

Sistema de avaliação baseado em técnicas de aprendizado de máquina

Evaluation system based on machine learning techniques

Lorraine Nayara Porto Siqueira

Graduanda do curso de Sistemas de Informação (UNIPAM).
e-mail: lorrannenpsiqueira@gmail.com

José dos Reis Mota

Professor orientador (UNIPAM). e-mail: josereis@unipam.edu.br

Resumo: Este artigo descreve o desenvolvimento de um sistema de avaliação web e a utilização de técnicas de aprendizado de máquina empregando análise preditiva para validação dos resultados obtidos nas avaliações. O objetivo de desenvolvimento desse sistema é proporcionar dados para que seja feito o acompanhamento do desenvolvimento dos alunos e, assim, para que os envolvidos no processo educativo tenham condições de propor intervenções pedagógicas e, como consequência, buscar resultados ainda melhores para a instituição e para os alunos, tanto no Exame Nacional de Desempenho dos Estudantes (ENADE), como no decorrer do seu percurso acadêmico. Para o desenvolvimento do trabalho, foram utilizadas tecnologias como o SCRUM e a linguagem de programação PHP com a utilização do modelo MVC (*Model, View, Controller*), banco de dados MySQL, para processo de desenvolvimento; e o ambiente Microsoft Azure Machine Learning Studio para o aprendizado de máquina.

Palavras-chave: Aprendizado de máquina. Sistema de avaliação. Inteligência artificial.

Abstract: This article describes the development of a web evaluation system and the use of machine learning techniques using predictive analysis to validate the results obtained in the evaluations. The goal of this system is to provide data to monitor the development of the students, so that those involved in the educational process are able to propose pedagogical interventions and as a consequence seek even better results for the institution and for the students, both in the National Student Performance Exam (ENADE) and in the course of their academic career. For the development of the work we used technologies such: SCRUM; PHP programming language using the MVC (*Model, View, Controller*) model, MySQL database for the development process; and the Microsoft Azure Machine Learning Studio, for machine learning.

Keywords: Machine learning. Evaluation system. Artificial Intelligence.

1. Introdução

O processo de tomada de decisões dentro de uma organização pode ser bastante difícil, já que a busca por uma melhor solução para uma oportunidade ou problema pode gerar consequências positivas ou negativas. A utilização de um sistema de apoio à tomada dessas decisões apresenta grande valia, pois o mesmo, por meio do cruzamento de informações, fornece apoio e aponta a melhor decisão a ser tomada.

Entretanto, esse é um modelo que pode ir além da área de negócios. Em projetos educacionais, um sistema como esse pode gerar indicadores que auxiliam instituições de ensino, professores, tutores e alunos, contribuindo, assim, para que decisões sejam baseadas em dados, evitando riscos e falhas que possam gerar impactos negativos.

Um ambiente de aprendizagem é um espaço em que um indivíduo está sujeito a oportunidades de aprendizagem. Entretanto, isso não precisa estar diretamente ligado a um ambiente físico. O avanço da tecnologia, o alcance da internet e, principalmente, a mobilidade, possibilitam que as atividades transcendam os limites de uma sala de aula.

Buscando efetivar essas oportunidades de aprendizagem, assim como obter informações para auxiliar decisões estratégicas, este trabalho tem como objetivo modelar e desenvolver um sistema para avaliação de alunos, assim como utilizar técnicas de aprendizado de máquina, empregando análise preditiva para validação dos resultados obtidos. O emprego de aprendizado de máquina possibilita a análise do possível resultado final do aluno em disciplinas, evitando reprovações e seus problemas correlatos, como o aumento do índice de evasão. Além disso, o sistema de avaliação trabalha com conteúdos avaliados pelo Exame Nacional de Desempenho dos Estudantes (ENADE), utilizado como instrumento governamental para avaliar e qualificar as instituições de ensino.

O sistema foi desenvolvido no Centro Universitário de Patos de Minas, buscando auxiliar o acompanhamento do desempenho do aluno, de forma que os envolvidos no processo educativo possam propor intervenções, buscando aperfeiçoá-lo. Como consequência, espera-se melhorar ainda mais os resultados da instituição no ENADE.

