolar irregular

"Quero aprender, mas as dificuldades são muitas..."

Lucas é um símbolo do potencial sufocado pelas circunstâncias, cuja história buscamos reverter com a IA.

Dados Estatísticos Alarmantes



Taxa Urbana

Evasão em centros urbanos.



Contextos Específicos

Evasão em áreas de vulnerabilidade.

Dados do IBGE/PNADC 2023 revelam que 9 milhões de jovens (18-24 anos) abandonaram a escola antes de concluir o ensino médio. Em 2024, são 8,7 milhões de jovens (14-29 anos) sem essa formação. A evasão se intensifica aos 16 anos, atingindo 20% entre 17-18 anos.

COMPREENDENDO OS MOTIVOS

Principais Causas da Evasão Escolar



Necessidade de Trabalhar

Cerca de **30%** dos casos são impulsionados pela pressão econômica.



Falta de Interesse

Aproximadamente **25%** dos estudantes desistem por desmotivação.



Outros Fatores

O restante (~**45%**) abrange questões socioeconômicas diversas, problemas familiares e ambiente escolar.

A evasão é um fenômeno multifacetado, com raízes em fatores econômicos, sociais e pedagógicos. Identificar e endereçar essas causas é crucial para desenvolver intervenções eficazes.

O Marco Legal da Educação em Goiás

Legislação Chave

- **Projeto de Lei Estadual GO nº 14915/24:** "Política Estadual Aluno Presente", focando na garantia da permanência do estudante.
- **Plano Estadual de Educação GO (2015-2025):** Lei nº 18.969/2015, estabelecendo diretrizes e metas para a educação.

A necessidade de identificar e antecipar momentos de evasão é um pilar dessas leis.

Iniciativas Complementares

- **Selo "Escola que Cuida":** Reconhece e incentiva boas práticas escolares na proteção de crianças e adolescentes.
- **Programa Jovem do Futuro:** Foca em apoiar estudantes para garantir a conclusão de sua trajetória educacional.



A INOVAÇÃO EM AÇÃO

Nossa Solução: O Protótipo de IA

Sistema Preditivo

Identificação precoce do risco de evasão.



Análise Multivariada

Considera variáveis educacionais e socioeconômicas.

Algoritmos de ML

Uso de Machine Learning para previsão.



Interface Intuitiva

Plataforma para gestores educacionais.

Nosso protótipo utiliza inteligência artificial para antecipar os sinais de que um estudante pode abandonar os estudos, permitindo intervenções proativas e personalizadas.

COMO CONSTRUÍMOS

Fluxo do Protótipo de IA

01

Análise Exploratória de Dados

Limpeza e preparação de dados educacionais, identificando padrões e correlações.

03

Modelos de Classificação e Regressão

Aplicação de algoritmos avançados para prever riscos e categorizar perfis de estudantes.

A metodologia de desenvolvimento seguiu um rigoroso processo de ciência de dados, garantindo a robustez e eficácia do protótipo.

02

Implementação em Google Colab

Desenvolvimento e treinamento dos modelos em um ambiente de nuvem acessível e escalável.

04

Validação Estatística

Teste e ajuste dos modelos para garantir precisão e confiabilidade nos resultados preditivos.

EVIDENCIANDO O SUCESSO

Resultados e Visualizações

Planilha gerada aleatoriamente

	A	B	C	D	E	F	G	H	I	J	K	L
1	Aluno_ID	Data_Nascimento	Ano_Escolar	Periodo	Materia	Frequencia_Periodo	Nota_Prova	Nota_Participacao	Auxilio_Financeiro	Necessidade_Especial	Valor_Mensalidade	Pagou_Mensalidade
2	ALU001	2008-06-10	2023	2023-01	Português	40	86,4	6,9	Não	Nenhuma	786,42	Sim
3	ALU001	2008-06-10	2023	2023-01	Matemática	40	75,6	7,3	Não	Nenhuma	786,42	Sim
4	ALU001	2008-06-10	2023	2023-01	Ciências Biológicas	88,6	65,1	7,8	Não	Nenhuma	786,42	Sim
5	ALU001	2008-06-10	2023	2023-01	Educação Física	93,6	94,4	8,6	Não	Nenhuma	786,42	Sim
6	ALU001	2008-06-10	2023	2023-01	Inglês	92,9	72,3	6,7	Não	Nenhuma	786,42	Sim
7	ALU001	2008-06-10	2023	2023-01	Estudo Religioso	86,9	82,9	7,3	Não	Nenhuma	786,42	Sim
8	ALU002	2009-04-07	2023	2023-01	Português	74,3	63,6	8,3	Não	Nenhuma	810,43	Não
9	ALU002	2009-04-07	2023	2023-01	Matemática	72,7	70,1	9,7	Não	Nenhuma	810,43	Sim
10	ALU002	2009-04-07	2023	2023-01	Ciências Biológicas	77,9	65,8	8,3	Não	Nenhuma	810,43	Sim
11	ALU002	2009-04-07	2023	2023-01	Educação Física	79,1	45	10	Não	Nenhuma	810,43	Sim
12	ALU002	2009-04-07	2023	2023-01	Inglês	77,5	76,1	9,7	Não	Nenhuma	810,43	Não
13	ALU002	2009-04-07	2023	2023-01	Estudo Religioso	75,5	79,6	10	Não	Nenhuma	810,43	Sim
14	ALU003	2010-08-06	2023	2023-01	Português	89,4	85,2	6,8	Não	Nenhuma	807,15	Sim
15	ALU003	2010-08-06	2023	2023-01	Matemática	92,2	98,7	6,3	Não	Nenhuma	807,15	Sim
16	ALU003	2010-08-06	2023	2023-01	Ciências Biológicas	99,1	65	7,2	Não	Nenhuma	807,15	Sim
17	ALU003	2010-08-06	2023	2023-01	Educação Física	89,6	79,2	6,4	Não	Nenhuma	807,15	Sim
18	ALU003	2010-08-06	2023	2023-01	Inglês	95,1	82,3	6,2	Não	Nenhuma	807,15	Sim
19	ALU003	2010-08-06	2023	2023-01	Estudo Religioso	88,9	79,2	5,9	Não	Nenhuma	807,15	Sim
20	ALU004	2009-06-22	2023	2023-01	Português	73,9	64,8	6,8	Não	Nenhuma	807,9	Sim
21	ALU004	2009-06-22	2023	2023-01	Matemática	75,5	77,2	8,2	Não	Nenhuma	807,9	Sim
22	ALU004	2009-06-22	2023	2023-01	Ciências Biológicas	78,5	66,9	6,2	Não	Nenhuma	807,9	Sim

