

Aplicação para previsão de abandono estudantil

Alex Se Yoon Koo¹, Gabriel Novaes², João Victor Pacheco³

¹Faculdade de Computação e Informática - Universidade Presbiteriana Mackenzie (UPM)

Abstract. *Student dropout is a complex and multifaceted phenomenon, which refers to students giving up or interrupting their studies before completing a course. The causes of this dropout can vary, including socioeconomic factors, mental health problems, lack of motivation, lack of educational support, challenges in balancing work and study, among others. This problem affects both basic and higher education, and has serious implications for individual and social development. Student dropout can result in economic losses, lower professional qualifications, increased social inequalities and limitations in career opportunities. Therefore, this research aims to create a classification algorithm to define whether a student will complete the course or not.*

Resumo. *O abandono estudantil é um fenômeno complexo e multifacetado, que se refere à desistência ou interrupção dos estudos por parte dos alunos antes da conclusão de um curso. As causas desse abandono podem variar, incluindo fatores socioeconômicos, problemas de saúde mental, desmotivação, falta de suporte educacional, desafios na conciliação de trabalho e estudo, entre outros. Esse problema afeta tanto o ensino básico quanto o superior, e tem implicações graves para o desenvolvimento individual e social. O abandono estudantil pode resultar em perdas econômicas, menor qualificação profissional, aumento das desigualdades sociais e limitações nas oportunidades de carreira. Com isso, esta pesquisa tem como objetivo de criar um algoritmo de classificação para definir se um aluno concluirá o curso ou não.*

1. Introdução

1.1. Contextualização

No cotidiano que vivemos a educação apresenta ser algo importante ou crucial, desenvolvendo conhecimento e pessoas, além de expandir o horizontal. Porém isso não apresenta ser algo que ocorre em todo o momento uma vez que mesmo em situações de educação ocorre casos de evasão, onde perdemos talentos por motivos econômicos ou sociais.

1.2. Justificativa

Diante disso, com a evasão escolar aumentando cada vez mais com o tempo, alunos deixando os estudos, perdemos a uma quantidade de talentos potenciais para a pesquisa ou acadêmico como também no âmbito profissional, fazendo-se assim a desvalorização da cultura brasileira nesse requisito.

1.3. Objetivo

Esta pesquisa tem como objetivo de fazer uma classificação dos alunos da chance de graduar ou não o curso de bacharelado presente de acordo com os seus atributos sociais e econômicos como uma forma de prevenir a evasão escolar.

1.4. Opção do projeto/Objetivo Específico

Como objetivo específico, temos como alvo de, além da classificação, buscar entender quais são os principais fatores para evasão escolar e a correlação entre eles, e se diferentes algoritmos de classificação afeta a probabilidade final.

2. Descrição do Problema

In some conferences, the papers are published on CD-ROM while only the abstract is published in the printed Proceedings. In this case, authors are invited to prepare two final versions of the paper. One, complete, to be published on the CD and the other, containing only the first page, with abstract and “resumo” (for papers in Portuguese).

3. Dataset

O dataset que utilizamos foi dados csv disponível na Kaggle sobre evasão escolar. O dataset possui os seguintes atributos:

Estado Civil (Marital Status): O estado civil do estudante (por exemplo, solteiro, casado, divorciado).

Modo de Inscrição (Application Mode): Refere-se ao modo ou tipo de inscrição que o estudante submeteu para se matricular no curso.

Ordem de Inscrição (Application Order): Indica a ordem em que o estudante se candidatou ao curso. Por exemplo, se foi a primeira, segunda ou terceira opção do estudante.

Curso (Course): O curso ou programa de graduação em que o estudante está matriculado (por exemplo, Ciência da Computação, Engenharia, etc.).

Turno de Frequência (Daytime/Evening Attendance): Especifica se o estudante frequenta o curso durante o dia ou à noite, representando seu horário de frequência.

Qualificação Anterior (Previous Qualification): O tipo de qualificação acadêmica que o estudante tinha antes de se matricular no curso (por exemplo, diploma de ensino médio, curso técnico).

Qualificação Anterior (Nota) [Previous Qualification (Grade)]: A nota final ou pontuação associada à qualificação anterior do estudante.

Nacionalidade (Nationality): A nacionalidade do estudante.