Ao estabelecer um prognóstico do resultado de alunos com base em dados já existentes, é possível antecipar ações e intervenções pedagógicas de forma que a instituição possa estabelecer metas para melhorar um determinado resultado, ou de forma que o próprio aluno possa antecipar ações, visando o alcance dos objetivos acadêmicos.

2. Revisão de literatura

Nesta seção, são apresentados conceitos referentes a aprendizado de máquina e avaliações no contexto educacional, assim como sobre as tecnologias utilizadas para conduzir o trabalho.

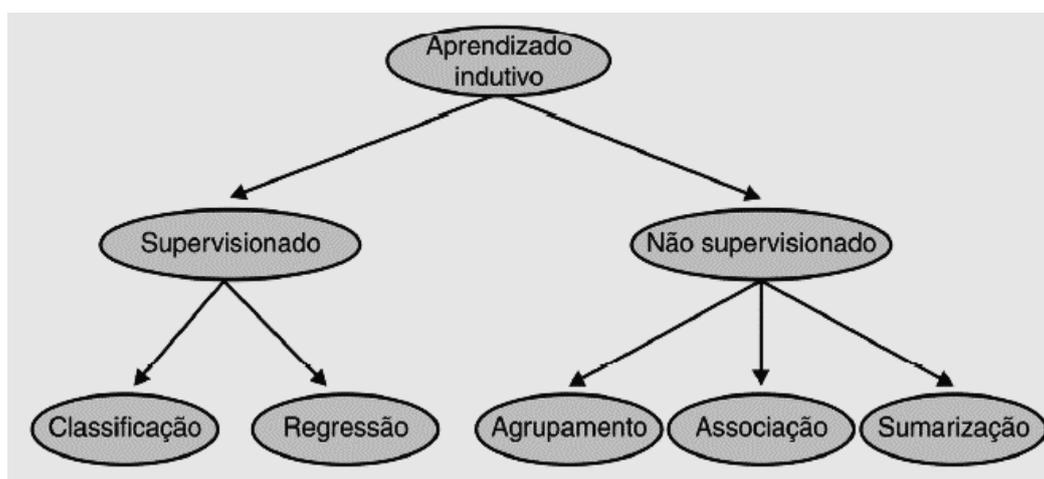
2.1. Aprendizado de máquina

A área de aprendizado de máquina (AM) é considerada um ramo da Inteligência Artificial, cujo objetivo é construir sistemas que sejam capazes de aprender de forma automatizada a partir de dados (BRINK; RICHARDS, 2014). Algoritmos de AM aprendem a induzir uma função ou hipótese capaz de resolver um problema a partir desses dados que representam instâncias do problema a ser resolvido (CARVALHO, 2011).

Segundo Carvalho (2011), algoritmos de AM são utilizados em diversas tarefas, que podem ser organizadas de acordo com diferentes critérios. Essas tarefas de aprendizado podem ser divididas em descritivas e preditivas.

Nas tarefas descritivas, a meta é explorar ou descrever um conjunto de dados. Já nas tarefas preditivas, a meta é encontrar uma função a partir dos dados de treinamento que possa ser usada para prever valores desconhecidos ou futuros de outras variáveis. Os tipos de tarefas de aprendizado podem ser visualizados na Figura 1.

Figura 1. Hierarquia de Aprendizado



Fonte: CARVALHO, 2011, p. 6

No topo da hierarquia, temos o aprendizado indutivo, em que são realizadas generalizações a partir dos dados. Em seguida, os tipos de aprendizado supervisionado e não supervisionado que correspondem, respectivamente, aos tipos preditivo e descritivo. As tarefas supervisionadas se diferenciam pelo tipo dos rótulos de dados: no caso de classificação, ele é discreto e, no caso de regressão, ele é contínuo.

Já as tarefas não supervisionadas são divididas em agrupamento, associação e sumarização. No agrupamento, os dados são agrupados de acordo com a similaridade; na associação, o objetivo é encontrar padrões frequentes de associações entre os atributos de um conjunto de dados; já a sumarização consiste em encontrar uma descrição simples e compacta para um conjunto de dados.

