

## Análise de dados acadêmicos baseado em previsão, recomendação e visualização

Bruno Bastos Stoll<sup>1</sup>, Davidson Cury<sup>1</sup>, Orivaldo de Lira Tavares<sup>1</sup>, Crediné Silva de Menezes<sup>2</sup>

<sup>1</sup> Departamento de Informática – Universidade Federal do Espírito Santo (UFES)

<sup>2</sup> Programa de Pós-Graduação em Educação – Universidade Federal do Rio Grande do Sul (UFRGS).

brunobstoll@hotmail.com, {dede, tavares, credine}@inf.ufes.br

**Abstract.** *Technological support to assist teachers in the monitoring of students' needs becomes a perspective of improvement in the learning process. In virtual environments, records of student productions, activities, and interactions are collected automatically, enabling the complete analysis of learning. However, the volume of data may be very large. The purpose of this paper is to present a complete academic data analysis solution. To validate the work, the analysis of a set of public data was performed, performing the import, treatment, generation of predictive models, recommendations and panels for teachers. The proposed approach has good preliminary results.*

**Resumo.** *Apoio tecnológico para auxiliar professores no acompanhamento das necessidades de alunos torna-se uma perspectiva de melhoria no processo de aprendizagem. Nos ambientes virtuais, os registros das produções, atividades e interações dos alunos são coletados automaticamente, possibilitando a análise completa do aprendizado. Contudo, o volume de dados poderá ser muito grande. O objetivo deste trabalho é apresentar uma solução de análise de dados acadêmicos completa. Para validar o trabalho foi realizada a análise de um conjunto de dados públicos, realizando a importação, tratamento, geração de modelos preditivos, recomendações e painéis para professores. A abordagem proposta apresenta bons resultados preliminarmente.*

### 1. Introdução

A tecnologia vem promovendo uma grande transformação digital no mundo. O uso de tecnologias como a internet e armazenamento em larga escala permite que produções e interações sejam armazenadas. Essas informações armazenadas são fonte de conhecimento profundo sobre os usuários, suas preferências e necessidades. Isso não é diferente na educação. Existe uma área da pesquisa chamada de Análise do Aprendizado, do inglês *Learning Analytics* (LA), que é um campo no qual ferramentas de análise de dados são usadas e projetadas para melhorar o aprendizado e a educação (Elias, 2011).

Para professores, a percepção das individualidades e a intervenção no processo de aprendizagem de cada aluno se torna uma tarefa desafiadora. As interações dos alunos devem ser periodicamente analisadas, que por vezes, se torna uma tarefa

impossível, pois o volume de dados poderá ser muito grande e as turmas poderão ter muitos alunos. O tempo imposto pelas instituições também poderá ser outro problema. Por isso, entre a coleta de dados, a análise e a intervenção no aprendizado dos alunos, o professor poderá ter pouco tempo.

O objetivo deste trabalho é fornecer uma ferramenta completa de análise de dados acadêmicos, que seja possível prever o resultado de cada aluno, gerar recomendações e painéis de acompanhamento para professores. A solução proposta é flexível para analisar dados de qualquer ambiente de aprendizado, focada em predição, recomendação e visualização, com uma abordagem dinâmica ao contexto. O desenvolvimento deste trabalho foi realizado em dois estágios. No primeiro foi projetado e desenvolvido um framework de previsão de desempenho e geração de recomendações. No segundo estágio foi desenvolvido um sistema que faça a importação dos dados, use o framework proposto na etapa anterior e apresente painéis totalmente customizáveis para professores. Para validar a solução, foi usado um conjunto de dados públicos para uma análise completa dos dados. Na metodologia, foi utilizada a abordagem qualitativa, de natureza aplicada, com o objetivo descritivo e com procedimentos experimentais. As contribuições deste trabalho são: propor uma ferramenta completa de análise de quaisquer dados acadêmicos; gerar modelos preditivos para classificação de alunos em risco e seleção automática do algoritmo que tiver a melhor acurácia; estabelecer uma forma inovadora de recomendações, usando uma métrica derivada do sistema preditivo; e obter uma ferramenta que reduza o tempo entre a coleta e análise para apoio de tomada de decisões de professores.

O artigo foi organizado em 5 seções. Na seção 2 são apresentados os trabalhos correlatos. Na seção 3 é descrita a solução. Na seção 4 são apresentados os experimentos. E na seção 5 é apresentada a conclusão e os trabalhos futuros.

