

Núcleo de Avaliação: Núcleo 2

Área temática: Ciências Exatas e da Terra

Área do Conhecimento: Ciência da computação

Uma revisão sistemática sobre mineração de dados educacionais para a prevenção de insucesso e abandono escolar.

Alice dos Santos Pinheiro, Reudismam Rolim de Sousa, Geiser Chalco Chalco

A evasão escolar é um problema significativo que impacta o desempenho educacional e as oportunidades futuras dos estudantes. Neste sentido, este trabalho teve como objetivo identificar as técnicas de aprendizado de máquina mais utilizadas na predição de evasão escolar e os fatores mais comuns empregados como preditores nessas abordagens. Com base nesses resultados e fazendo-se uma análise dos fatores presentes nos microdados fornecidos pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), busca-se determinar que dados e técnicas podem ser usados para a predição da evasão. A pesquisa foi realizada usando a metodologia de revisão sistemática da literatura na base de dados do Scopus e foram incluídos estudos publicados entre 2010 e 2023. Inicialmente, 751 artigos foram selecionados, e após a aplicação dos critérios de exclusão, o número foi reduzido para 9, que foram utilizados na análise final. Dos artigos selecionados, constatou-se que, entre os 13 algoritmos de machine learning empregados (AdaBoost Classifier, Artificial Neural Networks, Decision Tree, Gradient Boost, K-Nearest Neighbors, Linear Discriminant Analysis, Logistic Regression, Multilayer Perceptron, Naïve Bayes, Random Forest, Sequential Minimal Optimization, Stochastic Gradient e Support Vector Machine), três tiveram maior destaque: Random Forest, Support Vector Machine e Neural Artificial Network. Com base nos fatores mais utilizados como preditores de abandono nos artigos revisados, foram realizadas análises dos microdados do Censo Escolar, ENEM e Ensino Superior, fornecidos pelo INEP em 2022. Tal abordagem permitiu identificar que as características significativas que influenciam a evasão escolar são os dados demográficos, como idade e gênero, além de condições socioeconômicas e medidas de desempenho acadêmico. Os resultados desta revisão destacam o potencial de determinados algoritmos de machine learning para futuras análises de EDM, além de evidenciar a importância de considerar preditores significativos, como dados demográficos e socioeconômicos, na compreensão da evasão escolar.

Palavras-chave: EDM, dropout, machine learning, evasão escolar, data mining



Agência financiadora: PIVIC (IC)

Campus: Pau dos Ferros
