



MODELO PREDITIVO DE EVASÃO ESCOLAR INTEGRADO A SISTEMAS ACADÊMICOS

Palavras-Chave: EVASÃO ESCOLAR, CIÊNCIA DE DADOS, APRENDIZADO DE MÁQUINA

Autores:

ADOLFO LEPE TONAKI, FATEC OURINHOS – FATEC

JOÃO LUCAS MENDES PEDROSO, FATEC OURINHOS – FATEC

ROSEMEIRY DE CASTRO PRADO, FATEC OURINHOS – FATEC

THAIS GARCIA, FATEC OURINHOS – FATEC

Prof. Dr. SIDNEY CARLOS FERRARI (orientador), FATEC OURINHOS – FATEC

INTRODUÇÃO:

A evasão acadêmica é um desafio enfrentado por diversas instituições de ensino superior, incluindo a Fatec Ourinhos, instituição pública de Ensino Superior, que integra a rede de instituições públicas do Centro Paula Souza, criada com o objetivo de oferecer formação tecnológica de nível superior.

Inicialmente, a Fatec Ourinhos esteve diretamente subordinada à Fatec São Paulo, funcionando como uma extensão de campus e reproduzindo as diretrizes administrativas e pedagógicas da unidade da capital (Prado, 2018). A iniciativa de sua criação partiu de um pedido da Prefeitura Municipal de Ourinhos, em 1991, ao Centro Paula Souza, com o objetivo de instalar uma unidade de ensino superior tecnológico no município. Em 09 de dezembro de 1997, é criada a Fatec Ourinhos, conforme Decreto do Governador Mário Covas, publicado em 10 de dezembro de 1997 (Prado, 2018).

Contudo, ao longo de sua existência, um problema sempre tem demandado uma atenção especial da faculdade: a evasão acadêmica, que possui aspecto complexo, sendo resultante de um produto da interação de vários fatores, tornando-se um conjunto que envolve questões financeiras, desajustes acadêmicos, problemas psicológicos e de escolha inapropriada da profissão, dentre tantos outros. Esses fatores podem ser agravados ainda mais por dificuldades institucionais, como a falta de um sistema de apoio pedagógico e políticas eficazes de permanência. Além disso, as circunstâncias socioeconômicas e culturais também devem ser levadas em consideração, o que influencia muito a decisão final do aluno (Santos et al., 2024).

Assim, pode-se definir a evasão escolar como sendo a desistência de estudantes antes da conclusão do curso. Segundo Riffel e Malacarne (2010), o termo evasão remete ao ato de

evadir-se, fugir ou abandonar. Quando se pensa em tal questão relacionada ao ensino técnico profissionalizante, a evasão tem implicações não apenas na trajetória educacional do aluno, mas também na formação de profissionais para o mercado de trabalho (Reis; Reis; Pereira, 2025, p. 2511).

Para Barboza (2024), as questões financeiras continuam sendo uma das principais razões para a desistência dos alunos. Contudo, também é importante mencionar que a falta de interesse no curso, uma educação anterior defasada e a incapacidade de equilibrar estudo e trabalho também são fatores determinantes (Barboza, 2024). “Além dos aspectos mencionados, a evasão escolar costuma apresentar maior incidência já no primeiro ano de curso” (Da Silva et al., 2022 apud Barboza, 2024, p. 224).

Compreender os fatores que levam os alunos a abandonarem seus cursos é essencial para a criação de estratégias eficazes de retenção. Dessa forma, o presente projeto aborda o desenvolvimento de um modelo preditivo que será capaz de identificar os estudantes com maior risco de evasão. A solução será integrada aos sistemas acadêmicos existentes, permitindo que a instituição de ensino tome medidas proativas para minimizar esse problema, além de alinhar os esforços da faculdade na modernização de seus sistemas acadêmicos, integrando a inteligência artificial e a Ciência de Dados na gestão acadêmica.

