

A Inteligência Artificial como ferramenta para redução da Evasão Escolar

Fábio de Albuquerque Silva
Instituto Federal da Paraíba
fabio.albuquerque@ifpb.edu.br

Matheus Cordeiro de Melo
Instituto Federal da Paraíba
matheus.cordeiro@ifpb.edu.br

Victor Régis Beserra da Silva
Instituto Federal da Paraíba
victor.regis@ifpb.edu.br

Resumo

A evasão escolar é um problema que atinge instituições de ensino públicas e privadas gerando perdas financeiras e sociais. No Instituto Federal da Paraíba, segundo dados da plataforma Nilo Peçanha, a taxa de Evasão é de 16,8% no nível médio. Diante disto, a utilização de métodos de aprendizagem de máquina para identificar alunos com potencial de evasão, pode ser um ponto de partida importante para prever o perfil dos estudantes que evadem da nossa instituição, e a partir daí, traçar ações voltadas à prevenção/reversão desta evasão. Com este objetivo, foram extraídos dados de alunos do ensino médio que frequentaram os cursos do *Campus* João Pessoa do IFPB nos anos de 2015 a 2018, utilizando-se os dados de 2015 a 2017 para o treinamento do modelo preditivo e os dados de matrículas em 2018 para testar a eficiência do modelo em identificar os alunos com potencial de evasão. A base de treinamento continha, inicialmente, 4.195 alunos, sendo que a base utilizada para a construção do modelo ficou com 2.207 registros, devido às exclusões de registros na fase de pré-processamento/limpeza. A base de treinamento obtida continha 20 atributos iniciais, compostos por dados pessoais, familiares, escolares e socioeconômicos. Para a redução da dimensionalidade do modelo foi verificada a correlação destes atributos com a classe alvo – evasão = SIM. Com isto, ficamos com 5 atributos: idade, periodo_atual, cod_curso, modalidade, ira. Realizamos uma análise comparativa entre 6 classificadores para identificar o que apresentasse a melhor acurácia na identificação dos alunos com potencial de evadir da nossa instituição. O classificador K-Nearest Neighbors (KNN) apresentou a melhor acurácia (83%) para a base de treinamento e, com a utilização da base de testes (matrículas em 2018), demonstrou que o modelo obtido tem um alto grau de acurácia (85%) para prever evasão antes que aconteça.

Palavras-Chave: Evasão Escolar. Aprendizagem de Máquina. Inteligência Artificial.

Artificial Intelligence as a Tool to Reduce School Evasion

Abstract

School dropout is a problem that affects public and private educational institutions generating financial and social losses. At the Federal Institute of Paraíba, according to data from the Nilo Peçanha platform, the Evasion rate is 16.8% at the average level. Therefore, the use of machine learning methods to identify students with potential for evasion can be an important starting point for predicting the profile of the students who leave our institution, and from there, to map out prevention / reversion actions of this evasion. To this end, data were extracted from high school students who attended the courses of the IFPB campus João Pessoa in the years 2015 to 2018, using the data from 2015 to 2017 for the training of the predictive model and enrollment data in 2018 to test the model's effectiveness in identifying students with dropout potential. The training base initially contained 4,195 students, and the base used for the construction of the model was 2,207 records, due to the exclusions of registrations in the pre-processing / cleaning phase. The training base obtained contained 20 initial attributes, composed of personal, family, school and socioeconomic data. To reduce the dimensionality of the model, we verified the correlation of these attributes with the target class - avoidance = SIM. With this, we have 5 attributes: age, current_time, cod_current, modality, anger. We performed a comparative analysis among 6 classifiers to identify the one that presented the best accuracy in the identification of students with potential to evade our institution. The K-Nearest Neighbors (KNN) classifier presented the best accuracy (83%) for the training base and, using the test database (enrollment in 2018), demonstrated that the model obtained has a high degree of (%) to predict evasion before it happens.

Keywords: School Evasion, Machine Learning, Artificial Intelligence.