

USO DA MINERAÇÃO DE DADOS PARA PREDIÇÃO DE REPROVAÇÃO EM SELEÇÃO PARA RESIDÊNCIA MÉDICA.

Taciana Barbosa Duque¹, Seiji Isotani², Bruno Elias Pentead¹

¹ Pós-Graduação EAD em Computação Aplicada à Educação (ICMC/USP)

² Universidade de São Paulo (USP)

1. INTRODUÇÃO

A mineração de dados educacionais se concentra em dados provenientes do contexto educacional; busca entender como se dá a aprendizagem e que aspectos podem contribuir. Previsões baseadas em desempenho de grupos específicos compõem o grande *pool* de pesquisas e aplicações da mineração de dados na educação. Dentro dessa perspectiva, destacamos a residência médica que constitui uma etapa importante no processo de formação continuada do médico. A predição mais precoce de insucesso através de indicadores de avaliação de desempenho durante a graduação pode trazer benefícios com mudança no planejamento de ações educacionais específicas visando reverter esse resultado.

2. OBJETIVO DO ESTUDO

Identificar se os resultados da performance dos estudantes de medicina obtidos do sistema de avaliação durante os primeiros quatro anos de curso são capazes de prever o desfecho no processo de seleção para a residência médica.

3. MATERIAS E MÉTODOS

Foi utilizado o algoritmo de árvore de decisão Weka-classifiers-trees J48. Os atributos foram compostos pelo conjunto de notas obtidas por estudantes de medicina de uma Faculdade no Nordeste do Brasil, que concluíram o curso nos anos de 2018 e 2019. A reprovação na Residência Médica foi a variável preditiva.

Não existiam dados faltantes e todas as variáveis numéricas (notas) estavam na mesma escala, de 0 a 10. Foram utilizados os parâmetros padrão do Weka versão 3.7: valor do *confidence factor* foi 0,25, número mínimo de instâncias por nó (*minNumObj*) igual a 2, divisão entre treinamento e teste foi por cross-validation utilizando-se o método *K-fold* com 10 subconjuntos.

4. RESULTADOS e DISCUSSÃO

- Entre os 190 concluintes, 97 (51,05%) foram reprovado na RM, sendo a classe majoritária para a qual buscou-se estudar um modelo preditivo
- A melhor acurácia foi para os atributos notas obtidos no quarto ano (67,8947%).
- As acurácias do algoritmo J48 em cada um dos períodos do quarto ano foram semelhantes, sendo iguais a 65,26%.
- O modelo final do 8º período, trouxe Módulos relacionados à clínica do adulto: estudo das doenças do sistema Nervoso (MAB08-71), estudo das doenças hematológicas (MAB08-70) e um módulo de atividade prática, em cenário real (MAB08-00), onde o estudante na fase pré-internato, realiza atendimento a pacientes no hospital de ensino, sob supervisão, e as avaliações se dão pela observação direta do desempenho do estudante nessa atividade, incluindo todas as etapas da consulta.

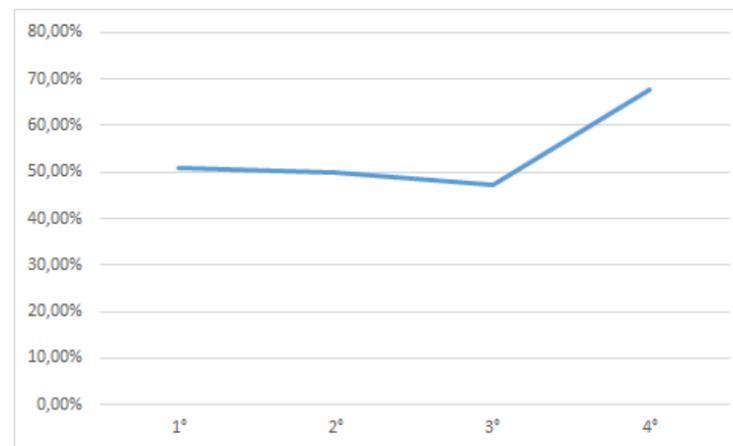


Figura 1 - Acurácia do algoritmo J48 testado em base de dados das médias por ano de curso de estudantes de medicina no período pré-internato para predição do desfecho reprovação no processo seletivo para Residência Médica.

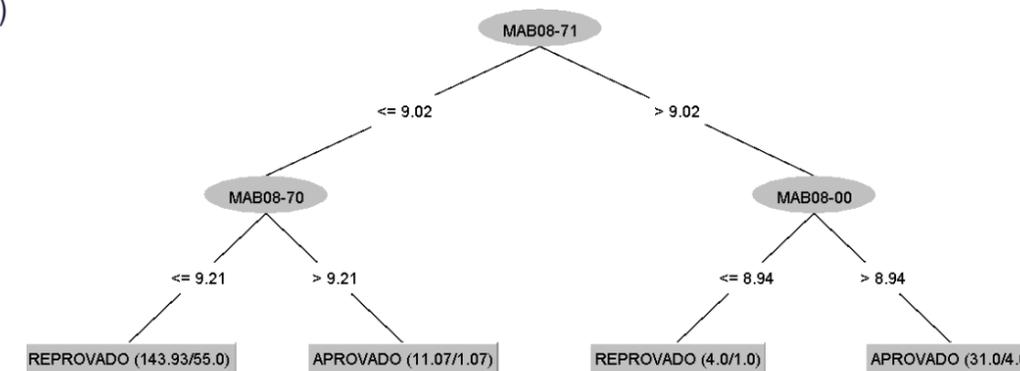


Figura 2 - Árvore de decisão final referente às notas do 4º ano - 8º período.

- Amjed Abu Saa et al [2019] -revisão sistemática identificaram que a utilização de resultados de avaliação curricular foi o critérios mais frequentemente estudado, utilizando a técnica de árvore de decisão.
- Creton e Gomes [2018] aplicaram a técnica de mineração de dados na base do ENADE relacionada aos resultados obtidos por estudantes de medicina com ênfase no nível de dificuldade apontado pelos estudantes
- Aragão (2018) –Importância do estudo da Taxonomia das questões da prova de Residência médica.
- Pugh, [2016], Sesato, [2017];Yassin, [2018]- estudos sobre desempenho de estudantes de Medicina mas utilizando análises estatísticas

5. CONCLUSÃO E COMENTÁRIOS

O modelo proposto apresenta como melhor momento de predição o quarto ano do curso que corresponde ao período imediatamente anterior ao internato.

O número de dados gerados em um período de doze semestres de um curso de medicina, pode responder a muitas indagações com a utilização da MDE.