MINERAÇÃO DE DADOS EDUCACIONAIS CONSIDERANDO A REDE FEDERAL DE ENSINO

VERSIANE, Felipe Fernandes[1]; SOUSA, Daniel Xavier de[1*]

[1]Instituto Federal de Goiás, Câmpus Anápolis,
[*]daniel.sousa@ifg.edu.br.

Este estudo teve como objetivo aplicar técnicas de mineração de dados no contexto educacional, buscando identificar padrões e tendências que possam aprimorar o processo de tomada de decisão do corpo docente e administrativo do Instituto Federal de Goiás (IFG). A crescente disponibilidade de dados educacionais e a necessidade de utilizá-los de forma eficaz para gerar insights fundamentaram a pesquisa. O desenvolvimento do trabalho envolveu a coleta de dados de diversas bases públicas, abrangendo informações acadêmicas e socioeconômicas dos alunos, seguidas de um rigoroso tratamento para assegurar a qualidade dos dados que seriam posteriormente utilizados nas análises e na descoberta de padrões por meio de métodos de Aprendizado de Máquina. Dentre as técnicas de pré-processamento, foram adotadas estratégias para o tratamento de variáveis nulas, identificação e remoção de outliers, agregação de colunas e aplicação de métodos para o tratamento de variáveis categóricas, além do balanceamento dos dados no contexto de machine learning. A aplicação do método de Aprendizado de Máquina teve como finalidade identificar as principais variáveis que influenciam a evasão dos alunos no instituto, permitindo uma compreensão mais aprofundada do perfil de evasão. As evasões ocorrem, principalmente, nos primeiros anos e períodos, resultando em um número significativamente reduzido nos últimos anos. Em virtude disso, observou-se uma queda na acurácia do nosso modelo ao decorrer dos anos, no entanto, essa redução não se deu de forma proporcional à diminuição do número de evasões citadas, o que evidencia a forte relevância dos dados analisados. Dessa forma, os dados revelam padrões informativos que permanecem significativos, mesmo diante da diminuição da taxa de evasão. Para além dessas análises, conseguimos construir diversos gráficos de alunos para que possam apoiar as gestões acadêmicas do IFG, mostramos gráficos que informam o quantitativo de alunos por turmas, as porcentagens de alunos que recebem auxílios da instituição, e números de formandos.

Palavras-chave: mineração de dados educacionais, análise de dados, modelos de predição.

Agradecimentos: O presente trabalho foi realizado com apoio do Instituto Federal de Goiás (nº20/2023). Versiani, Felipe Fernandes agradece a bolsa concedida pelo CNPQ.