Resultados e Visualizações

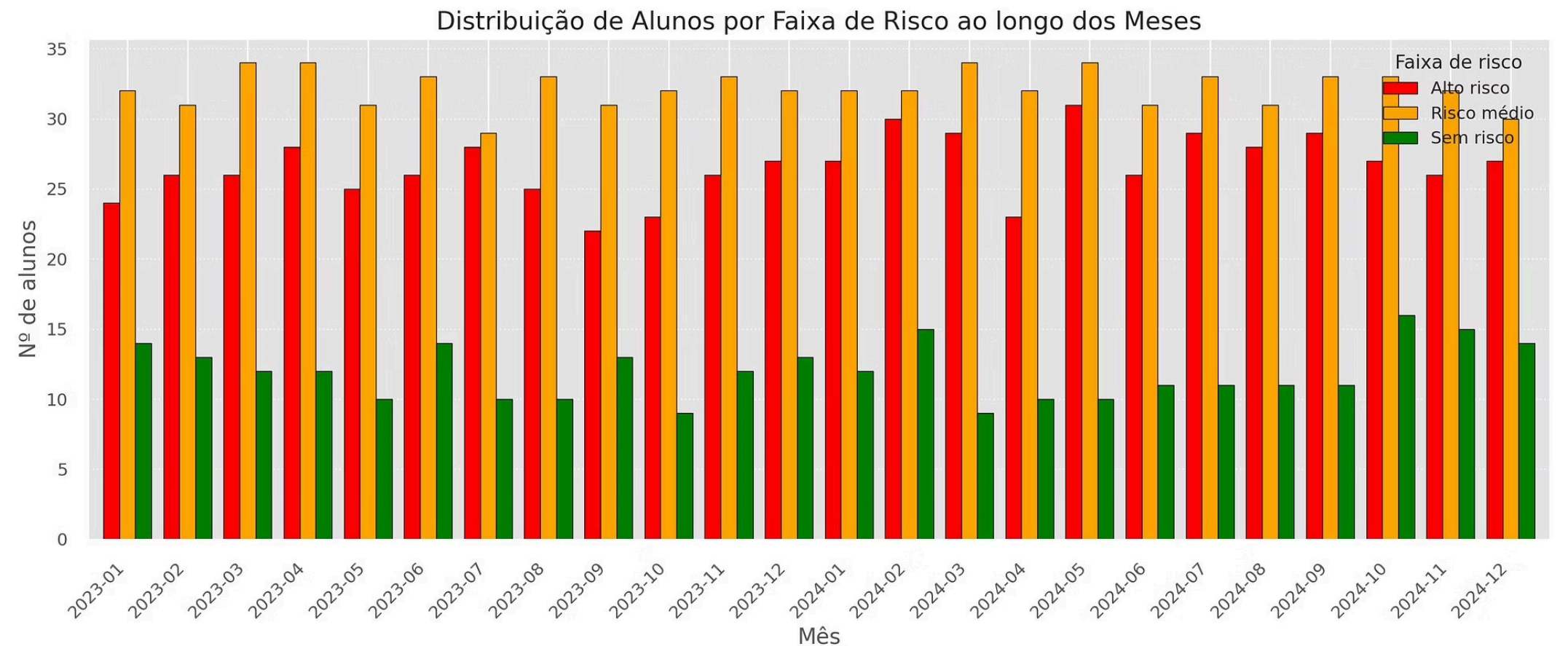
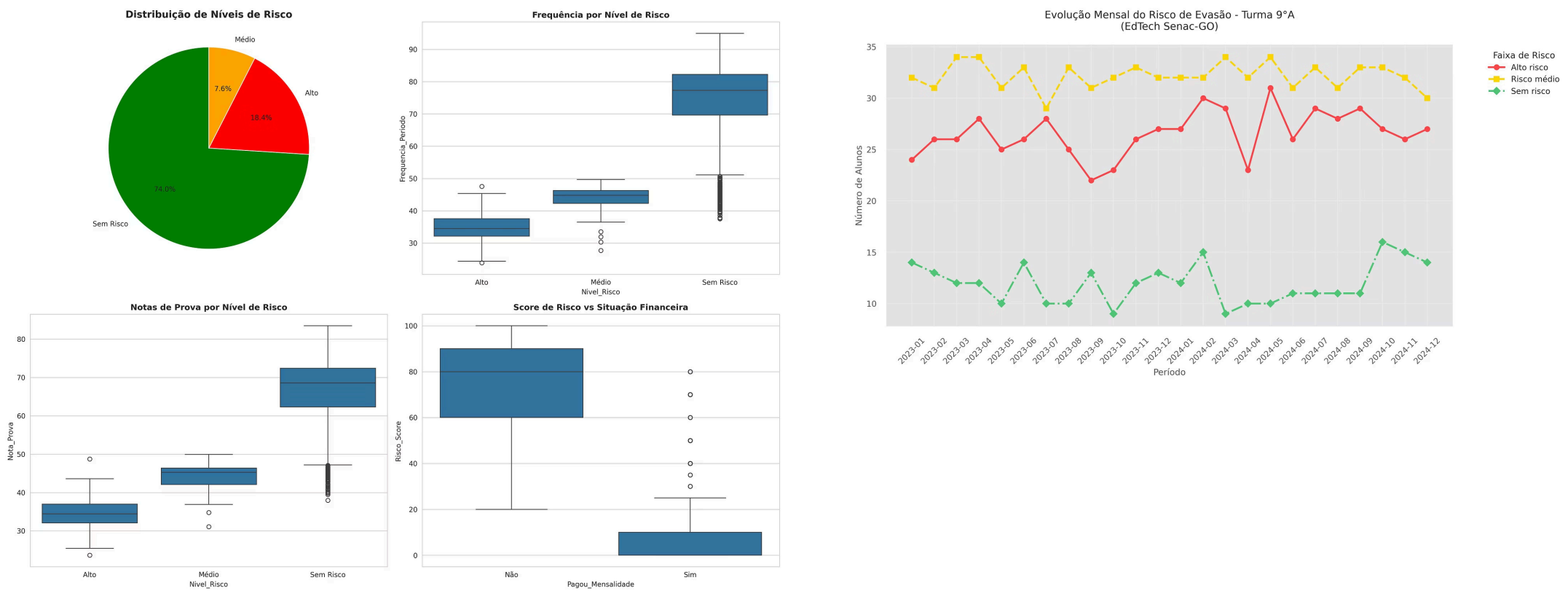
Métricas de Performance do Modelo

Resultados iniciais demonstram alta precisão na identificação de estudantes em risco.