Qualificação da Mãe (Mother's Qualification): O nível mais alto de qualificação acadêmica alcançado pela mãe do estudante.

Qualificação do Pai (Father's Qualification): O nível mais alto de qualificação acadêmica alcançado pelo pai do estudante.

De acordo com a análise exploratória de dados, podemos observar que metade dos alunos tem chance de terminar o curso de graduação e 0.32 de porcentagem em largar a faculdade de acordo com a figura 3.

Também é possível observar a distribuição dos dados de acordo com a figura 4, onde criamos gráficos entre dois atributos quaisquer para ver como está a distribuição dos dados.

Também realizamos um bucket do target label para cada coluna para observar o comportamento dos dados para cada coluna como mostra na figura 1 e 2.

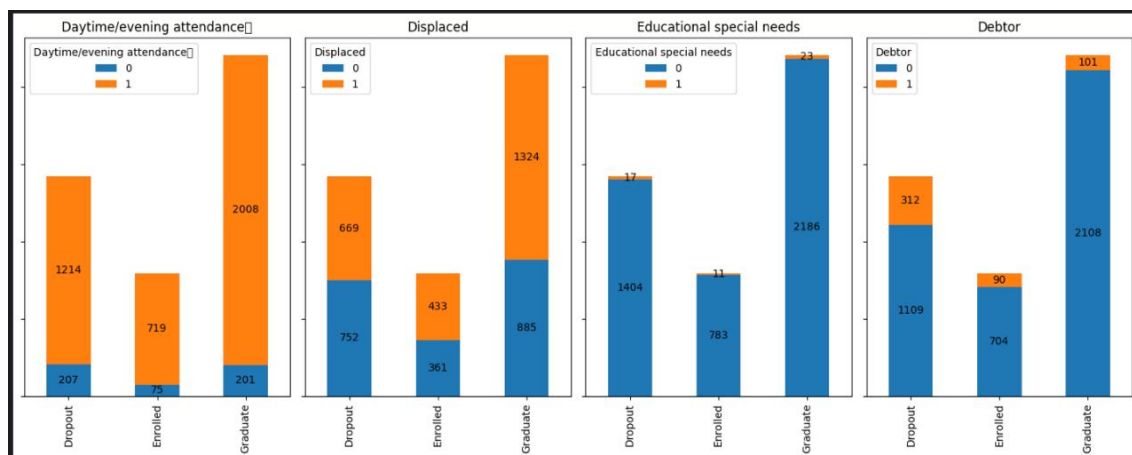


Figura 1. Bucket dos alunos de acordo com cada coluna.

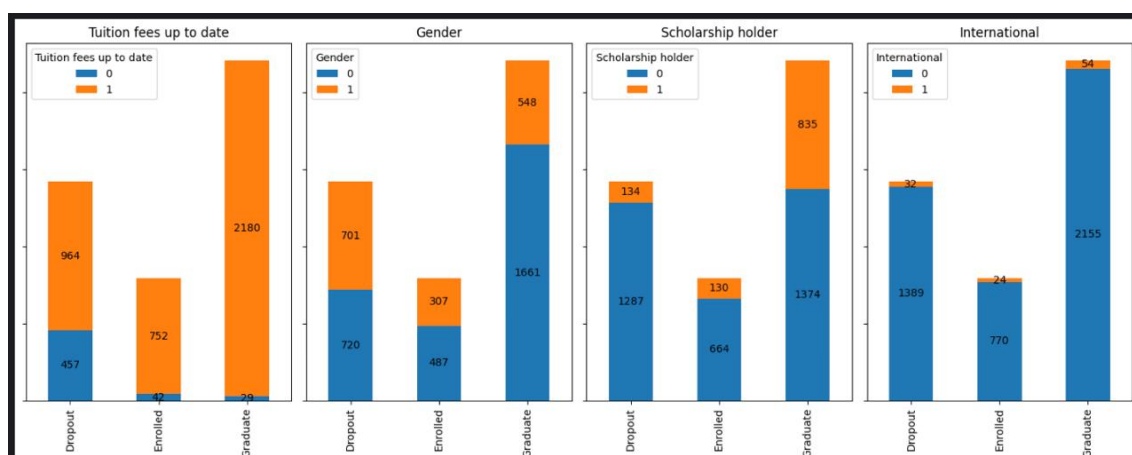


Figura 2. Bucket dos alunos de acordo com cada coluna.