No desenvolvimento deste projeto, o foco foi nas tarefas de aprendizado indutivo supervisionado, ou seja, a partir de dados já existentes, foram obtidos modelos

preditivos em relação ao aproveitamento acadêmico dos discentes.

2.2. Avaliação da aprendizagem

Nos últimos anos, segundo Cervi (IESDE, 2001), o conceito de avaliação da aprendizagem tem evoluído muito e sensivelmente. Se antes ela era vista como uma quantificação no processo de ensino e como aquilo que propositadamente o aluno tenha aprendido dentro do contexto escolar, hoje ela está mais estabelecida com complexidade e amplitude.

No trabalho docente, a avaliação é tarefa necessária para que professores e instituição possam acompanhar o processo de ensino. É a partir dessa avaliação que o professor consegue comparar os objetivos propostos e, dessa forma, constatar progressos e dificuldades. São os resultados que os auxiliam a tomar decisões sobre o seu trabalho.

2.3. Tecnologias na educação

Uma grande evolução da humanidade é a facilidade com que o conhecimento pode chegar em todos os lugares com apenas alguns cliques. Essa modalidade de educação é efetivada a partir do intenso uso de tecnologias de informação e comunicação, podendo ou não apresentar momentos presenciais (MORAN, 2009).

De acordo com Paulo Freire (2003, p. 21), “ensinar não é transferir conhecimento, mas criar as possibilidades para a sua produção ou sua construção”, e isso pode ocorrer em qualquer ambiente, seja ele presencial ou virtual. Diante das influências tecnológicas de informação e comunicação, a mobilidade tecnológica vem contribuindo para um novo modelo de educação, e as instituições de ensino vêm se adaptando a esse novo modelo, adotando esses sistemas.

2.4. Scrum

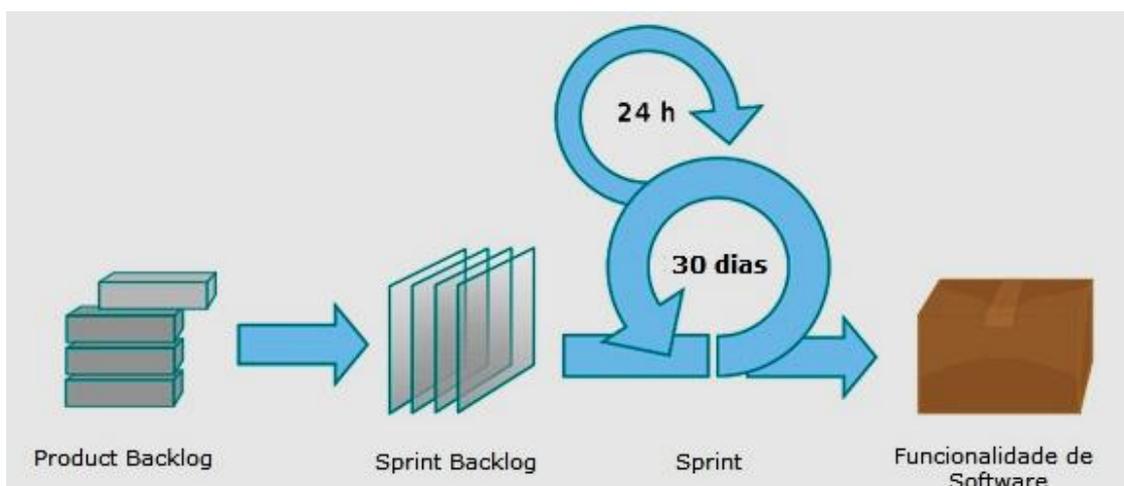
O framework Scrum é uma metodologia ágil para gestão e desenvolvimento de projetos. É uma ferramenta capaz de controlar de forma eficaz e eficiente o trabalho, unindo equipes e cliente em função de um objetivo em comum (PRESSMAN, 2006).

É importante saber que o Scrum trabalha com grupos altamente gerenciáveis, que trabalham nos seguintes papéis (KNINBERG, 2007):

- *Product Owner*: é o que define os requisitos do produto e prioriza os itens do *Product Backlog*, defende os interesses do cliente e motiva a equipe.
- *Scrum Master*: garante que o time esteja produtivo e funcional, ou seja, é responsável por eliminar impedimentos e colabora com o *Product Owner* na priorização de requisitos.
- *Time Scrum*: normalmente um time formado por cinco a nove pessoas, que são responsáveis por estimar tempo e valores para tarefas de desenvolvimento.