Análises Preditivas

Visualizações claras que permitem aos gestores identificar e agir sobre os casos críticos.

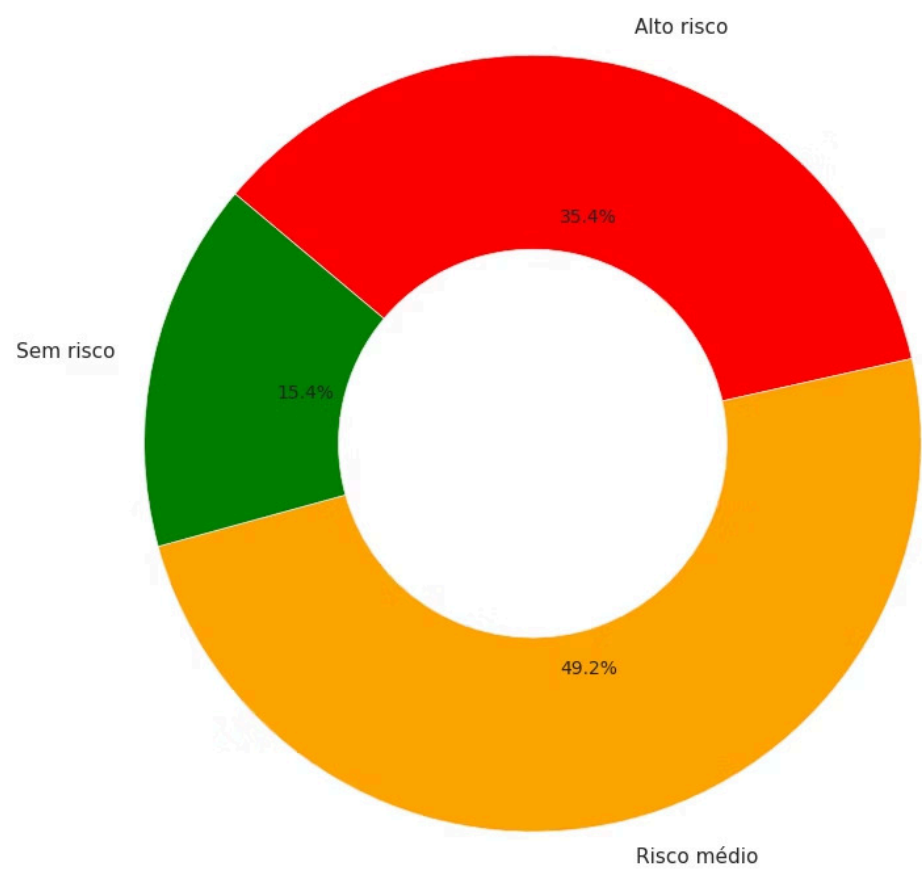
A análise preditiva do nosso protótipo oferece uma visão clara e acionável para a tomada de decisão no ambiente educacional.



