

Relatório de Previsão de Evasão de Alunos: Estratégias para Aumentar a Retenção nos Primeiros 30 Dias

Introdução

A EdTech tem enfrentado um desafio significativo: a alta taxa de evasão de alunos nos primeiros 30 dias após a matrícula. Para lidar com isso, desenvolvemos uma solução de Ciência de Dados e Machine Learning capaz de prever quais alunos têm maior risco de abandonar o curso cedo. Este relatório apresenta os principais insights e recomendações práticas para ajudar a equipe pedagógica e de produto a reter mais alunos, transformando dados em ações estratégicas.

Quais perfis de aluno apresentam maior risco de evasão?

Nossa análise identificou padrões comportamentais e características que tornam alguns alunos mais propensos a evadir. Conhecer esses perfis é o primeiro passo para intervir de forma eficaz:

- **Baixo Engajamento Inicial:** Alunos com **poucas sessões totais** na plataforma (`total_sesoes`) nos primeiros dias são um forte indicativo de risco. Quanto menor o número de acessos, maior a chance de evasão.
 - **Não Conclusão do Primeiro Quiz:** Aqueles que **não conseguem ou não completam o primeiro quiz** do curso têm uma taxa de evasão ligeiramente maior (17.1%) em comparação com quem o conclui (13.8%). Este é um marco importante de engajamento e compreensão inicial.
 - **Dispositivo de Acesso:** Usuários de **"Tablet" (19.8%) e "Desktop" (17.4%)** apresentam taxas de evasão mais altas do que os usuários de **"Mobile" (10.2%)**.
 - **Categoria de Curso:** Alunos matriculados em cursos da categoria **"Design"** demonstraram a maior taxa de evasão (20.5%), enquanto **"Idiomas"** teve a menor (12.8%).
 - **Tipo de Assinatura:** Assinantes do tipo **"Grátis"** têm maior propensão à evasão (17.4%) se comparados aos assinantes **"Pago"** (12.1%).
 - **Nível Educacional "Não Informado" ou "Ensino Médio":** Alunos que não informaram seu nível educacional (18.7%) ou estão no **"Ensino Médio"** (16.9%) mostram taxas de evasão mais elevadas.
-

Quais ações poderiam ser tomadas para reter esses alunos nos primeiros dias?

Com base nos perfis de risco, sugerimos ações proativas e direcionadas:

- **Intervenções para Baixo Engajamento:**

- **Notificações Inteligentes:** Enviar lembretes automatizados (e-mail, push no aplicativo) para alunos que apresentarem baixo número de sessões ou pouca atividade nos primeiros 3-5 dias. A mensagem deve ser de apoio, oferecendo ajuda ou destacando o valor do curso.
 - **Contato de Tutores:** Acionar tutores para um contato proativo e personalizado com alunos identificados com risco de evasão devido ao baixo número de sessões. O tutor pode oferecer suporte, tirar dúvidas e incentivar a continuidade.
 - **Jornadas de Boas-Vindas Otimizadas:** Para todos os alunos, mas especialmente para os de risco, criar uma trilha de boas-vindas com pequenos desafios e conquistas que incentivem o engajamento e a familiarização com a plataforma nos primeiros dias.
 - **Suporte para o Primeiro Quiz:**
 - **Lembretes e Incentivos:** Enviar notificações específicas para alunos que ainda não concluíram o primeiro quiz, reforçando sua importância e oferecendo recursos de apoio.
 - **Sessões de Dúvidas:** Agendar sessões rápidas de tira-dúvidas ao vivo ou gravadas, focadas nos tópicos do primeiro quiz, para alunos que parecem ter dificuldades.
 - **Otimização por Dispositivo e Curso:**
 - **Experiência Multiplataforma:** Investigar possíveis pontos de atrito ou dificuldades na experiência do usuário em "Tablet" e "Desktop". Há alguma funcionalidade que funciona melhor no mobile e não é tão intuitiva em outras plataformas?
 - **Conteúdo Engajador para "Design":** Explorar o que torna os cursos de "Design" mais propensos à evasão. Talvez um "onboarding" mais aprofundado, exemplos práticos logo no início ou um suporte mais próximo para os alunos dessa categoria.
 - **Estratégias para Assinaturas "Grátis" e Nível Educacional:**
 - **Demonstração de Valor para Assinantes Grátis:** Destacar, nos primeiros dias, os benefícios e o valor agregado da assinatura paga, incentivando a transição antes da evasão.
 - **Identificação de Necessidades Específicas:** Para alunos com nível educacional "Ensino Médio" ou "Não Informado", considerar oferecer materiais de nivelamento ou apoio pedagógico adicional para garantir que eles se sintam aptos a progredir no curso.
-