Na Figura 2, é apresentado o ciclo de vida de um desenvolvimento de software utilizando-se o Scrum.

Figura 2. Ciclo de vida da metodologia Scrum



Fonte: DevMedia - Desenvolvimento ágil com Scrum: uma visão geral, 2012

Inicia-se o ciclo com o chamado *Product Backlog*, um registro que contém as funcionalidades a serem desenvolvidas. Através do *Product Backlog* são criados ciclos que são chamados de *Sprints*. No início de cada *Sprint*, é realizado um *Sprint Planning Meeting*, uma reunião de planejamento em que o *Product Owner* prioriza itens do *Product Backlog* e a equipe seleciona as atividades que serão capazes de serem realizadas dentro da *Sprint* que se inicia. Essas tarefas que são definidas na *Sprint* são movidas de *Product Backlog* para *Sprint Backlog*.

O *Daily Scrum* é uma breve reunião feita a cada dia de uma *Sprint* que acontece normalmente na primeira parte do dia, em que o objetivo é informar sobre o que foi feito no dia anterior, identificar o que será feito no dia e resolver possíveis impedimentos.

Ao final de uma *Sprint*, as funcionalidades implementadas são apresentadas pela equipe em uma *Sprint Review Meeting*. É nessa reunião que o projeto é avaliado em relação aos objetivos da *Sprint*. Finalmente, é realizada uma *Sprint Retrospective* que serve para identificar o que funcionou bem, o que pode ser melhorado e como pode ser melhorado. Assim, o ciclo se reinicia (SCHWABER; SUTHERLAND, 2013).

3. Metodologia

Inicialmente, foi feito um estudo bibliográfico sobre aprendizado de máquina, área da Inteligência Artificial, que foi utilizado no trabalho para a obtenção de resultados. Também foram estudadas tecnologias utilizadas no desenvolvimento do projeto, levando em consideração recursos, disponibilidade e custos para a adoção do sistema.

O sistema de avaliação baseado em técnicas de aprendizado de máquina foi desenvolvido utilizando-se o Scrum. A cada *Sprint*, novas funcionalidades foram desenvolvidas, testadas e acopladas ao sistema.

O Quadro 1 lista ferramentas e tecnologias utilizadas no desenvolvimento do sistema.

Quadro 1. Ferramentas e tecnologias utilizadas

Nome	Descrição
Sublime Text	IDE de desenvolvimento que dá suporte a linguagem PHP
Visual Paradigm for UML	Software utilizado para desenvolver os diagramas UML
MockupBuilder	Software para modelagem de protótipos
Navicat Premium	Ferramenta para gerenciar banco de dados MySQL.
Microsoft Office Word	Editor de texto utilizado para gerar a documentação do projeto
Microsoft Azure Machine Learning Studio	Ambiente disponibilizado pela Microsoft para criar experiências de aprendizado de máquina

Fonte: Dados do trabalho

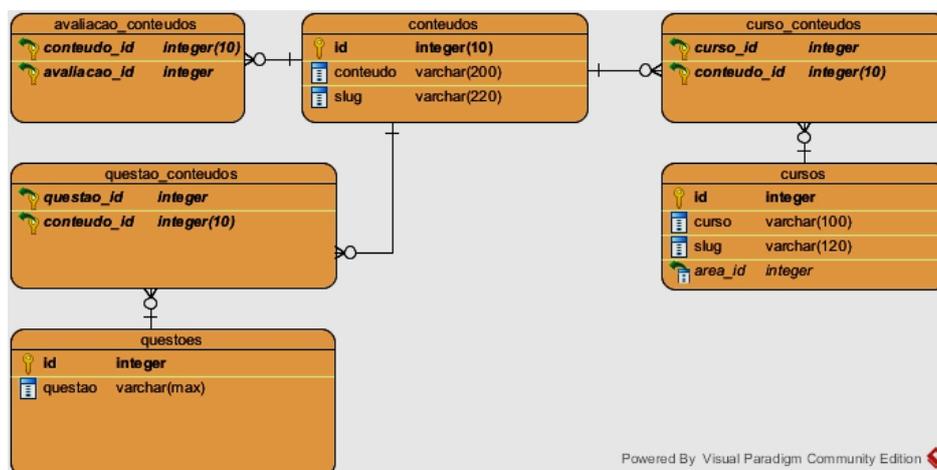
Na fase de estruturação do projeto, foram utilizadas as ferramentas *Microsoft Office Word* e *Visual Paradigm for UML* para documentação dos requisitos e para a construção dos diagramas de caso de uso: diagrama entidade relacionamento e diagrama de classes. Na fase de desenvolvimento, foi utilizado o *Sublime Text* para codificação do sistema e o *Navicat Premium* para gerenciar o banco de dados, em que os dados são armazenados. Para criar experiências com os dados e empregar as técnicas de aprendizado de máquina, foi utilizada a plataforma oferecida pela *Microsoft*, *Azure Machine Learning Studio*.