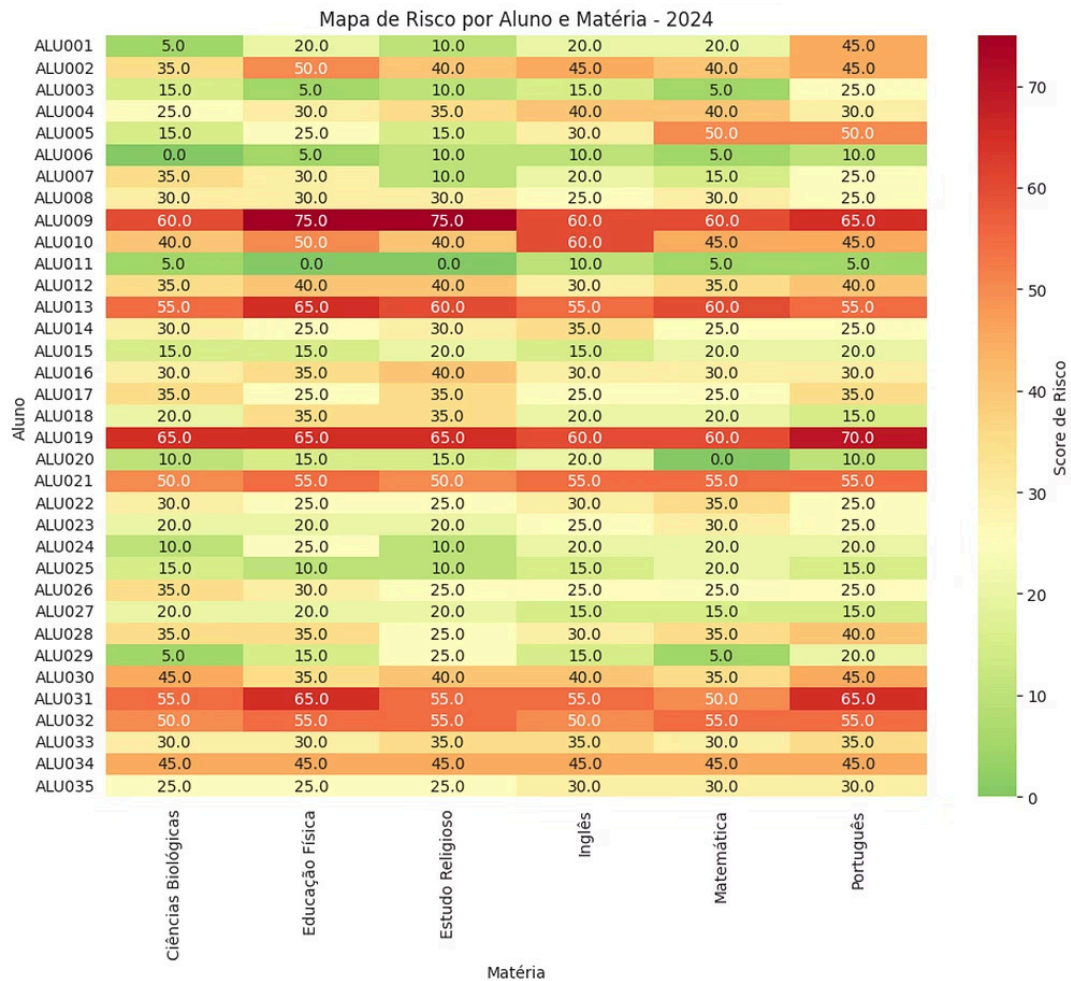
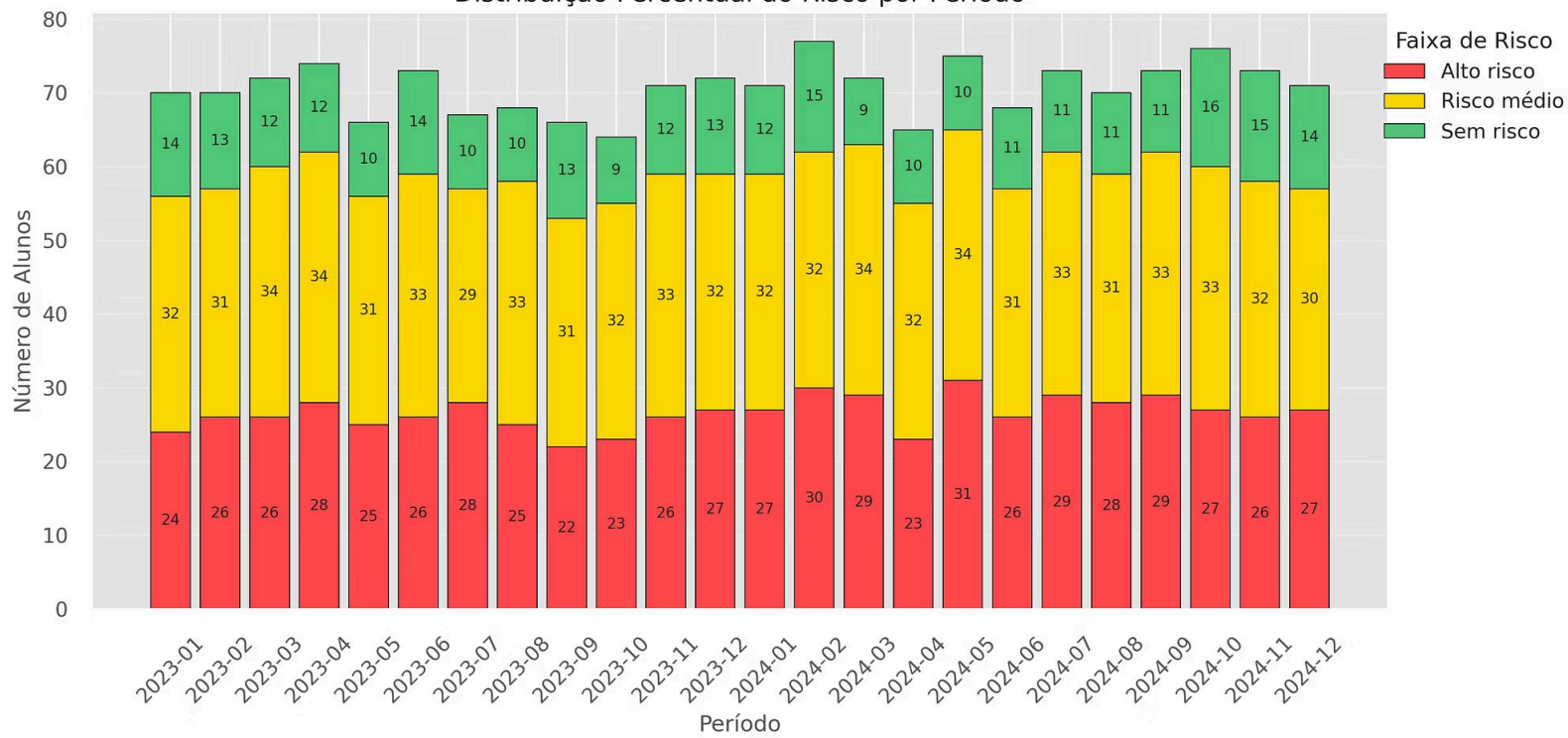
EVIDENCIANDO O SUCESSO