## 2. Trabalhos Correlatos

Existem várias pesquisas voltadas à análise de dados acadêmicos. O trabalho (Dyckhoff et al. 2012) apresenta detalhes teóricos sobre a implementação de uma ferramenta que permita professores explorarem e correlacionarem o uso de objetos de aprendizado e o comportamento de alunos com resultados de avaliações com base em indicadores gráficos. A pesquisa (Greller; Drachsler, 2012) propõe um framework genérico dando suporte a práticas educacionais e discute sobre as dimensões chave da análise do aprendizado no contexto educacional. O trabalho (Gotardo; Cereda; Junior, 2013) apresenta uma abordagem que usa algoritmos de aprendizado acoplados para integrar diferentes técnicas de aprendizado, mineração de dados e explorar um conjunto de dados educacionais oferecendo recomendações que possam indicar preferências do aluno que obtém melhores resultados. A pesquisa (Verbert et al., 2013) apresenta um framework conceitual que ajuda professores e alunos a analisar dados sobre o aprendizado. O trabalho (De Brito et al., 2014) propõe a previsão de desempenho dos alunos no primeiro período do curso de Ciência da Computação da UFPB, através das suas notas de ingresso no vestibular. O trabalho (Martin; Ndoye, 2016) faz uso das ferramentas *Tableau* e *Many Eyes* para analisar dados acadêmicos. A pesquisa (Padilha; Souza, 2017) apresenta o uso das ferramentas *Gephi* e *RapidMiner* para análise de interações entre alunos e professores. A pesquisa (Lanes; Alcântara, 2018) visa identificar estudantes que apresentam risco de evasão a partir do seu primeiro ano no curso de graduação.

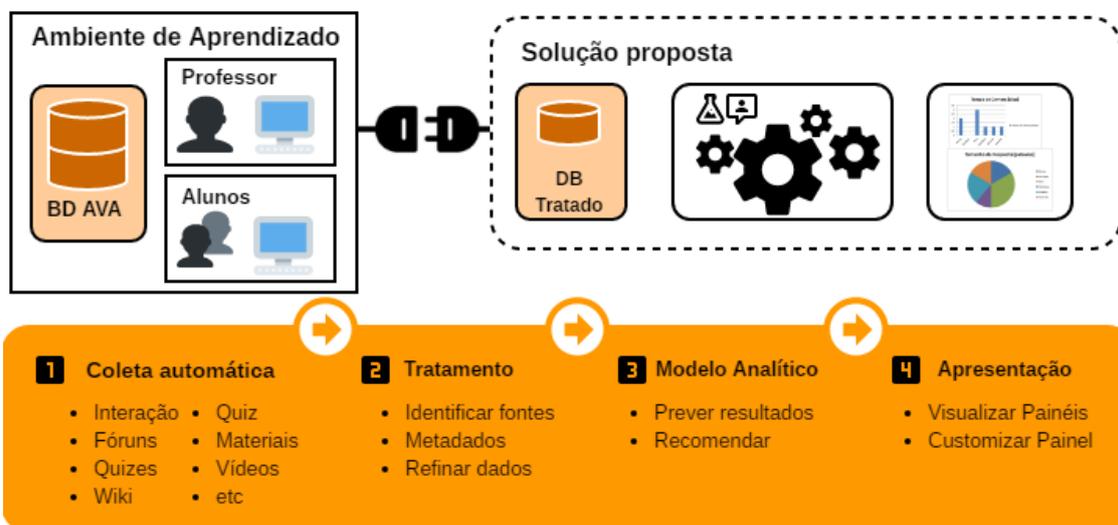
Os trabalhos descritos acima usam técnicas de predição ou descrição, recomendações e visualização, propondo diferentes soluções e abordagens. Em geral, as propostas na literatura são restritas e apresentam apenas experimentos ou um escopo fechado e reduzido. Conforme apresentado na Tabela 1, a maioria dos trabalhos não combina as diversas abordagens utilizadas neste trabalho, que são: Previsão de Desempenho (PD), Sistemas de Recomendações (SR) Sistemas de Visualização (SV) e acesso a Múltiplas Bases de dados (MB).

