

7.08.07 - Educação / Tópicos Específicos de Educação.

USO DE TÉCNICAS DE MINERAÇÃO DE DADOS EDUCACIONAIS PARA PREVER O DESEMPENHO DO PROFESSOR DE EAD EM AMBIENTES VIRTUAIS DE APRENDIZAGEM.

Humberto Rabelo^{12*}, Aquiles M. F. Burlamaqui¹³, Ricardo A. M. Valentim¹⁴, Danieli S.S.R.⁴

1. PPgECC/UFRN - Programa de Pós-graduação em Engenharia Elétrica e Computação/UFRN

2. DCT/UFRN – Departamento de Ciência e Tecnologia

3. ECT/UFRN – Escola de Ciência e Tecnologia / Orientador

4. SEDIS/UFRN – Secretaria de Educação a Distância

Resumo:

Este texto descreve a aplicação de técnicas de mineração de dados em ambiente virtual de aprendizagem educacional, que utiliza o Moodle®[7] como plataforma para realização de cursos de graduação a distância. O experimento utiliza dados de uma base histórica, que contém dez cursos de graduação, é parte de um estudo que procura melhorar o processo de ensino a distância, elaborado por doutorandos do Programa de Pós-graduação em Engenharia Elétrica e Computação e por profissionais da Secretaria de Educação a Distância (SEDIS) da Universidade Federal do Rio Grande do Norte (UFRN).

Para realizar a pesquisa, foram selecionados 105 professores de um semestre letivo, e através das suas interações no ambiente, projetaram-se os indicadores representativos por meio de estatística descritiva para construir o modelo preditivo que visa prever se um professor terá ou não um bom desempenho no ambiente virtual. A técnica de mineração de dados aplicada foi a Classificação Bayesiana, que se mostrou eficaz para realizar a previsão anunciada.

Palavras-chave: Ambiente virtual de aprendizagem; Indicador de desempenho; Mineração de Dados Educacionais.

Introdução:

A mineração de dados é definida por [6, p.6], [8] como o processo de descobrir padrões úteis ou conhecimento a partir de dados. A mineração de dados educacionais produz métodos e técnicas que objetivam a descoberta de padrões que forneçam conhecimentos utilizáveis na melhoria dos processos de ensino e aprendizagem [5].

Cada ação realizada no Moodle é registrada no ambiente, afirma [9,p.4]. Uma grande dificuldade em ambientes virtuais de aprendizagem é realizar acompanhamento efetivo das ações dos usuários no ambiente[1], devido a grande massa de dados gerada na utilização do ambiente, e a complexidade de exploração destes dados.

Um elevado grau de colaboração e participação é indicativo de um ambiente que propicia uma grande quantidade de relacionamentos e usos entre atores e objetos [4], [9]. As interações entre professores, tutores, alunos, com as atividades e recursos disponíveis no ambiente, são valiosos indicadores [3] sobre a efetividade do processo de aprendizagem.

Atualmente, o interesse em mineração de dados e sistemas educacionais tem aumentado, tornando-se uma comunidade de pesquisa em crescimento.

Diante deste cenário, este trabalho tem como objetivo trazer destaque ao tema, ao descrever a utilização de técnicas de mineração de dados para prever o desempenho do professor no ambiente virtual de educação a distância, e para possibilitar um melhor acompanhamento da atuação do professor com o propósito de melhorar o seu desempenho na utilização do ambiente

Metodologia:

Esta é uma pesquisa tecnológica, que analisa os dados históricos de um ambiente virtual de aprendizagem e emprega técnicas de mineração de dados, com vista a apoiar e aprimorar o processo de aprendizagem. É também uma pesquisa exploratória, ao fazer levantamento bibliográfico sobre estudos disponíveis sobre a utilização de algoritmos de mineração de dados e modelos preditivos, e analisar dados reais de um ambiente virtual de aprendizagem em pleno funcionamento na [11]. O processo de descoberta do conhecimento através dos dados históricos analisados neste estudo foi realizado em três etapas:

A primeira aborda a preparação dos dados [2, p.5], realizada a partir das interações armazenadas no repositório de dados histórico do ambiente de pesquisa [11], utilizando SQL, ou seja, Linguagem de Consulta Estruturada identificou um conjunto de aproximadamente 60 ações (atributo *action* encontrado nas tabelas de *Log* do Moodle) realizadas por usuários com perfil de professor[10]. Por meio de estatística descritiva, média, desvio padrão, coeficiente de correlação de Pearson[12], selecionamos nove ações mais representativas (relativas a login, utilização de recursos e atividades como tarefas, questionários, fóruns, chat, mensagem no ambiente), como indicadores de desempenho que adotamos para a construção do modelo preditivo, utilizado neste trabalho para prever o desempenho do professor quanto a utilização do ambiente virtual nos cursos de graduação a distância (EaD) da Universidade Federal do Rio Grande do Norte.

A segunda etapa emprega técnicas de mineração de dados educacionais [9], para identificar precocemente se um professor terá ou não terá um bom desempenho no ambiente virtual, além de identificar padrões de comportamento que levam a um professor ter ou não ter um bom desempenho no ambiente virtual, e auxiliar a identificar quais ações ou atributos são mais importantes ou expressivos ou estejam relacionados ao desempenho.

Para fazer a mineração de dados propriamente dita fizemos uso da ferramenta computacional WEKA[13]. A técnica de mineração de dados empregada neste trabalho foi o algoritmo de Classificação Bayesiana que recebe este nome por ser baseado no teorema de probabilidade de Bayes[2]. Ele realiza a predição dos resultados baseado em probabilidades, ou seja, faz uma classificação estatística, e tem como objetivo calcular a probabilidade que uma amostra desconhecida pertença a cada uma das classes possíveis, ou seja, predizer a classe mais provável estatisticamente. O Algoritmo foi empregado com os parâmetros de sua implementação padrão. Existem parâmetros que podem ser alterados na técnica de mineração de dados aplicada no experimento, em trabalhos futuros ajustes podem ser feitos para melhorar a precisão, e estudos mais detalhados para cada caso podem ser realizados.

A terceira e última etapa avalia os resultados obtidos após a etapa de mineração dos dados, e objetiva Identificar padrões relevantes, e prever a classificação de novos registros.

Resultados e Discussão:

Com Modelo preditivo adotado obtivemos alto índice de acertos na predição do desempenho do professor no ambiente virtual, a acurácia foi superior a 90% de precisão como pode ser observado.

<i>Correctly Classified Instances</i>	96 = 91.4286 %
<i>Incorrectly Classified Instances</i>	9 = 8.5714 %
<i>Kappa statistic</i>	0.8885
<i>Total Number of Instances</i>	105

O experimento foi realizado com a finalidade de obter um Modelo Preditivo de desempenho contendo um número de 105 instâncias de professores. Na validação do modelo foram encontradas 96 instâncias classificadas corretamente, o que corresponde a 91,5% de acertos, e foram encontradas 09 instâncias classificadas incorretamente, o que corresponde a 8,5% de erros.

Tão importante quando o número de instâncias corretamente classificadas é a estatística de Kappa, métrica que indica o nível de concordância (coesão) dos dados classificados, no caso o valor do índice Kappa deste experimento foi de 0,89, considerado excelente, uma vez que sua variação vai de 0 a 1 de magnitude.

Conclusões:

Através da utilização de mineração de dados e com o modelo preditivo de desempenho do professor no ambiente virtual de educação a distancia, não é necessário esperar pelo final da disciplina ou do semestre para saber se o professor terá um desempenho positivo. Neste experimento, através da técnica de mineração de dados que utiliza o algoritmo de Classificação Bayesiana foi possível prever com precisão de 91,5% se um professor terá ou não um desempenho satisfatório. Tal previsão ao ser realizada possibilita que, durante o andamento de uma disciplina, a equipe de supervisão pedagógica e coordenação do curso, realizem um melhor acompanhamento do desempenho dos diversos professores envolvidos, e assim possam vir a propor a determinado professor, ações de ajuste de conduta personalizadas a serem tomadas durante o processo de ensino no propósito de melhorar seu desempenho na utilização do ambiente, e assim melhorar também o desempenho de seus alunos e diminuir a evasão do curso[10].