Resultados e Visualizações

Proporção das Faixas de Risco em 2024-04

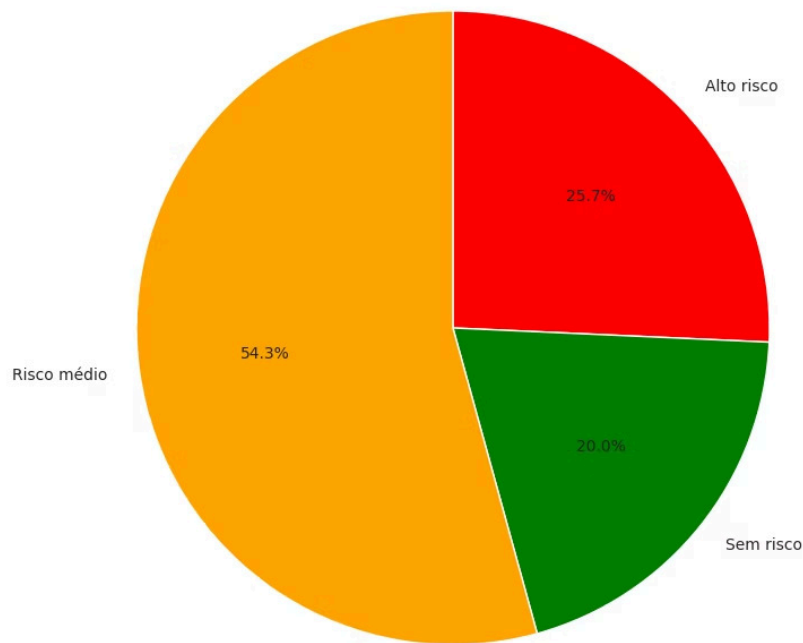


Distribuição Percentual do Risco por Período

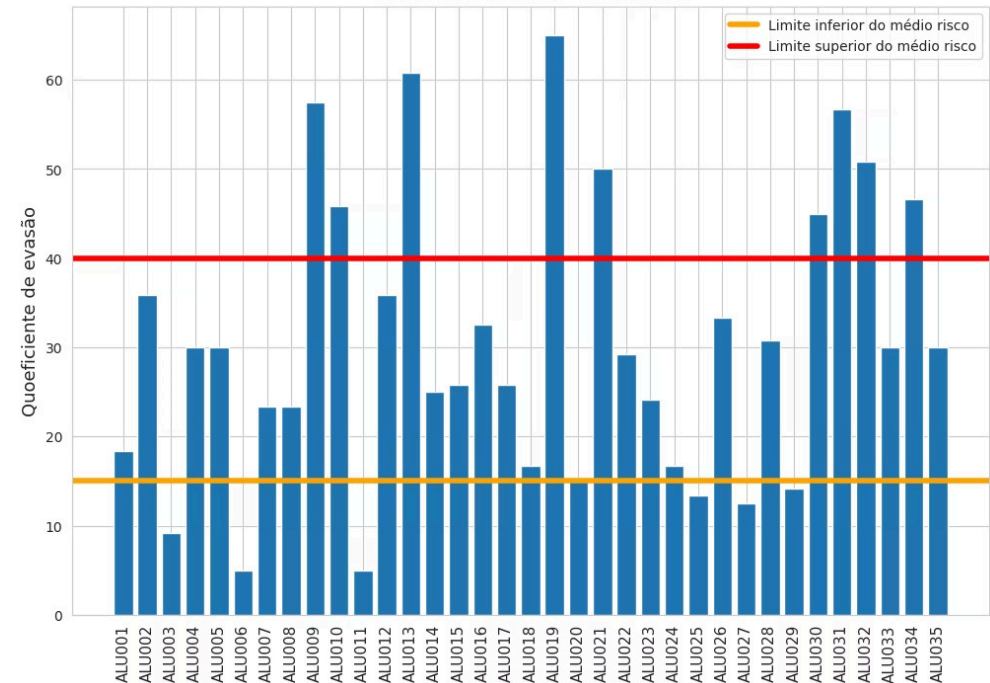


Resultados e Visualizações

Proporção de faixas de risco da turma - 2023



Risco dos alunos - Ano 2023



EVIDENCIANDO O SUCESSO

Resultados e Visualizações

Link no colab para o arquivo: ED.TECH_Senac_ANALISE PREDITIVA_v.4

https://colab.research.google.com/drive/1-F1ihglaa1mDfuPhesfBc_h4kWzUBENp?usp=sharing

🔍 INICIANDO ANÁLISE DETALHADA COMPLETA...

📊 ANÁLISE COMPLETA GERADA: 420 linhas (Aluno x Matéria x Ano)

🎯 GERANDO INTERVENÇÕES PERSONALIZADAS (por linha agregada)...

✅ Intervenções aplicadas ao dataframe agregado.

- Média intervenções por registro: 9.4

🚨 CASOS CRÍTICOS IDENTIFICADOS: 108

🔴 ALU009 | Estudo Religioso | 2024

Score: 75.0 | Freq: 77.7 | Nota: 56.27

Fatores: Frequência Baixa, Notas Baixas, Participação Crítica

Intervenções sugeridas (exemplo):

- academias: ['📖 MONITORIA ACADÊMICA: Acompanhamento semanal com monitor', '💡 TÉCNICAS DE ESTUDO: Workshop sobre métodos
- frequencia: ['📊 MONITORAMENTO SEMANAL: Relatório semanal de frequência', '🎁 INCENTIVOS À PRESENÇA: Sistema de recompensa
- participacao: ['🎮 GAMIFICAÇÃO: Implementar elementos lúdicos nas aulas', '🧑‍🤝‍🧑 DESENVOLVIMENTO SOCIAL: Workshops de express
- necessidades_especiais: ['🔗 ADAPTAÇÕES CURRICULARES: Personalização do conteúdo', '🔧 RECURSOS ASSISTIVOS: Ferramentas te
- familiares: ['👨‍👩‍👧‍👦 REUNIÕES FAMILIARES: Encontros quinzenais com família', '🗣️ COMUNICAÇÃO CONSTANT
- motivacionais: ['🏆 PLANO DE METAS: Estabelecer objetivos alcançáveis', '💬 SESSÕES DE COACHING: Orientação vocacional e m

🔴 ALU009 | Educação Física | 2024

Score: 75.0 | Freq: 75.62 | Nota: 59.2

Fatores: Frequência Baixa, Notas Baixas, Participação Crítica

Intervenções sugeridas (exemplo):

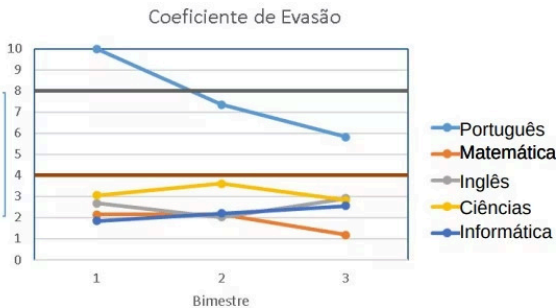
- academias: ['📖 MONITORIA ACADÊMICA: Acompanhamento semanal com monitor', '💡 TÉCNICAS DE ESTUDO: Workshop sobre métodos
- frequencia: ['📊 MONITORAMENTO SEMANAL: Relatório semanal de frequência', '🎁 INCENTIVOS À PRESENÇA: Sistema de recompensa
- participacao: ['🎮 GAMIFICAÇÃO: Implementar elementos lúdicos nas aulas', '🧑‍🤝‍🧑 DESENVOLVIMENTO SOCIAL: Workshops de express
- necessidades_especiais: ['🔗 ADAPTAÇÕES CURRICULARES: Personalização do conteúdo', '🔧 RECURSOS ASSISTIVOS: Ferramentas te
- familiares: ['👨‍👩‍👧‍👦 REUNIÕES FAMILIARES: Encontros quinzenais com família', '🗣️ COMUNICAÇÃO CONSTANT
- motivacionais: ['🏆 PLANO DE METAS: Estabelecer objetivos alcançáveis', '💬 SESSÕES DE COACHING: Orientação vocacional e m

EVIDENCIANDO O SUCESSO

Resultados e Visualizações



Nome: LUCAS GABRIEL
Turma: 5º ano B
Telefone:(62) 9 2001-7015
Email: lucasgabriel@gmail.com



Causas do coeficiente:
Nota baixa, Participação baixa _____

Sugestões para melhorar:
Monitoria acadêmica, projetos colaborativos _____

O FUTURO DA EDUCAÇÃO

Impacto Esperado e Transformação de Vidas



Redução da Evasão

Mais estudantes concluindo suas jornadas educacionais.