**Tabela 1 - Comparação das contribuições**

Proposta	PD	SR	SV	MB
Design and implementation of a learning analytics toolkit for teachers (Dyckhoff et al. 2012)	Sim	-	Sim	-
Translating learning into numbers: A generic framework for learning analytics (Greller; Drachslar, 2012)	Sim	Sim	Sim	-
Predição do Desempenho do Aluno usando Sistemas de Recomendação e Acoplamento de Classificadores (Gotardo; Cereda; Junior, 2013)	Sim	Sim	-	-
Learning analytics dashboard application (Verbert et al., 2013)	-	-	Sim	-
Predição de desempenho de alunos do primeiro período baseado nas notas de ingresso utilizando métodos de aprendizagem de máquina (De Brito et al., 2014)	Sim	-	-	Sim
Using learning analytics to assess student learning in online courses (Martin; Ndoye, 2016)	Sim	-	Sim	-
Uma Experiência do Uso de Ferramentas de Learning Analytics para Análise de Interações (Padilha; Souza, 2017)	Sim	-	Sim	Sim
Predição de Alunos com Risco de Evasão: estudo de caso usando mineração de dados (Lanes; Alcântara, 2018)	Sim	-	Sim	-

### 3. Descrição da Solução

Este trabalho consiste em uma ferramenta de análise de aprendizado acadêmico, que após ser devidamente configurada, será provedora de informações importantes aos alunos e professores. O sistema dá suporte às funcionalidades: (a) analisar qualquer dado acadêmico, ou seja, oferece a possibilidade de ser usado em diferentes situações; (b) usar técnicas de predição; (c) gerar recomendações; e (d) visualizar painéis totalmente customizáveis.



**Figura 1 - Arquitetura proposta**

Para uma experiência mais rica, o sistema está disponível para a plataforma web e para dispositivos móveis, ou seja, o usuário poderá acessar o sistema no navegador ou no celular, quando e onde quiser. Conforme apresentado na Figura 1, o sistema está dividido nos módulos: Coleta, Processamento, Transformação, Predição, Recomendação e Visualização. E os módulos são agrupados nas etapas da análise de dados: Coleta e Tratamento, Modelo Analítico e Apresentação.

### 3.1. Coleta e Tratamento

Na etapa de Carga e Tratamento ocorre a coleta, processamento e transformação dos dados. No módulo de Coleta é feita a carga inicial de dados, também pode ser necessário catalogar as fontes de dados e gerenciar os metadados. No módulo de Processamento é possível alterar configurações de campos e valores nas tabelas importadas, limpar e transformar valores contínuos em valores discretos.

### 3.2. Modelo Analítico

Na etapa de Análise de dados foi utilizado o framework projetado, que faz a previsão de desempenho e recomendações para os alunos classificados “em risco” (Stoll et al. 2018). O módulo de predição faz uso de algoritmos de aprendizado de máquina para previsão de desempenho dos alunos, podendo classificá-los em “sucesso” ou “em risco” acadêmico. A classificação “sucesso” acadêmico ocorre quando o aluno atende aos propósitos educacionais de aprender e tirar notas boas. Já o resultado “em risco” ocorre quando o aluno dá sinais que não está aprendendo, com probabilidade de não tirar notas boas, ou até mesmo desistir (evadir) de uma determinada disciplina ou do curso. No módulo de recomendações, as mensagens de recomendações são sugeridas aos professores, que deverão julgar sua pertinência antes do envio.

Para um problema específico, vários algoritmos podem ser apropriados, no entanto, um algoritmo pode ser mais adequado do que outros, mas nem sempre é possível saber com antecedência qual é a melhor opção (Pakalra; Olprod, 2018). Como o sistema está preparado para ler quaisquer dados acadêmicos, foi utilizada a seleção automática do algoritmo que tiver melhor acurácia, dentre as opções: Árvore de Decisão; K Vizinhos Próximos, do inglês *K-Nearest Neighbors* (KNN); Máquina de Vetores de Suporte, do inglês *Support Vector Machine* (SVM); Perceptron em Multicamadas, em inglês *Multi-Layer Perceptron* (MLP); e Naive Bayes (Faceli et al. 2011). Esses algoritmos foram selecionados, pois cada um deles tem uma abordagem diferente.