Quando comparamos turmas em que o desempenho do professor foi positivo, com turmas que o desempenho do professor não foi positivo, constatamos que isso também ocorre no desempenho dos alunos da turma, quer dizer, o desempenho do professor, seja

positivo ou negativo, tem influência e correlação no desempenho dos seus alunos no ambiente virtual.

Diante do exposto, constatamos que o desempenho do professor no ambiente virtual tem influência no desempenho dos seus alunos, e assim consideramos que o acompanhamento da predição do desempenho do professor pode ser um fator relevante na redução da evasão de alunos, e contribuir para melhoria da qualidade do ensino em educação a distancia.

Referências bibliográficas

[1] ANDRADE, Fábio Goulart; ROSITO, Mauricio Covolan; DA SILVA, Julia Marques Carvalho. Análise de Ferramentas de Suporte às Tarefas do Professor-Tutor em um Ambiente Virtual de Ensino e Aprendizagem. In: Anais dos Workshops do Congresso Brasileiro de Informática na Educação. 2016. p. 92.

[2] CAMILO, Cássio Oliveira; SILVA, João Carlos da. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. Universidade Federal de Goiás (UFG), p. 1-29, 2009.

[3] FERRAZ, Paulo Fernando Ozório; OLIVEIRA, Pedro Torres de; HORNINK, Gabriel Gerber. Desenvolvimento e Implementação de Indicadores de Colaboração e Participação no Moodle. Informática na Educação: teoria e prática, Porto Alegre, v. 18, n. 1, p. 85-95, jan./jun. 2015.

[4] FORBECK, VERA LÚCIA ALMEIDA. Análise da Interatividade em Ambientes Virtuais de Aprendizagem usando Métricas baseadas em Regras. 2008. Tese de Doutorado. Dissertação (Mestrado em Ciência da Computação)–Universidade Metodista de Piracicaba UNIMEP.

[5] GOTTARDO, Ernani; KAESTNER, Celso; NORONHA, Robinson Vida. Previsão de desempenho de estudantes em cursos ead utilizando mineração de dados: uma estratégia baseada em séries temporais. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2012.

[6] LIU, Bing. Web data mining: exploring hyperlinks, contents, and usage data. Springer Science & Business Media, 2007.

[7] MOODLE, Moodle.org: open-source community-based tools for learning. Disponível em <<http://moodle.org/>>. Acesso em: 13/05/2016.

[8] ROMERO, Cristobal; VENTURA, Sebastian. Educational data mining: A survey from 1995 to 2005. Expert systems with applications, v. 33, n. 1, p. 135-146, 2007.

[9] ROMERO, Cristóbal; VENTURA, Sebastián; GARCÍA, Enrique. Data mining in course management systems: Moodle case study and tutorial. Computers & Education, v. 51, n. 1, p. 368-384, 2008.

[10] QUEIROGA, Emanuel; CECHINEL, Cristian; ARAÚJO, Ricardo. Um Estudo do Uso de Contagem de Interações Semanais para Predição Precoce de Evasão em Educação a Distância. In: Anais dos Workshops do Congresso Brasileiro de Informática na Educação. 2015. p. 1074.

[11] SEDIS, Secretaria de Educação a Distância – UFRN. Disponível em <www.sedis.ufrn.br/> Acesso em: 13/04/2016.

[12] SILVESTRE, António. Análise de dados e estatística descritiva. Escolar editora, 2007.

[13] WEKA - Waikato Environment for Knowledge Analysis. Weka 3: Data Mining Software in Java. Disponível em <<http://www.cs.waikato.ac.nz/ml/weka/>> Acesso em: 13/04/2016.