Intervenções Personalizadas

Apoio direcionado para cada necessidade.



Gestão Aprimorada

Decisões baseadas em dados e inteligência.



Transformação de Vidas

Oportunidades concretas para jovens como o Lucas.

Nosso protótipo não é apenas uma ferramenta tecnológica, é um catalisador para um futuro onde o potencial de cada estudante pode ser plenamente realizado.

O FUTURO DA EDUCAÇÃO com a Inteligência Artificial

COMO A INTELIGÊNCIA ARTIFICIAL CONTRIBUIRÁ COM O PROJETO ?



Análise Preditiva de Risco

- **Detecção Precoce de Risco** Nossos modelos de IA analisam o histórico acadêmico, frequência, indicadores socioeconômicos e engajamento digital para identificar estudantes com até 85% de precisão que correm risco de evasão;
- **Insights Ocultos** Usando técnicas avançadas de machine learning, nossos modelos treinados revelam correlações surpreendentes, como por exemplo: o impacto do transporte irregular no desempenho em matemática;
- **Análise Integrada** Processamos simultaneamente dados socioeconômicos, desempenho acadêmico e padrões de frequência para obter uma visão holística de cada estudante;
- **Alertas Antecipados** Nossos modelos de machine learning identificam estudantes com mais de 70% de risco de evasão até **6 meses antes** do possível abandono, permitindo intervenções precoces;
- **Perfis de Risco** Algoritmos avançados de classificação, como Random Forest e Regressão Logística, categorizam os alunos de acordo com seus níveis de vulnerabilidade.

O FUTURO DA EDUCAÇÃO com a Inteligência Artificial

COMO A INTELIGÊNCIA ARTIFICIAL CONTRIBUIRÁ COM O PROJETO ?



Personalização de Intervenções

- **Mapeamento de perfis:** Classifica estudantes em categorias como exemplo: "risco econômico" ou "desmotivação pedagógica", direcionando ações específicas;
- **Acompanhamento dinâmico:** Reavalia riscos mensalmente com dados atualizados;
- **Recomendações Adaptativas:** Sistema sugere estratégias específicas para cada perfil;
- **Plano de Ação Individualizado:** Geração automática de roadmaps educacionais baseados em lacunas de aprendizagem detectadas.

O FUTURO DA EDUCAÇÃO com a Inteligência Artificial

COMO A INTELIGÊNCIA ARTIFICIAL CONTRIBUIRÁ COM O PROJETO ?



Integração com Gestão Educacional

- **Dashboard intuitivo:** Visualizações em tempo real (ex.: heatmaps de evasão por região) auxiliam decisões estratégicas;
- **Otimização de recursos:** Prioriza escolas com maior Índice de Vulnerabilidade Educacional (IVE) para receber programas de apoio;
- **Sistema de Notificações:** Chatbots educacionais enviam mensagens motivacionais personalizadas via:
 - WhatsApp;
 - Aplicativo móvel próprio;
 - E-mail automatizado;
- **Gamificação Adaptativa:** Mecânicas de recompensa ajustadas ao perfil comportamental de cada aluno;
- **Feedback Loop Automatizado:** Modelos que se retroalimentam com dados de efetividade das intervenções;
- **Painel de Gestão em Tempo Real:** Dashboard interativo com:
 - Taxa de risco atualizada diariamente;
 - Progresso das ações implementadas;
 - Projeções de impacto nas métricas educacionais.

O FUTURO DA EDUCAÇÃO com a Inteligência Artificial

COMO A INTELIGÊNCIA ARTIFICIAL CONTRIBUIRÁ COM O PROJETO ?



Aprendizado contínuo do sistema

- *Feedback loops* atualizam algoritmos com novos dados educacionais e resultados de intervenções;
- Adaptação a contextos locais: Modelo reconhece particularidades de áreas rurais vs. urbanas.

"A IA não substitui educadores, mas potencializa sua capacidade de agir preventivamente".

Dados reais

Insights do Dataset MEC

Dados com mais de 1 milhão de linhas para identificar fatores que influenciam a evasão escolar e construir previsões. Trabalhamos com 299 mil linhas.

<https://app.powerbi.com/view?r=eyJrljoiZDhkNGNiYzgtMjQ0My00OGVILWJjNzYtZWQwYjI2OThhYWWM1liwidCI6IjllNjgyMzU5LWQxMjgtNGVkYi1iYjU4LTgyYjJhMTUzNDBmZiJ9>

- | | |
|----------------|--------------------------|
| → COR/RAÇA | → EIXO TECNOLÓGICO |
| → FAIXA ETÁRIA | → FONTE DE FINANCIAMENTO |
| → IDADE | → RENDA FAMILIAR |
| → SEXO | → TURNO |