São usados os dados históricos de turmas anteriores para fazer o treinamento e testes dos algoritmos preditivos. Foi utilizada a técnica de validação cruzada para avaliar a acurácia dos algoritmos. Onde é feita a divisão dos dados já rotulados para treinamento e teste para verificar a quantidade de acertos (Zaki et al. 2014). O algoritmo selecionado, que teve a melhor acurácia, classificará instâncias ainda não rotuladas, ou seja, realizará a previsão de desempenho dos alunos.

Existem as categorias de intervenções automáticas e a semiautomática. Na intervenção automática o sistema realiza ações pré-determinadas automaticamente com pouca ou nenhuma participação. Já as intervenções semiautomáticas são alertas e outros indicadores relevantes dos padrões de aprendizagem enviados para os envolvidos no

processo de ensino (Brown, 2012). A intervenção semiautomática foi escolhida para que o professor tenha a oportunidade de visualizar e alterar as mensagens antes de enviá-las ou simplesmente as ignorar caso não sejam apropriadas.

Para cada aluno classificado “em risco” o sistema deverá sugerir uma mensagem ao professor, que deverá julgar sua pertinência antes do envio, podendo editá-las ou ignorá-las. Para a geração das mensagens o sistema utilizará um método evolucionário, no qual (re)submete instâncias, com parâmetros alterados, até que a previsão de desempenho do aluno seja igual a “sucesso”. Ou seja, para cada aluno classificado “em risco”, o sistema alterará cinco características/comportamento e (re)submete a instância ao modelo preditivo selecionado. Caso o aluno ainda não tenha sido classificado como “sucesso”, o sistema adicionará mais características e o submete ao modelo até que ele seja classificado como “sucesso”. As características alteradas vão compor a mensagem, que correspondem aos dados coletados do ambiente de aprendizado.

### 3.3. Apresentação

O sistema dispõe de um mecanismo de visualização de dados, onde o professor poderá visualizar graficamente informações sobre o andamento da turma. É possível configurar várias Visões e agrupá-las em Painéis. Os tipos de visões são: Gráfico de Pizza; Gráfico de barras; Gráfico de Linha; Listagem em grade (*Grid*); e informações em cartões (*Cards*). O sistema filtra os dados, para que o professor só visualize os dados de suas turmas.

### 3.4. Arquitetura Tecnológica

Os recursos tecnológicos utilizados foram: linguagem de programação Python <sup>1</sup> para desenvolver o sistema; Bottle <sup>2</sup> como framework Web; frameworks Javascript JQuery <sup>3</sup> e Bootstrap <sup>4</sup>; SQLAlchemy <sup>5</sup> para mapeamento das entidades, um framework de persistência de banco de dados; framework scikit-learn <sup>6</sup> para aprendizado de máquinas, com o uso dos classificadores Decision Tree, KNN, SVM, MLP e Naive Bayes; as bibliotecas Numpy <sup>7</sup> e o Pandas <sup>8</sup> para manipulação dos dados; o banco de dados SQLite <sup>9</sup>; e o sistema foi hospedado no PythonAnywhere <sup>10</sup>, devido a facilidade de instalação de bibliotecas Python.

## 4. Experimentos

Para validar as funcionalidades do sistema, foi usado um conjunto de dados públicos chamado OULA, *Open University Learning Analytics*, disponível no site *UCI Machine Learning Repository* <sup>11</sup>. Esses dados foram disponibilizados para apoiar pesquisas

---

<sup>1</sup> <https://www.python.org/>

<sup>2</sup> <https://bottlepy.org/>

<sup>3</sup> <https://jquery.com/>

<sup>4</sup> <https://getbootstrap.com/>

<sup>5</sup> <https://www.sqlalchemy.org/>

<sup>6</sup> <https://scikit-learn.org/>

<sup>7</sup> <http://www.numpy.org/>

<sup>8</sup> <https://pandas.pydata.org/>

<sup>9</sup> <https://www.sqlite.org/>

<sup>10</sup> <https://www.pythonanywhere.com/>

<sup>11</sup> <https://archive.ics.uci.edu/ml/>

relacionadas à análise de dados educacionais, contendo informações de cursos apresentados na Universidade Aberta (*Open University* - UO). O conjunto de dados contém as informações sobre 22 cursos, 32.593 alunos com seus resultados de avaliação e registros de suas interações com o AVA representados por resumos diários de cliques de alunos (10.655.280 entradas) (Kuzilek et al. 2017). Com esses dados foi realizada a validação do sistema. Esse processo de validação consiste: na importação dos dados; no tratamentos dos metadados; no processamentos e transformações; na geração de modelos preditivos; previstos os resultados dos alunos; na geração das recomendações para os alunos classificados “em risco”; e realizada a parametrização de painéis para visualização dos dados.

