

ANÁLISE DE FATORES DETERMINANTES DE AUSÊNCIA NO EXAME NACIONAL DO ENSINO MÉDIO (ENEM) ATRAVÉS DA APLICAÇÃO DE MODELOS LOGIT E CROSS-VALIDATION

XXXI Encontro de Iniciação à Docência

Gabriel Cordeiro Holanda, João Victor Batista Lopes, Marcelo de Castro Callado

O Exame Nacional do Ensino Médio (ENEM) foi criado em 1998 com o intuito de estabelecer uma prova padrão para a admissão em universidades públicas no Brasil. A edição de 2021 contou com cerca de 3,1 milhões de participantes inscritos, mas uma taxa recorde de 26% de ausentes. Serão aplicadas técnicas de cross-validation dentro de um modelo Logit através de Machine Learning. O objetivo é encontrar quais fatores são determinantes de probabilidade de um participante ausentar-se. Para isso utilizaremos a base de microdados dos participantes fornecida pelo INEP em domínio público, a qual contém informações demográficas e sociais de cada participantes, bem como seu resultado nas provas aplicadas e se o mesmo compareceu ou não a cada dia de aplicação. Esperamos que os determinantes sejam os que já aparecem na literatura como variáveis de risco, ou seja: gênero, renda e raça afetarão negativamente a chance de presença quando um candidato for racializado e/ou mulher e quanto menor sua renda. Também espera-se que essas características tenham efeitos cumulativos, ou seja, uma mulher negra tem menor chance de presença que uma mulher branca, que por sua vez tem menor chance de presença que um homem branco. Ademais, será efetuada revisão da literatura disponível em relação ao tema para podermos solidificar nossas hipóteses, evitar inferências precipitadas e extrair conclusões com o devido embasamento.

Palavras-chave: Enem. Logit. Machine Learning.