🎓 ANÁLISE DE EVASÃO ESCOLAR - DADOS MEC

=====

📊 ANÁLISE DETALHADA DA ESTRUTURA DOS DADOS

=====

📄 Dimensões do dataset: 299999 linhas x 38 colunas

💻 Memória utilizada: 502.39 MB



Dados reais

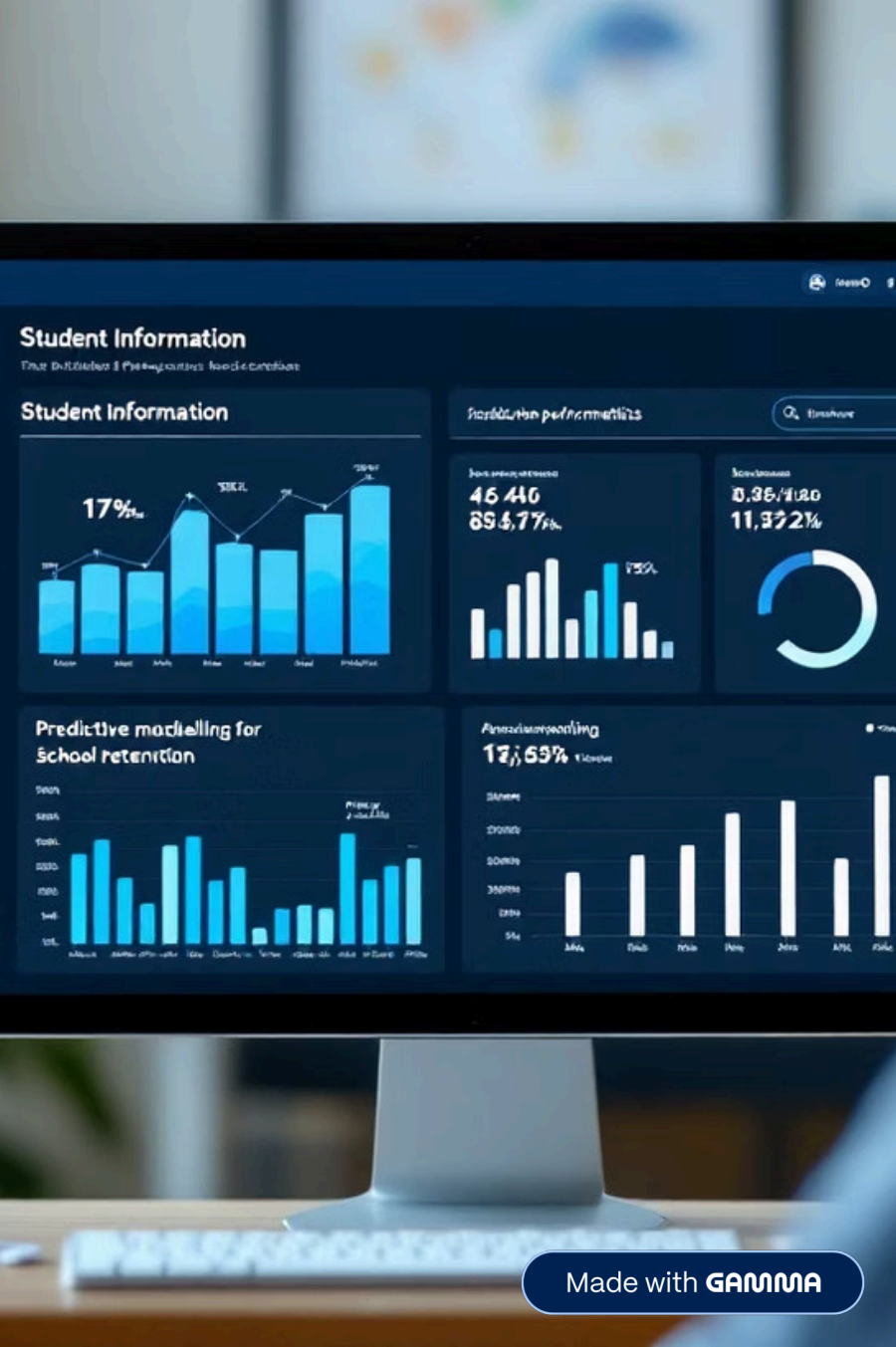
Insights do Dataset MEC

PARTE 1: ASSOCIAÇÃO COM EVASÃO

- Analizando variáveis categóricas com Cramér's V...
- Analizando variáveis numéricas com Correlação...

TABELA DE ASSOCIAÇÃO COM EVASÃO:

1.	Habilitação	Cramér's V	0.156 (p= 0.000)	✓	Fraca
2.	Turno	Cramér's V	0.094 (p= 0.000)	✓	Muito Fraca
3.	Fonte de Financiamento	Cramér's V	0.084 (p= 0.000)	✓	Muito Fraca
4.	Renda Familiar	Cramér's V	0.067 (p= 0.000)	✓	Muito Fraca
5.	Eixo Tecnológico	Cramér's V	0.062 (p= 0.000)	✓	Muito Fraca
6.	Idade	Correlação de Pearson	0.035 (p= 0.000)	✓	Muito Fraca
7.	Idade	Correlação de Pearson	0.035 (p= 0.000)	✓	Muito Fraca
8.	Cor / Raça	Cramér's V	0.023 (p= 0.000)	✓	Muito Fraca
9.	Sexo	Cramér's V	0.008 (p= 0.000)	✓	Muito Fraca