Na Tabela 2 são descritas as informações dos arquivos disponíveis no conjunto de dados, que foram importados no sistema. O metadado “classe” do atributo *final\_result* foi alterado para “sim”, ou seja, quando não estiver preenchido será previsto pelo sistema. Foi apagado o resultado de cinco alunos aleatoriamente de cada um dos sete módulos, gerando um total de 35 alunos para compor as turmas do professor, no caso de teste. Foi usada uma base de nomes para atribuir aos 35 alunos para melhor o cenário de testes. Para cada um dos sete módulos atribuímos uma descrição para auxiliar na montagem dos relatórios. Foram discretizados os dados das interações dos alunos no ambiente (*click\_homepage*, *click\_glossary*, *click\_quiz*, *click\_ouwiki*, etc). Os valores desses atributos foram transformados de contínuos em valores categóricos. E foram categorizados conforme os quartis, com as descrições “pouco uso”, “uso moderado” e “muito uso”, onde: “pouco uso” corresponde ao primeiro quartil; “uso moderado” corresponde ao segundo e o terceiro quartil; e “muito uso” corresponde ao quarto quartil. Desta forma, as mensagens de recomendações ficam de fácil entendimento para os alunos. Por fim, alteramos a configuração, para que esses valores sejam apresentados em colunas.

**Tabela 2 - Descrição dos dados OULA**

Tabela	Descrição
assessments.csv	informações sobre avaliações dos módulos
courses.csv	lista de todas as disciplinas disponíveis
studentAssessment.csv	resultados das avaliações dos alunos
studentInfo.csv	informações demográficas sobre os alunos, juntamente com os seus resultados
studentRegistration.csv	registro do aluno no módulo
studentVle.csv	informações sobre as interações de cada aluno com os materiais no AVA
vle.csv	Informações sobre os materiais disponíveis no AVA

No módulo de Predição foram gerados modelos preditivos, esse processo consiste no treinamento e teste dos algoritmos de aprendizado de máquina. Conforme apresentado na Figura 2, são apresentadas as medidas: Acurácia (**Acur**), Sensibilidade (**Sen**), Especificidade (**Esp**), Eficiência (**Efic**), Valor Preditivo Positivo (**VPP**) e Valor Preditivo Negativo (**VPN**). Nesse experimento, foi realizada a separação de 70% dos dados para treinamento e 30% para testes. O algoritmo que teve a melhor acurácia foi a MLP, com um valor de 75,9521. Após a seleção automática do algoritmo, foram realizadas as previsões das instâncias não rotuladas, ou seja, a previsão do resultado de cada aluno, classificando-os em “em risco” ou “sucesso” acadêmico.

