





















Como resultados, as análises dos experimentos realizados mostraram melhores resultados da regressão robusta em relação a regressão linear, para o âmbito do ensino fundamental do estado de Pernambuco. A regressão robusta obteve menor erro de predição, comprovada por tabelas e gráficos *boxplots*.

Portanto, a principal contribuição desse trabalho consiste em fornecer uma aplicação de análise correlacional e de regressores que minimiza o erro de predição e estima com maior precisão o relacionamento entre variáveis e indicadores educacionais, como a evasão e a reprovação escolar. Assim, utilizar a mineração de dados educacionais possibilita a identificação prévia de aspectos que podem precisar de melhorias e investimentos mais adequados, aprimorando aspectos do ensino e mitigando problemas. Além disso, melhor estimar as variáveis relacionadas a esse cenário traz um grande ganho a literatura e aos diversos interessados como estudantes, educadores, pesquisadores, governo, entre outros. São ferramentas que podem ser aplicadas de forma extensiva, gerando conhecimento, servindo como base para soluções de problemas e desenvolvimento de mecanismos em apoio ao ensino

Dessa forma, como trabalhos futuros, pretende-se refinar as técnicas de previsão, assim como explorar outros aspectos educacionais, expandir o escopo de estudo para outros níveis de escolaridade e para outros cenários como o regional ou nacional.

## Referências

BAKER, R. S. J., ISOTANI, S., CARVALHO, A. de: Mineração de dados educacionais: Oportunidades para o Brasil. **Revista Brasileira de Informática na Educação**, 2011, v. 12, n. 2, p. 3 – 13.

CALIXTO, K., SEGUNDO, C., and DE GUSMÃO, R. P.: Mineração de dados aplicada a educação: um estudo comparativo acerca das características que influenciam a evasão escolar. **In Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**, 2017, volume 28, page 1447.

CHAI, T., DRAXLER, R. R.: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. **Geoscientific Model Development**, 7, 1247-1250, 2014.

CHAPMAN, P. et al.: CRISP-DM 1.0 step-by-step data mine guide. **CRISP-DM Consortium**, 2000.

DA SILVA, J.M.C., IMRAN, H.: Um estudo sobre as variáveis para predição de alunos não concluintes em cursos suportados por Ambientes Virtuais de Ensino e Aprendizagem. **Revista Novas Tecnologias na Educação**, v. 13, n. 2, 2015.

DETONI, D., CECHINEL, C., ARAÚJO, R. M.: Modelagem e Predição de Reprovação de Acadêmicos de Cursos de Educação a Distância a partir da Contagem de Interações. **Revista Brasileira de Informática na Educação**, 2015, v. 23, n. 3.

INEP, Portal: Inep divulga dados inéditos sobre fluxo escolar na educação básica. 20 jun. 2017. **Notícias**. Disponível em: <<http://portalinep.gov.br/>>. Acesso em 14/02/2018.

LAISA, J., and NUNES, I.: Mineração de Dados Educacionais como apoio para a classificação de alunos do Ensino Médio. **In Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**, 2015, volume 26, page 1112.

MACHADO, R. D., NARA, E. O. B., SCHREIBER, J. N. C., and SCHWINGEL, G. A.: Estudo bibliométrico em mineração de dados e evasão escolar. **XI Congresso Nacional de Excelência em Gestão**, 2015.

MORO, S., LAUREANO, R., CORTEZ, P.: Using data mining for bank direct marketing: An Application of the crisp-dm methodology. EUROSIS. **Proceedings of European Simulation and Modelling Conference-ESM**, 2011, page 117–121.

ORTIZ, M. C., SARABIA, L. A. and, HERRERO, A.: Robust regression techniques: A useful alternative for the detection of outlier data in chemical analysis. **Talanta**, 2006, v. 70, n.3, pages 499-512.

RODRIGUES, R. L., DE MEDEIROS, F. P., and GOMES, A. S.: Modelo de regressão linear aplicado à previsão de desempenho de estudantes em ambiente de aprendizagem. **In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**, 2013, volume 24, page 607.