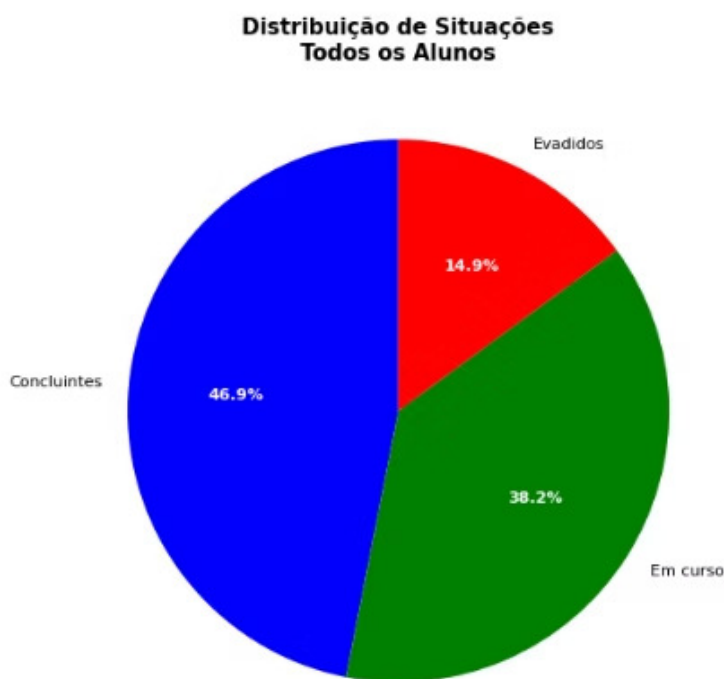
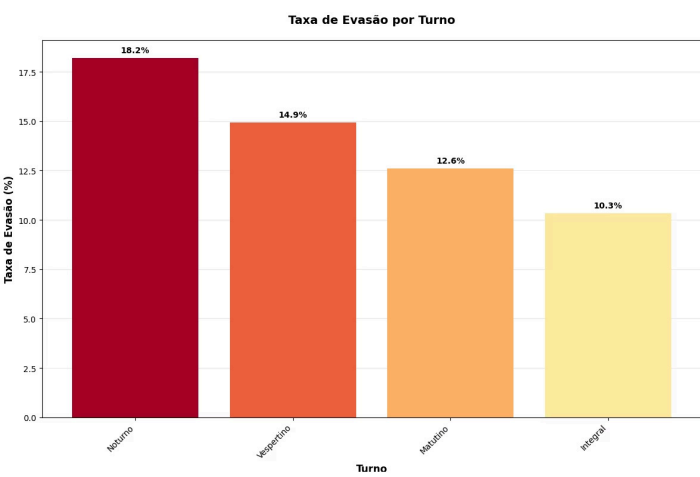
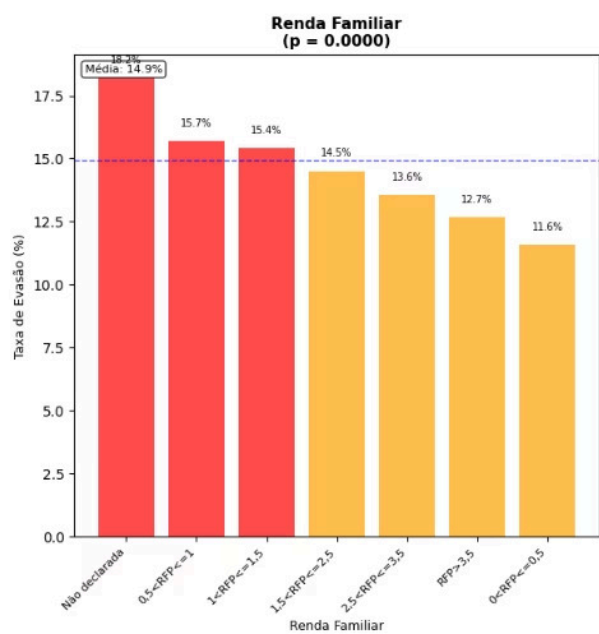
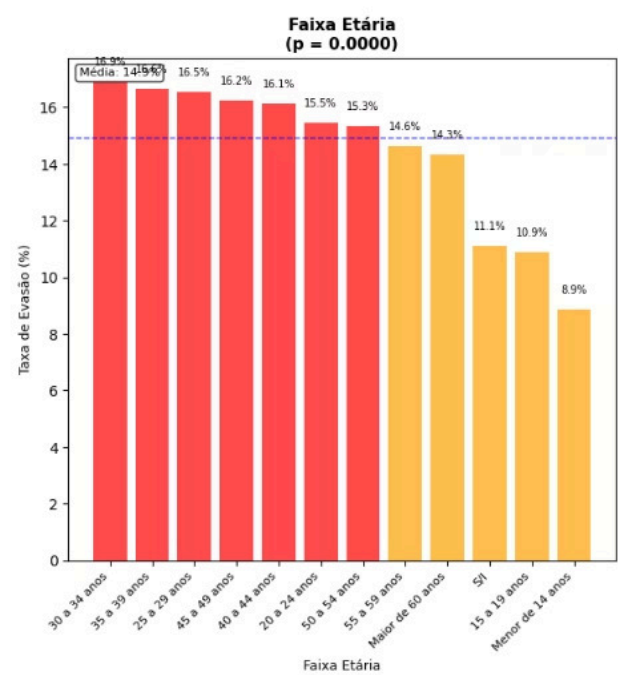
Dados reais

Insights do Dataset MEC

Dados com mais de 1 milhão de linhas para identificar fatores que influenciam a evasão escolar e construir previsões. Trabalhamos com 299 mil linhas.



Insights do Dataset MEC



=== Taxa predita por categoria -> Idade (top 10) ===

	Cod Unidade	n	taxa_predita	taxa_predita_pct
19	66.0	1	0.623333	62.333
7	22.0	5	0.289456	28.946
12	42.0	4	0.166667	16.667
5	19.0	1	0.126416	12.642
30	92.0	1	0.108958	10.896
18	65.0	1	0.108925	10.892
21	70.0	4564	0.102099	10.210
28	85.0	6	0.090625	9.063
8	29.0	1	0.074703	7.470
20	67.0	11	0.072012	7.201

=== Taxa predita por categoria -> Renda Familiar (top 10) ===

	Renda Familiar	n	taxa_predita	taxa_predita_pct
5	NÃO declarada	1396	0.123491	12.349
0	0,5<RFP<=1	795	0.094483	9.448
2	1,5<RFP<=2,5	467	0.093997	9.400
1	0<RFP<=0,5	1284	0.088903	8.890
4	2,5<RFP<=3,5	222	0.084004	8.400
6	RFP>3,5	196	0.075371	7.537
3	1<RFP<=1,5	639	0.071008	7.101

=== Taxa predita por categoria -> Turno (top 10) ===

	Turno	n	taxa_predita	taxa_predita_pct
3	Noturno	983	0.180700	18.070
4	Vespertino	461	0.149737	14.974
0	Integral	757	0.072016	7.202
1	Matutino	654	0.071655	7.165
2	NA_MISSING	2144	0.063580	6.358

Estrutura do código

Leitura e preparação do dataset

- Carrega o CSV.
- Detecta automaticamente os nomes das colunas.
- Cria a coluna EVADIDO (binário).

Seleção das features

- As colunas serão usadas para prever a evasão:
- Cor/Raça.
 - Eixo tecnológico.
 - Faixa etária.
 - Idade.
 - Renda familiar.
 - Sexo.
 - Turno.

Limpeza e transformação

- Idade e Renda** → convertidas para numéricas, imputando valores faltantes com a mediana.
- Categóricas** → valores nulos viram "NA_MISSING" e categorias muito raras são agrupadas em "OUTROS".
- One-hot encoding** → converte categorias em variáveis binárias.
- Padronização** → aplica escala em variáveis numéricas (como Idade).

Treinamento dos modelos

- Dois modelos são treinados:
- Logistic Regression** → modelo mais simples, baseline.
 - Random Forest** → modelo mais robusto e não linear, geralmente captura melhor relações complexas.
- Ambos produzem:
- Probabilidades de evasão (predict_proba)
 - Previsões binárias (0 = não evadiu, 1 = evadiu).



Avaliação

- Relatórios de classificação (accuracy, precision, recall, F1).
- Matriz de confusão.
- Curva ROC e AUC para comparar os modelos.
- Importância de variáveis (permutation importance no RandomForest).



Predição

- Para **cada categoria** (ex: cada raça, cada turno, cada faixa etária...), o código faz um **agrupamento**:
 - n: quantidade de registros naquela categoria.
 - taxa_predita: média da probabilidade de evasão naquela categoria (valor entre 0 e 1).
 - taxa_predita_pct: a mesma taxa, mas em **percentual** (0–100%).

OBRIGADO !!!