O objetivo geral do projeto é abordar o desenvolvimento e a integração de um modelo preditivo de evasão escolar baseado em *machine learning*, fornecendo subsídios para intervenções personalizadas que reduzirão a taxa de desistência dos alunos. Já os objetivos específicos têm como intenção, selecionar e testar algoritmos de *machine learning* para prever a evasão estudantil; levantar e analisar dados históricos de evasão; avaliar a eficácia do modelo e integrar o modelo ao Sistema Acadêmico da instituição (SIGA).

METODOLOGIA:

A metodologia deste estudo propõe uma abordagem sistematizada para o desenvolvimento de um modelo preditivo voltado à identificação da evasão escolar na Faculdade de Tecnologia de Ourinhos (Fatec Ourinhos). A proposta integra técnicas de análise exploratória de dados (EDA) com métodos de aprendizado de máquina supervisionado. A metodologia adotada é delineada em etapas sequenciais, abrangendo desde os procedimentos de coleta e a preparação dos dados até a avaliação e a interpretação dos resultados obtidos pelos modelos preditivos.

Com o objetivo de garantir a integridade e a consistência da base de dados, já foi realizado um processo rigoroso de limpeza e preparação. As etapas perpassam pela remoção de colunas consideradas irrelevantes para a análise, como a variável “unidade”; a padronização de cadeias de texto (eliminação de espaços em branco e acentuação); a conversão de tipos de dados,

especialmente no caso de variáveis numéricas associadas à frequência e ao desempenho acadêmico; o tratamento de valores ausentes, com ênfase nas colunas relativas às datas de cancelamento e conclusão; e a filtragem de cursos selecionados, de forma a delimitar o escopo da análise.

RESULTADOS E DISCUSSÃO:

Estão sendo realizados testados algoritmos como Regressão Logística, Random Forest, SVM, KNN e Naive Bayes, com foco na identificação de padrões associados à evasão. Além da modelagem, há a realizada de análise exploratória detalhada, destacando correlações entre frequência e desempenho, bem como os padrões geográficos e temporais de evasão. A partir dos resultados, observar-se-á o potencial de integração do modelo aos sistemas acadêmicos proporcionando visualizações úteis para a tomada de decisão.

Os modelos desenvolvidos até o momento, tais como Random Forest, SVM, Decision Tree e Logistic Regression alcançaram excelentes desempenhos, com valores elevados de F1 Score (acima de 0.95) e ROC AUC (próximos de 1), além de manterem níveis de overfitting baixos, o que garante maior confiabilidade na generalização para novos dados. O modelo KNN, apesar de apresentar métricas de desempenho semelhantes, teve um percentual de overfitting mais elevado (5,26%), sendo classificado como “Moderado”, o que reduziu sua classificação final para “Bom”. Já o Naive Bayes, mesmo com métricas um pouco inferiores, destacou-se por apresentar overfitting praticamente nulo, sendo também considerado “Excelente”. Esses resultados indicam que a maioria dos modelos é robusta, equilibrando alta performance e baixo risco de superajuste. As visualizações em gráficos reforçaram esses achados, evidenciando o perfil dos estudantes mais propensos à evasão. A ferramenta, ainda em desenvolvimento, poderá ser integrada futuramente aos sistemas acadêmicos da Fatec, possibilitando alertas precoces aos gestores acadêmicos.