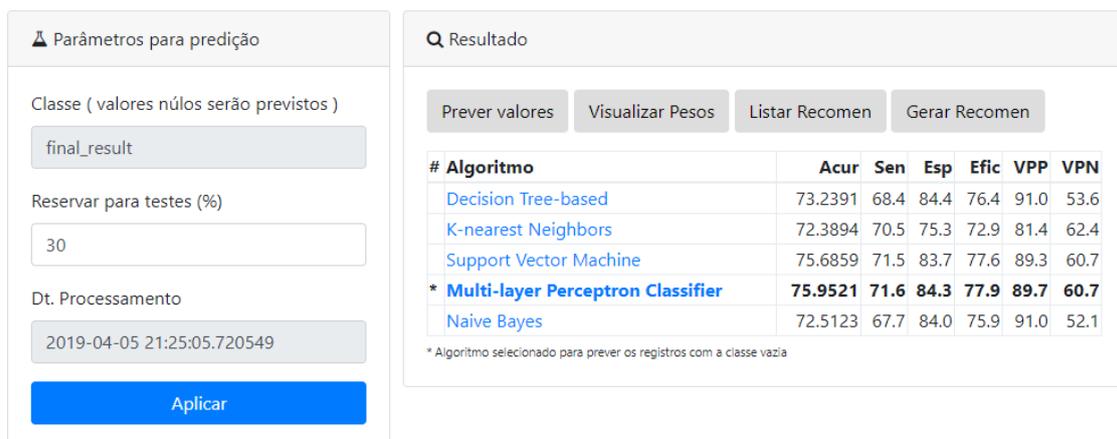


Figura 2. Medidas preditivas dos algoritmos

A base das mensagens de recomendações são as importâncias/peso das características processadas pelo módulo de Predição. Conforme apresentado na Figura 3, a característica “Acessos ao portal AVA”, com uma importância de 0,51518. A maior porcentagem de alunos com “sucesso” acadêmico teve “muito acesso” nessa característica. Com um total de 93,53% para alunos com sucesso acadêmico e consecutivamente 6,47%, com insucesso acadêmico. Todas as características são apresentadas com diferentes importâncias e porcentagens.

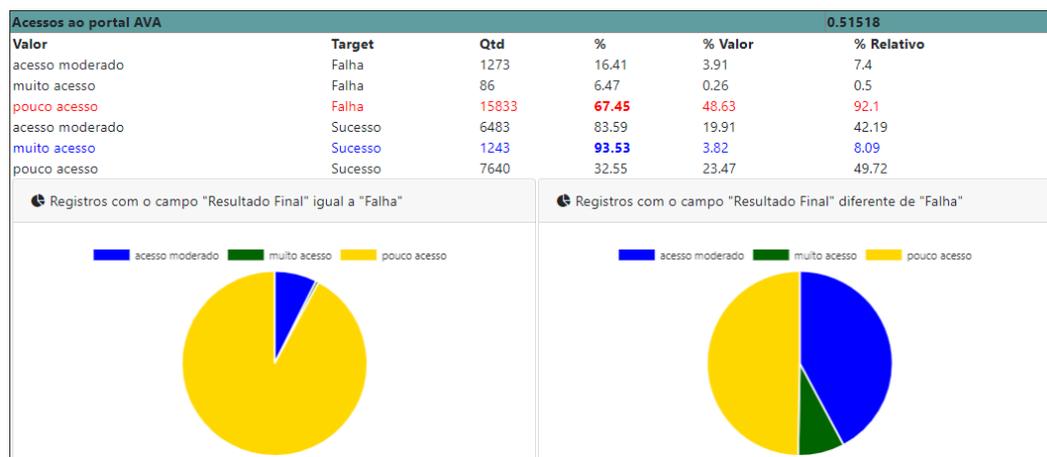


Figura 3. Peso da característica acesso ao portal acadêmico

No módulo de recomendação, foram geradas as mensagens para os alunos classificados “em risco”. Conforme apresentada na Figura 4, bloco A, foi apresentada a mensagem candidata sugerida ao professor. O professor pode editar cada item da mensagem ou até mesmo ignorar qualquer um que não considere apropriado. Na Figura 4, bloco B, no qual a tela foi capturada com o auxílio de software de conexão remota, é apresentada a mensagem recebida pelo aluno com as recomendações do professor.

**(A) Professor**

Caro aluno, **Rebecca Arnold**. Você foi classificado "em risco" por um sistema de previsão com uma taxa de acerto de 70.2027%

Ignorar

Na característica "Acessos ao portal AVA" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Ignorar

Na característica "Acesso a quiz" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Ignorar

Na característica "Acesso aos materiais" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Ignorar

Na característica "Acesso ao wiki" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Ignorar

Na característica "Acesso a sub-páginas" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Enviar recomendação

**(B) Aluno**

LA :: Estudante

Caro aluno, **Rebecca Arnold**. Você foi classificado "em risco" por um sistema de previsão com uma taxa de acerto de 70.2027%

Na característica "Acessos ao portal AVA" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

Na característica "Acesso a quiz" esta informado: "pouco acesso".  
Recomendamos que seja modificado para "muito acesso"

**Figura 4. Mensagem: (A) sugerida ao professor; (B) recebida pelo aluno**

Conforme a Figura 5, Após configurar uma Visão é possível exibir seus resultados. O professor tem acesso a um Painel com dados de sua turma, permitindo analisar as Visões de forma interativa, conforme apresentado na Figura 6. Esse painel foi parametrizado para que o professor visualize as informações: (a) quantidade de aluno por desempenho previsto das turmas; (b) previsão de resultado por disciplina; (c) quantidade de acessos ao portal; (d) quantidade de acessos aos quizzes; (e) quantidade de acessos aos materiais; (f) previsão de desempenho por disciplina; e (g) lista da previsão de desempenho por aluno. Ele também pode incluir, alterar ou remover qualquer visualização em seu painel.

