

Machine Learning como ferramenta de auxílio na predição de rendimento de alunos do curso de Sistemas para Internet

Ariel Corrêa Pezzoli, Karen Selbach Borges (orientador)

Afiliação: Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul –
Campus Porto Alegre

ariel_pezzoli@hotmail.com, karen.borges@poa.ifrs.edu.br

Na educação superior, especialmente em cursos de TI, existe muita evasão e retenção de alunos. Como hipótese, acredita-se que um dos motivos é que os alunos ingressantes nos cursos ainda não se encontram no último estágio de desenvolvimento cognitivo, o qual foi identificado por Jean Piaget como Operatório Formal. Diversos trabalhos estudam os motivos pelos quais os alunos têm dificuldades, porém, em sua maioria, esses estudos levam em consideração aspectos sociais e desempenho acadêmico. Este trabalho tem como objetivo criar um sistema que irá aplicar técnicas de *machine learning* sobre dados obtidos a partir do desempenho dos alunos na disciplina de Lógica de Programação do IFRS e nos resultados dos Testes Operatórios de François Longeot, os quais irão auxiliar na identificação do nível cognitivo dos alunos, concreto ou formal. Espera-se, assim, identificar aqueles com propensão à evasão ou retenção e, com isso, contribuir para melhorar a qualidade da formação dos futuros profissionais de TI. Ao se identificar o nível em que cada aluno recém-ingressado no curso se encontra, torna-se possível que a instituição ofereça formas para ajudar àqueles que ainda não se encontram no nível formal, seja com aulas extras, acompanhamento psicológico, mudanças nas metodologias de aprendizagem ou outras técnicas que a instituição julgue pertinente. Para a execução do sistema foram utilizadas técnicas de mineração de dados e aprendizado de máquina, utilizando-se de algoritmos de classificação e clusterização para identificar padrões nos dados em questão, além da necessidade da utilização da metodologia KDD (*Knowledge Discovery in Databases*), necessária para preparar os dados que serão utilizados no algoritmo de *machine learning*. Ao final do projeto, os algoritmos de aprendizado de máquina elaborados para o sistema, acabaram por não conseguir encontrar padrões relevantes capazes de estabelecer uma relação entre os testes operatórios e as notas finais da disciplina. Por conseguinte, também não foi possível alcançar uma porcentagem relevante de acerto nas previsões, de novos alunos, do sistema. Há diversos casos em que alunos com notas satisfatórias para aprovação acabam por reprovar ou evadir. Há também casos em que o aluno não tem um bom rendimento e acaba por aprovar. A conclusão que se chega é que os algoritmos não foram capazes de encontrar um padrão bem definido se baseando apenas nas notas obtidas dos testes operatórios, não conseguindo identificar uma relação entre o nível cognitivo dos alunos com a aprovação da disciplina. Existem diversos outros fatores, alheios aos resultados acadêmicos e ao nível cognitivo, como renda, local onde o aluno reside, situação familiar, situação psicológica, entre outros, que afetam o desenvolvimento dos alunos, não sendo suficiente, portanto, apenas o uso dos dados utilizados por este trabalho. Cabendo assim, em trabalhos futuros, a adição de mais dados aos modelos além de novos atributos.

Palavras-Chave: Mineração de dados; Testes Operatórios;

Financiamento/Apoio: LECC (Laboratório de Estudos Cognitivos Apoiados por Computação).