RESULTADOS PRINCIPAIS:

	Model	F1 Score_test	ROC_AUC_test	Overfitting_%	Nível_Overfitting	Classificação_Final
1	Random Forest	0.98	0.99	1.93	★ BAIXO	★ EXCELENTE
2	SVM	0.98	0.99	0.44	★ BAIXO	★ EXCELENTE
4	Decision Tree	0.97	0.96	2.41	★ BAIXO	★ EXCELENTE
0	Logistic Regression	0.95	0.99	0.01	★ BAIXO	★ EXCELENTE
3	KNN	0.95	0.98	5.26	⚠ MODERADO	👍 BOM
5	Naive Bayes	0.93	0.96	-0.06	★ BAIXO	★ EXCELENTE

Figura 1 – Métricas de desempenho dos modelos

CONCLUSÕES:

Este estudo teve como foco central o desenvolvimento de um modelo preditivo capaz de antecipar casos de possível evasão entre estudantes da Fatec Ourinhos. Utilizando dados institucionais e técnicas de ciência de dados, foi possível identificar padrões recorrentes que indicam maior probabilidade de abandono, especialmente relacionados à frequência, rendimento acadêmico e perfil demográfico dos alunos.

Durante o processo, foram adotadas diversas etapas metodológicas, que incluíram desde a preparação e limpeza da base de dados até a aplicação e validação de diferentes algoritmos de aprendizado supervisionado. Estratégias como a criação de novas variáveis, o balanceamento das classes por meio do SMOTE e a utilização de validação cruzada contribuíram para que os modelos atingissem bons resultados, confirmando que os objetivos traçados inicialmente foram alcançados.

Como contribuição prática, o estudo demonstra a viabilidade de integrar esse tipo de modelo aos sistemas acadêmicos da instituição, oferecendo uma ferramenta de apoio à gestão educacional. A análise dos dados também trouxe à tona informações relevantes sobre o perfil dos estudantes que evadem, o que pode servir de base para a criação de estratégias mais direcionadas de prevenção.

Dessa forma, recomenda-se a continuidade deste trabalho, com a expansão da abordagem para outras unidades da rede Fatec e a experimentação com técnicas mais sofisticadas de modelagem. Espera-se que este estudo possa colaborar com a construção de políticas mais eficazes de permanência estudantil, valorizando a tomada de decisão baseada em dados e fortalecendo a atuação preventiva da instituição diante da evasão.

BIBLIOGRAFIA

BARBOZA, R. C. Evasão Escolar no Ensino Superior em Instituições Públicas e Privadas no Brasil. *Revista RG&PP*, 2024.

CHICON, P. M. M. et al. Análise da construção de modelos preditivos sob a perspectiva de indicadores de evasão. *RENOTE*, 2021.

DA SILVA, M. C. et al. Estudo longitudinal da evasão no ensino superior paulista. In: BARBOZA, L. A. (Org.). *Evasão no ensino técnico e superior: causas, impactos e soluções*. São Paulo: Editora Acadêmica, 2024. p. 223-239.

GARCIA, R. M. et al. Machine Learning aplicado à análise de evasão escolar. *Journal of Educational Data Science*, 2021.

HAN, J.; KAMBER, M.; PEI, J. *Data Mining: Concepts and Techniques*. 4. ed., 2022.

MEDEIROS, L. F.; ALMEIDA, T. R. *Predição de evasão no ensino superior utilizando IA*. R.SITE, 2021.

PRADO, Rosemeiry de Castro. *As Faculdades de Tecnologia do Estado de São Paulo: um histórico da instituição e aspectos relativos ao ensino de Matemática nela praticado*. 2018. 369 f. Tese (Doutorado em Educação para a Ciência) – Faculdade de Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Bauru, 2018.

REIS, C. A.; REIS, R. S.; PEREIRA, L. M. Evasão escolar no ensino técnico: um desafio para o mercado de trabalho. *Revista Brasileira de Educação Profissional e Tecnológica*, Brasília, v. 17, n. 2, p. 2508-2516, 2025.

RIFFEL, R.; MALACARNE, V. M. Evasão escolar: conceitos, causas e desafios para a educação profissional. *Educação em Foco*, Juiz de Fora, v. 15, n. 1, p. 45-58, 2010.

SANTOS, J. P. et al. A permanência estudantil como política institucional: desafios e caminhos. *Revista de Políticas Educacionais*, Salvador, v. 9, n. 1, p. 90-112, 2024.