**(A) Configurar**

Nome

[B] Resultado por Disciplina

Tipo

Gráfico Barras

Campo	Agrupador	Gráfico
Nome	Quantidade	Valor
descr_module		Categoria
final_result		Serie

Salvar

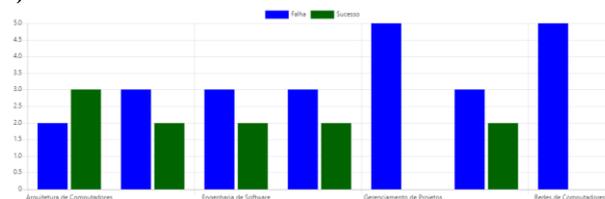
**(B) Visualizar****Figura 5. Visão: (A) Configuração; (B) Exibição**



Figura 6. Painel do professor

## 5. Conclusão

Durante os últimos anos houve um aumento nas pesquisas voltadas à análise de dados acadêmicos. Principalmente porque os dados são coletados automaticamente nesses ambientes. Neste artigo, foi apresentada uma proposta de sistema de análise de dados completa. Cujos resultados iniciais demonstram um potencial para professores visualizarem graficamente os dados de suas turmas, entenderem fatores que comprometem seus resultados, preverem resultados dos alunos permitindo-os atuarem de forma proativa em relação aos resultados e obterem sugestões de mensagens de recomendações. Portanto, conclui-se que uma análise de dados, quando realizada de forma integrada, é mais eficiente. Os impactos advindos de sua utilização são que professores tenham maior eficiência em ambientes de aprendizagem na modalidade online ou híbridos. Como trabalhos futuros pretende-se melhorar a eficiência dos

algoritmos através de técnicas de pré-processamento de dados em bases desbalanceadas e adicionar um módulo de análise de redes sociais, do inglês *Social Network Analysis* (SNA).

## Referências

- Brown, J. S. (2000). Growing up: Digital: How the web changes work, education, and the ways people learn. *Change: The Magazine of Higher Learning*, 32(2), 11-20.
- De Brito, D. M., de Almeida Júnior, I. A., Queiroga, E. V., & do Rêgo, T. G. (2014). Predição de desempenho de alunos do primeiro período baseado nas notas de ingresso utilizando métodos de aprendizagem de máquina. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 25, No. 1, p. 882).
- Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Journal of Educational Technology & Society*, 15(3), 58-76.
- Elias, T. (2011). Learning analytics. *Learning*, 1-22.
- Faceli, K., Lorena, A. C., Gama, J., & Carvalho, A. C. P. D. L. (2011). Inteligência Artificial: Uma abordagem de aprendizado de máquina.
- Gotardo, R., Cereda, P. R. M., & Junior, E. R. H. (2013). Predição do desempenho do aluno usando sistemas de recomendação e acoplamento de classificadores. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 24, No. 1, p. 657).
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics.
- Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific data*, 4, 170171.
- Lanes, M., & Alcântara, C. (2018, October). Predição de Alunos com Risco de Evasão: estudo de caso usando mineração de dados. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 29, No. 1, p. 1921).
- Martin, F., & Ndoye, A. (2016). Using learning analytics to assess student learning in online courses. *Journal of University Teaching & Learning Practice*, 13(3), 7.
- Padilha, T., & Souza, I. (2017, October). Uma experiência do uso de ferramentas de learning analytics para análise de interações. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação* (Vol. 6, No. 1, p. 644).
- Pakalra, Olprod, Página de dicas úteis do algoritmo para Machine Learning para o Microsoft, Azure Machine Learning Studio, <<https://docs.microsoft.com/pt-br/azure/machine-learning/studio/algorithm-cheat-sheet>>, acessado em agosto 2018.
- Stoll, B. B., Cury, D., & de Menezes, C. S. (2018). Framework para predições e recomendações em dados acadêmicos. *RENOTE*, 16(2).
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509.
- Zaki, M. J., Meira Jr, W., & Meira, W. (2014). *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press.