4. Desenvolvimento e resultados

Inicialmente, foi necessária a realização de uma pesquisa para entender o formato do Exame Nacional de Desempenho dos Estudantes (ENADE), a fim de colher todos os requisitos que seriam necessários para o desenvolvimento do sistema, uma vez que um dos objetivos do sistema é simular o exame de forma automatizada.

Para descrever o modelo de dados, foi elaborado o Diagrama Entidade Relacionamento (DER), representado parcialmente na Figura 3, que é composto de tabelas ou entidades, seus atributos e relacionamentos.

Figura 3. Diagrama Entidade Relacionamento (DER)



Fonte: Dados do trabalho

Entre as entidades, existem aquelas que são importantes para o funcionamento do sistema, como a entidade que mantém informações dos usuários (*users*), pois é a partir dela que será feito o controle de acesso ao sistema, já que o mesmo é restrito apenas a usuários autorizados. Existem também aquelas que são a chave para o aprendizado de máquina, como a entidade que salva o resultado dos alunos, pois a partir dos resultados dos alunos é que são gerados resultados para a instituição.

Outra entidade chave para o sistema é a entidade de conteúdos. A fim de obter um resultado efetivo, o conteúdo é relacionado com o curso, com a avaliação e com as questões. Ao criar uma avaliação, será necessário informar quais conteúdos serão avaliados e, dessa forma, o sistema buscará questões que foram relacionadas a esse conteúdo. A avaliação poderá ser aplicada para todos os alunos dos cursos que também possuem esse conteúdo associado.

Quando o aluno entra no sistema, é apresentado para ele um histórico das avaliações que foram realizadas, as avaliações que estão disponíveis e as futuras avaliações. Após a realização da avaliação, o aluno pode analisar seu resultado, conforme mostra a Figura 4.

Figura 4. Tela de resultados da avaliação



Fonte: Dados do trabalho.

Com a participação dos alunos, a instituição pode fazer análises com os dados, filtrando por aluno, por curso e por conteúdo, e trabalhar dentro desses resultados para tomar decisões e, possivelmente, adquirir melhoria nos processos de aprendizagem da instituição e contribuir para o aprendizado dos próprios alunos.

Além disso, foi modelado um *DataWarehouse* para coleta e análise dos dados de avaliação dos alunos em etapas anteriores, de forma que esses dados possam ser usados como treinamento para os algoritmos de aprendizado de máquina e comparação de resultados. A Figura 5 representa a tabela fato criada para efetuar essa análise.

Conforme se pode observar, várias dimensões podem ser utilizadas para a análise, como a etapa de avaliação que, no caso do UNIPAM, são: a) etapa de pontos distribuídos pelo professor (40,0 pontos); b) etapa de avaliação colegiada (20,0 pontos); c) etapa de avaliação integradora (20,0 pontos); d) etapa de projeto integrador (20,0 pontos). Com os dados históricos, pode-se fazer uma análise preditiva dos resultados dos alunos e compará-los aos resultados reais ao aplicar a avaliação gerada pelo sistema. As variáveis podem ser comparadas e ajustadas para verificar, por exemplo, qual a etapa de avaliação traduz melhor o resultado do aluno. Esse modelo é representado pela Figura 5.