4. Metodologia e Resultado Esperado

Nesta seção buscamos desdobrar a metodologia adotada e os materiais e estimar o resultado esperado da nossa pesquisa.

Como primeira etapa, buscamos estudos pesquisas anteriormente já realizadas como uma forma de referência e definir quais serão os próximos passos, em seguida na segunda etapa, buscamos coletar e interpretar os dados usando análise exploratória dos dados e definir quais os resultados que podemos obter. Na terceira etapa, realizamos o pré-processamento dos dados, limpando os dados, seja na sua duplicidade ou formatação, e consequentemente, na quarta etapa separamos os dados e realizamos a validação cruzada como uma forma de diminuir as chances de overfitting durante o treinamento. Na quinta etapa, buscamos algoritmos que podemos utilizar e comparar entre eles como também as métricas que iremos utilizar para avaliar o nosso modelo. Na sexta etapa, buscamos implementar o nosso algoritmo para treinamento e predição e como última etapa analisar os dados.

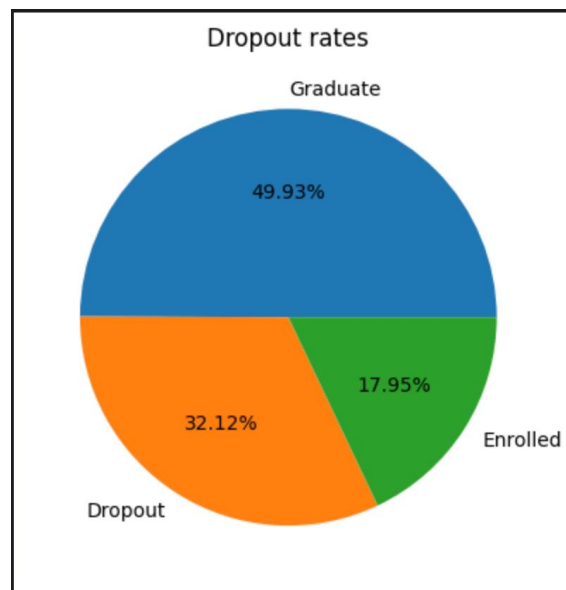


Figura 3. Gráfico de pizza mostrando a porcentagem de alunos em no label target.

5. Referência e Bibliografia

5.1. Fatores associados ao abandono escolar no ensino médio público de Minas Gerais

Esse artigo aponta os principais fatores que influenciam o abandono escolar no ensino médio, onde os dados se baseiam em uma série histórica constituída a partir do levantamento das pesquisas nacionais por amostra de Domicílios.

5.2. Abandono e evasão escolar sobre a ótica dos sujeitos envolvidos

O artigo busca refletir sobre as concepções relativas ao abandono e à evasão escolar por parte de agentes públicos que atuam junto à população juvenil em um município de porte médio do interior paulista.

5.3. A problemática da evasão escolar: Uma revisão bibliográfica integrativa

O artigo visa a educação no Brasil que enfrenta diversos problemas que tem onde ele aborda a importância do direito a educação e mostra as principais causas que estão ligadas à evasão escolar.

5.4. Abandono escolar: Entendendo as causas e buscando soluções

Nesse artigo ele mostra dados de acordo com o senso e a OCDE onde ele explica a realidade do abandono escolar e fala sobre a diferença entre abandono escolar e evasão escolar.

5.5. Evasão e abandono escolar na educação básica no Brasil: fatores, causas e possíveis consequências

O artigo objetiva trazer algumas considerações sobre a evasão e abandono escolar na educação brasileira e mostra diversos fatores entre eles é dizer que existem três dimensões conceituais para a investigação da evasão escolar.