Quais variáveis deveriam ser monitoradas ou coletadas com mais atenção?

Para refinar a identificação de risco e as estratégias de retenção, sugerimos um foco maior na coleta e monitoramento das seguintes variáveis:

- **total_sesoes e media_minutos_por_sessao:** Monitorar a frequência e a profundidade de uso é crucial. Podemos estabelecer gatilhos com base em quedas nesses indicadores.

- **concluiu_primeiro_quiz:** Essa é uma métrica binária de alto impacto. Monitorar a taxa de conclusão e os tempos de conclusão.
 - **dias_ativo_primeira_semana:** É importante ver a consistência da atividade do aluno.
 - **tipo_dispositivo:** Continuar monitorando qual dispositivo o aluno utiliza preferencialmente.
 - **categoria_curso, tipoassinatura, nivel_educacional:** Essas variáveis são essenciais para segmentar os alunos e personalizar as intervenções.
 - **Novas Variáveis a Considerar:** Para futuras melhorias do modelo e da compreensão do aluno, seria valioso coletar:
 - **Tempo médio de conclusão de módulos/aulas iniciais.**
 - **Participação em atividades síncronas (webinars, aulas ao vivo).**
 - **Número de vezes que o aluno acessa materiais de apoio (documentos, FAQs).**
 - **Metas ou expectativas que o aluno define ao se matricular (se possível de ser coletado).**
-

O modelo pode ser usado para disparar notificações automáticas ou acionar tutores? Com quais limitações?

Sim, o modelo desenvolvido tem um potencial significativo para **automatizar e otimizar as ações de retenção**:

- **Disparo de Notificações Automáticas:** O modelo pode ser integrado ao sistema da plataforma para, por exemplo, identificar um aluno em alto risco de evasão e automaticamente disparar um e-mail com dicas personalizadas ou um lembrete para acessar o curso.
- **Acionamento de Tutores:** A equipe de tutoria poderia receber alertas ou listas diárias/semanais de alunos com maior probabilidade de evasão. Isso permite que os tutores entrem em contato proativamente e ofereçam suporte direcionado, economizando tempo e recursos ao focar nos alunos que mais precisam de atenção.

No entanto, é crucial entender as limitações do modelo atual:

- **Falsos Negativos (75% dos casos de evasão não são detectados):** Nosso modelo, no momento, consegue identificar apenas **1 em cada 4 alunos que realmente irão evadir**. Isso significa que a grande maioria dos alunos que abandonam o curso ainda não será sinalizada pelo sistema. As intervenções, portanto, alcançarão apenas uma parcela dos alunos em risco.
- **Falsos Positivos (1 em cada 3 previsões de evasão está incorreta):** Embora a precisão para a classe "Evadiu" seja de 67%, ainda haverá **alarmes falsos**. Ou seja, para cada 3 alunos que o modelo prevê que vão evadir, aproximadamente 1 não irá. Isso pode levar a um pequeno desperdício de recursos (tempo dos tutores, envio de notificações desnecessárias), mas é um trade-off aceitável pela possibilidade de reter os alunos que realmente precisam.

- **Capacidade de Discriminação Moderada (AUC-ROC de 0.5990):** O modelo é um auxiliar, não um oráculo perfeito. Ele oferece uma boa direção, mas não substitui completamente o monitoramento humano e a adaptação das estratégias.
- **Necessidade de Validação Contínua:** O comportamento dos alunos e os padrões de evasão podem mudar ao longo do tempo. O modelo precisará ser monitorado regularmente, validado com novos dados e, se necessário, retreinado para manter sua eficácia.