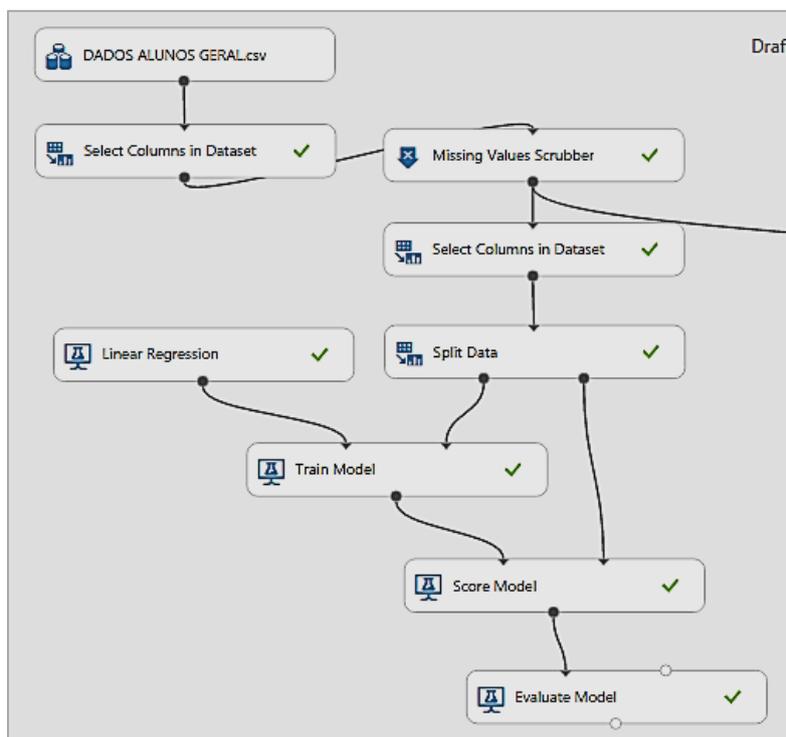
Figura 5. Diagrama Entidade Relacionamento



Fonte: Dados do trabalho.

Depois de realizar a coleta e a análise dos dados dos alunos em períodos anteriores, foram criados dois modelos utilizando regressão linear para prever a nota média final dos alunos, empregando algumas variáveis diferentes. No primeiro modelo, foram utilizadas as variáveis: tipo de bolsa, curso, turno, média final, média da nota distribuída por professor, nota do Projeto Integrador e nota da Avaliação Integradora. No segundo modelo, foram adicionadas todas as variáveis do primeiro modelo e acrescentadas as seguintes: quantidade de disciplinas aprovadas e reprovadas; quantidade de faltas e aulas; e percentual de falta e presença. A figura 6 traz a representação do primeiro modelo.

Figura 6. Modelo de regressão linear



Fonte: Dados do trabalho.

Cada etapa representada na Figura 6 é chamada de módulo. O ciclo se inicia na escolha de um *Dataset*, que corresponde aos dados extraídos. Pode-se escolher com quais variáveis o experimento irá trabalhar, utilizando-se o módulo de *Select Columns in Dataset*. O módulo de *Split Data* tem a função de dividir os dados. Nessa etapa, é possível escolher qual a porcentagem que será dedicada ao treinamento, e o restante será aplicado em teste.

Nos dois modelos que foram criados, 70% dos dados foram dedicados ao treinamento e 30% dedicado aos testes. No módulo *Linear Regression*, é utilizado um algoritmo de Regressão Linear. Regressão é um dos tipos de técnicas de aprendizado de máquina supervisionado e é utilizada para prever um número, por exemplo. No *Train Model*, define-se qual variável será prevista e, nesse caso, a variável é média final. O *Score Model* é o responsável por gerar as previsões usando um treinamento de classificação ou regressão. Para analisar e testar a qualidade dos resultados, utiliza-se o *Evaluate Model*.