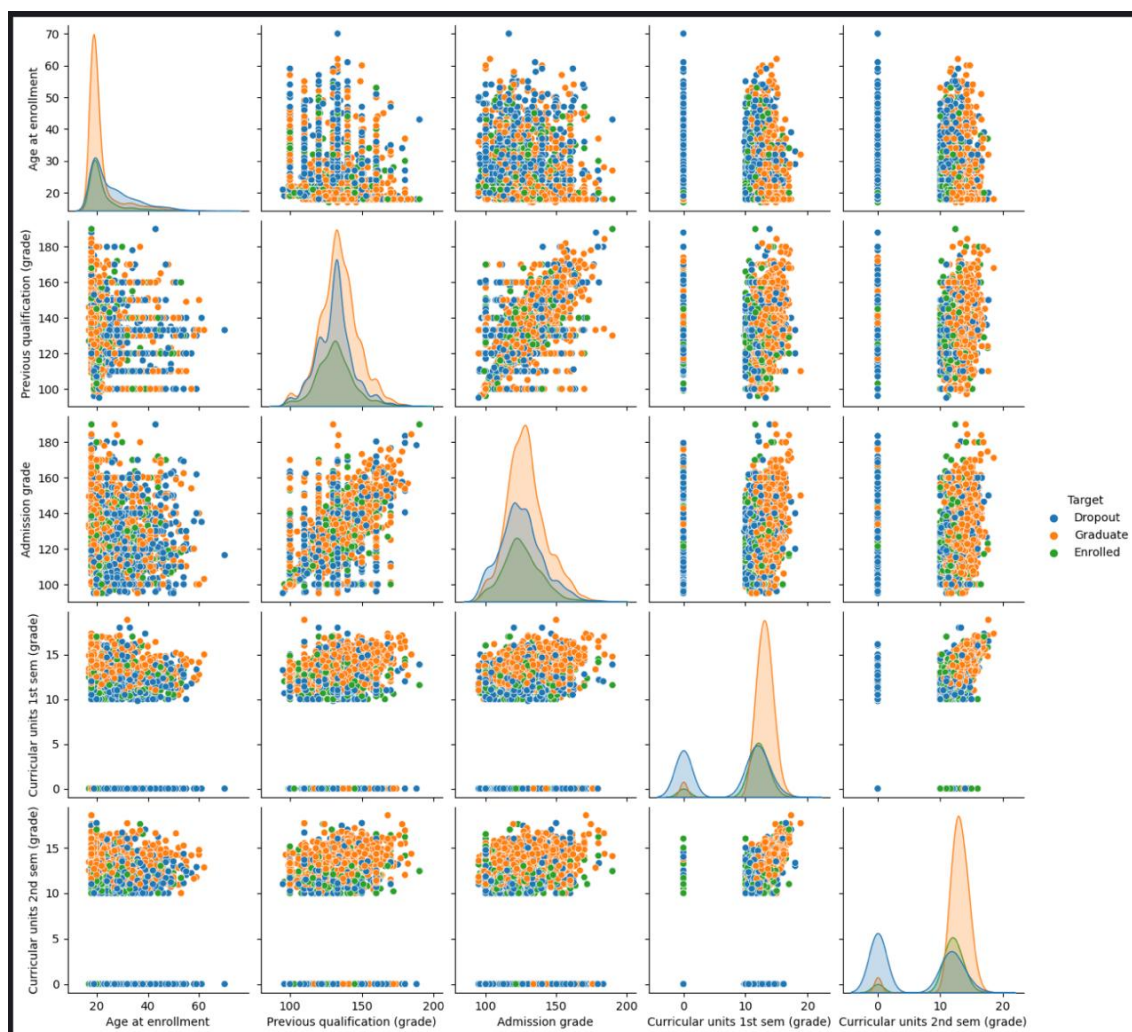


Figura 4. Correlação entre duas colunas ou atributos.

Referências

Brasil: fatores, causas e possíveis consequências Dropout and abandonment of school in Brazil: factors, causes and possible consequences Raimundo Barbosa Silva Filhoa. RONALDO MARCOS DE LIMA ARAÚJO: [s.n.].

RAMOS, A. C.; GONÇALVES JUNIOR, O. Abandono e evasão escolar sob a ótica dos sujeitos envolvidos. **Educação e Pesquisa**, v. 50, p. e268037, 2024.

SENNA, I. A. **Abandono escolar: entendendo as causas e buscando soluções**. Disponível em: <https://institutoayrtonsenna.org.br/abandono-escolar/>. Acesso em: 29 set. 2024.

SOARES, T. M. et al. Fatores associados ao abandono escolar no ensino médio público de Minas Gerais. **Educação e Pesquisa**, v. 41, n. 3, p. 757–772, 2015.

LINO, R. E.. A Problemática da evasão escolar: uma revisão bibliográfica interativa. **PUC-Goiás**, 2020.