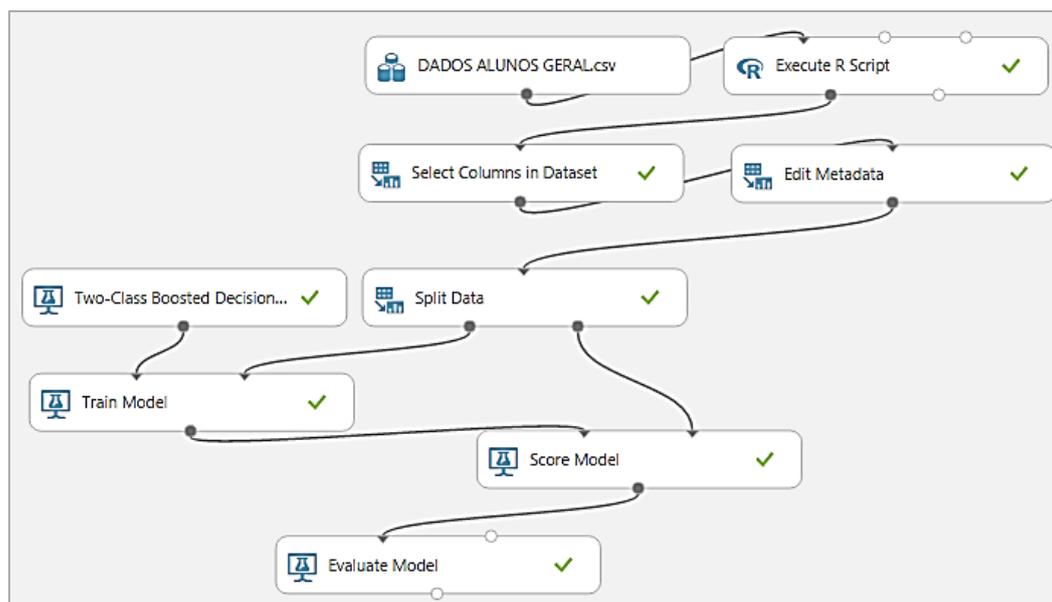
A Figura 7 mostra a análise de dados dos modelos I e II, respectivamente, calculados pelo *Evaluate Model*. Nessas variáveis, quanto menor o valor, melhor, pois maior é a chance de as previsões se aproximarem dos valores reais. Para o *Coefficient of Determination* (Coeficiente de Determinação), o ideal é que esse número esteja bem próximo de 1, pois indica que as previsões foram melhores. O modelo II segue a mesma proposta, havendo apenas um acréscimo de variáveis. Pode-se concluir que, com o acréscimo das variáveis de frequência e índice de aprovação, os resultados foram melhores.

Figura 7. Resultado Experiência 1 dos modelos I e II

TCC RL > Evaluate Model > Evaluation results		TCC RL > Evaluate Model > Evaluation results	
Metrics		Metrics	
Mean Absolute Error	1.208518	Mean Absolute Error	1.09485
Root Mean Squared Error	2.275941	Root Mean Squared Error	1.955589
Relative Absolute Error	0.155653	Relative Absolute Error	0.141013
Relative Squared Error	0.052103	Relative Squared Error	0.038468
Coefficient of Determination	0.947897	Coefficient of Determination	0.961532

Fonte: Dados do trabalho.

Utilizando o mesmo conjunto de dados, foi criado um experimento utilizando-se um algoritmo de classificação. Classificação também é uma das técnicas de aprendizado supervisionado e é utilizada para prever uma resposta de um conjunto definido de categorias. A Figura 8 apresenta o experimento realizado.

Figura 8. Modelo de classificação

Fonte: Dados do trabalho.

Nesse experimento, foi utilizado um módulo de execução de script. O script foi escrito em linguagem R e calcula a média final do aluno, baseado nas outras variáveis que foram utilizadas no primeiro experimento. Se a nota for maior ou igual a 60, indica que o aluno foi aprovado, se não, reprovado. Essa regra é baseada no modelo de avaliação utilizado na instituição. O módulo *Two-Class Boosted Decision Tree* é o responsável pela classificação binária (aprovado, reprovado).

Figura 9. Resultado Experimento 2

nota_media_professor	nota_pi	nota_avin	disciplinas_aprovadas	disciplinas_reprovadas	Scored Labels	Scored Probabilities
						
17.37	19.3	8	1	6	fail	0.0001
50.26	19	16.15	5	0	pass	0.99993
31.18	16	12	3	3	pass	0.82013
24.82	16	5.33	0	5	fail	0.000093

Fonte: Dados do trabalho.

A Figura 9 mostra alguns exemplos de resultados após a execução do algoritmo de classificação. À esquerda, são apresentadas cinco variáveis que pertencem ao conjunto de dados. A coluna *Scored Labels* mostra a previsão de cada registro. Pode-se observar que o somatório das notas do primeiro registro é de 44,67, o que, de fato, implicaria reprovação, como mostra o *Scored Labels*. Mas o somatório das notas não é a chave para a decisão, todos os dados são analisados, e isso faz com que o terceiro registro, apesar de não possuir 60,0 pontos, acabe sendo aprovado. A coluna *Scored Probabilities* apresenta a probabilidade de o registro de dados pertencer à classe aprovado. Quanto mais próximo de 1 esse valor for, maior é a chance de essa previsão ser correta. Se o valor é muito baixo, ele decide, então, que o registro não pertence a essa classe e atribui reprovação.

5. Conclusão

O desenvolvimento do trabalho teve como objetivo modelar e desenvolver um sistema de avaliação baseado em técnicas de aprendizado de máquina. O sistema permite que os alunos da instituição realizem avaliações com conteúdos avaliados pelo ENADE, a fim de buscar melhores resultados. O emprego das técnicas de aprendizado de máquina utilizando algoritmos de classificação e regressão sobre os dados da instituição tiveram resultados positivos.

No desenvolvimento do sistema, houve dificuldade para alimentá-lo com questões de maneira que ele pudesse ser utilizado. Para realização de testes, foram utilizadas questões das provas anteriores do ENADE. O incentivo aos alunos para utilizarem com seriedade o sistema é também um desafio que precisa ser trabalhado pela instituição.

Futuramente, pretende-se integrar o sistema de avaliação com o *webservice*, que é disponibilizado pela *Microsoft Azure Machine Learning Studio*, de maneira que, ao preparar-se para fazer uma avaliação, o aluno já consiga receber uma previsão de seus resultados, que pode ser comparada ao seu resultado real, estimulando-o a melhorar o aproveitamento na avaliação. Da mesma forma, os alunos poderão, ao longo do semes-

tre letivo, receber uma previsão em relação ao seu resultado final, alertando-os para uma possível reprovação ou estimulando-os a melhorarem seus resultados.

Os resultados dos alunos nas avaliações feitas no sistema também servirão de base para empregar outros algoritmos de análise preditiva, de forma a tentar buscar por resultados cada vez melhores que sirvam para tomada de decisões da instituição.

Referências

BRINK, H.; RICHARDS, J. *Real world Machine Learning*. [S.l.]: Manning Publications C.O, 2014.

CARVALHO, André Carlos Ponce de Leon F. *et al. Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro: LTC, 2011.

CERVI, Rejane de Medeiros. *Didática*. Curitiba: IESDE. 2001.

DEVMEDIA. *Desenvolvimento ágil com scrum: uma visão geral*. Disponível em: <<http://www.devmedia.com.br/desenvolvimento-agil-com-scrum-uma-visao-geral/26343>>. Acesso em: 20 fev. 2016.

FREIRE, Paulo. *Pedagogia da Autonomia: saberes necessários à prática educativa*. São Paulo: Paz e Terra, 2003.

KNINBERG, Henrik. *Scrum e XP direto das trincheiras: como nós fazemos Scrum*. C4media, 2007. Disponível em <https://www.cti.ufu.br/sites/cti.ufu.br/files/scrum-e-xp-direto-das-trincheiras.pdf>

MORAN, J. M. *O que é Educação a Distância*. Universidade de São Paulo, 2009. Disponível em: <<http://www.eca.usp.br/prof/moran/dist.htm>>. Acesso em: 21 fev. 2016.

PRESSMAN, Roger S. *Engenharia de software*. 6. ed. Rio de Janeiro: McGraw-Hill, 2006.

SCHWABER, Ken; SUTHERLAND, Jeff. *Guia do scrum: um guia definitivo para o Scrum: as regras do jogo*. São Paulo: Desc, 2013.