

SISTEMA DE IDENTIFICAÇÃO DE MÉTRICAS DE QUALIDADE PARA OBJETOS DE APRENDIZAGEM DENTRO DE REPOSITÓRIOS

VINICIUS VIEIRA DOS SANTOS¹; CRISTIAN CECHINEL²

¹Universidade Federal de Pelotas - vvdsantos@inf.ufpel.edu.br

²Universidade Federal de Pelotas - contato@cristiancechinel.pro.br

1. INTRODUÇÃO

Objetos de aprendizagem são definidos como “qualquer entidade, digital ou não digital, que possa ser utilizada, reutilizada ou referenciada durante o aprendizado apoiado por meios tecnológicos” (IEEE LTSC, 2002).

Com a grande quantidade de objetos de aprendizagem na internet tornou-se uma tarefa extenuante para o ser humano analisar a qualidade desses recursos e, em decorrência disto, foi originada uma necessidade de alternativas de avaliações automatizadas que possam minorar tal afazer. Com a grande variedade de objetos de aprendizagem dentro de repositórios (OCHOA; DUVAL, 2009), em companhia de metadados avaliativos - que são informações contextuais sobre objetos de aprendizagem fornecidas pelos usuários - pôde-se vislumbrar a possibilidade de se obter medidas intrínsecas dos objetos de aprendizagem, as quais possam ser utilizadas como possíveis indicadores de qualidade.

Na avaliação de medidas intrínsecas, determinadas informações são extraídas dos objetos para que medidas quantitativas de boa e má qualidade sejam comparadas com o intuito de descobrir quais são os atributos intrínsecos associados a qualidade do referido objeto. Atributos intrínsecos consistem de toda informação que possa ser obtida de maneira automática a partir do próprio objeto como: número de palavras, número de imagens, tamanho do objeto, número de links, etc. Estes atributos podem, assim que associados a qualidade dos objetos de aprendizagem, permitir a criação de perfis estatísticos de recursos bons e ruins que, por sua vez, servirão como base para a previsão automática da qualidade. Esta abordagem já foi usada anteriormente, com êxito, na análise automática sobre a usabilidade de sites (IVORI; HEARST, 2002).

Um dos trabalhos realizados com o intuito de avaliar a qualidade dos objetos de aprendizagem é o trabalho de Cechinel; Sanchez-Alonso e Sicilia (2012) que teve como fim explorar perfis estatísticos de objetos altamente pontuados e criar modelos para a avaliação automática de sua qualidade, usando como base os objetos contidos no repositório de materiais *online* para ensino-aprendizagem MERLOT¹. Neste trabalho, os experimentos estatísticos realizados foram feitos em um pequeno subconjunto de dados, mas existe uma grande quantidade de combinações de categorias de disciplinas e tipo de material a ser considerada. O trabalho proposto neste documento foi elaborado na forma de um software e utilizará estes mesmos experimentos, mas para todos os subconjuntos existentes e de forma automática e integrada à base de dados.

¹ <http://www.merlot.org>

2. METODOLOGIA

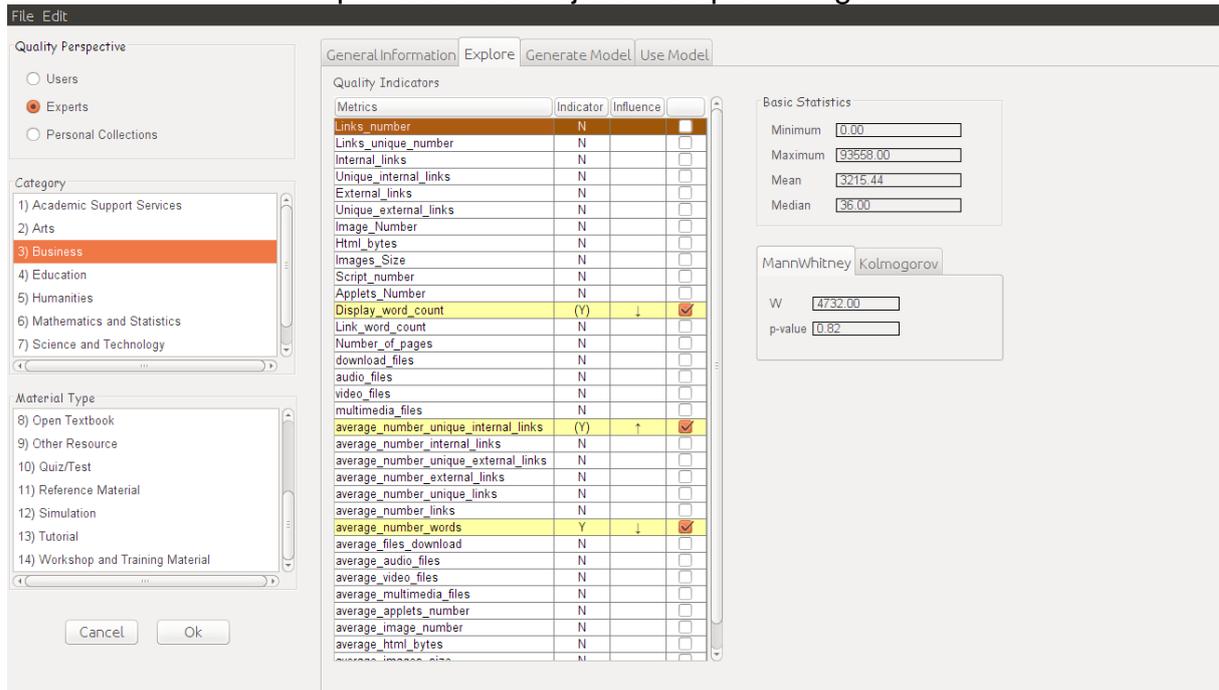
A análise para definição dos perfis estatísticos dos objetos de aprendizagem começa com o usuário clicando nas diferentes categorias, tipos de materiais e perspectivas de qualidade. Conforme vão sendo selecionados os subconjuntos, consultas utilizando estas intersecções são feitas no banco de dados e as informações existentes na tela sobre o número de objetos são atualizadas. Quando um subconjunto é selecionado a fim de fazer uma avaliação, fazer uma avaliação implica em pelo menos selecionar uma perspectiva de qualidade, e o botão ok da tela inicial é pressionado, os seguintes eventos acontecem na ordem descrita abaixo:

- Cálculo dos tercís do subconjunto;
- É criada uma pasta chamada graphics no diretório local em que o programa está sendo executado e um gráfico de densidade é salvo nesta pasta;
- Os objetos são separados em dois grupos, bons e não bons, de acordo com o valor dos tercís calculados;
- Cálculo das estatísticas de Mann-Whitney (Bussab; Morettin, 2009) e Kolmogorov-Smirnov (Sheskin, 2003) com o grupo dos bons e não bons para cada uma das 34 métricas existentes. Foi considerado que o grupo dos bons e não bons teve uma diferença de perfis sobre uma determinada métrica quando ambas as distribuições e medianas apresentaram diferença significativa ao nível de confiança de 90% para os limiares avaliados. Quando tal situação é observada (quando ambas as distribuições e medianas apresentaram diferença significativa) significa que a métrica avaliada está associada (ou correlacionada) com a qualidade do OA;
- Cálculo da mediana do grupo dos bons e dos não bons para cada uma das 34 métricas existentes;
- Cálculo da influência de cada uma das métricas. Este cálculo é feito comparando os valores das medianas dos dois grupos, se a mediana do grupo dos bons é maior que a mediana do grupo dos não bons para uma determinada métrica, significa que ela tem influência positiva na qualidade do OA, caso contrário ela tem influência negativa. Isto é válido apenas para métricas que são consideradas relevantes para definir a qualidade do objeto. Na tabela onde é mostrada a influência das métricas, uma métrica que influencia positivamente na qualidade do OA é representada pelo símbolo \uparrow e uma métrica que influencia negativamente é representada pelo símbolo \downarrow ;
- Calcula quais métricas estão associadas ou correlacionadas com a qualidade dos objetos de aprendizagem do subgrupo selecionado. Isto é feito comparando o valor p das estatísticas de Mann-Whitney e Kolmogorov-Smirnov, para cada uma das métricas. Se os valores do p de Mann-Whitney e do p de Kolmogorov-Smirnov calculados estão abaixo de 0.10 podemos dizer que as distribuições e medianas apresentaram diferença significativa ao nível de confiança de 90% e, portanto, a métrica analisada deve ser considerada um indicador de qualidade parcial do OA. Caso os valores p das estatísticas de Mann-Whitney e Kolmogorov-Smirnov estejam ambas abaixo de 0.05 as distribuições e medianas apresentaram diferença significativa ao nível de confiança de 95% e a métrica deve ser considerada como um indicador de qualidade.
- São adicionadas nos seus respectivos lugares as informações obtidas após o cumprimento dos passos acima descritos.

Após uma avaliação ser realizada, é possível criar um modelo para avaliação de novos objetos. Para isso, foi disponibilizada uma área em que ao pressionar o botão ok um modelo é criado, fazendo uso do algoritmo de árvore de decisão J48 do weka², e informações como a árvore criada e a avaliação do modelo podem ser visualizadas na tela.

A Figura 1 mostra as métricas encontradas que estão relacionadas com a qualidade após uma avaliação ser feita em um subconjunto.

Figura 1 – Visualização das métricas encontradas que estão relacionadas com a qualidade dos objetos de aprendizagem.



Assim que o modelo é criado existem duas opções a serem escolhidas, ou o modelo pode ser salvo para uso posterior, ou ele pode ser usado imediatamente para avaliar novos objetos de aprendizagem. Caso o usuário decida utilizar o modelo recém-criado para classificar um novo objeto, ele deve informar o endereço eletrônico do objeto no campo URL da aba use model do software e clicar no botão get metrics. Neste momento o web crawler, que foi integrado com o software, irá extrair as 34 métricas do objeto e armazená-las no banco de dados. Depois que as métricas foram armazenadas no banco de dados, são mostradas em uma tabela apenas as métricas que foram utilizadas para criar o modelo mas com os valores que foram obtidos do OA que está sendo analisado. Somente então, o usuário pode pressionar o botão classify data para classificar o OA como bom ou não bom.

3. RESULTADOS E DISCUSSÃO

Até o momento o software é capaz de criar modelos de avaliação automáticos da qualidade de objetos de aprendizagem e usar os mesmos para avaliar novos recursos.

² <http://www.cs.waikato.ac.nz/ml/weka/>

Alguns modelos gerados pelo software, usando como parâmetros as métricas intrínsecas identificadas como potenciais indicadores de qualidade para os subconjuntos alvo chegam a resultados satisfatórios.

4. CONCLUSÕES

O trabalho desenvolvido apresenta uma inovação no modo de avaliação de objetos de aprendizagem dentro do repositório MERLOT reduzindo significativamente o esforço humano empregado nesta tarefa. Um dos trabalhos futuros para o incremento do software é a inclusão de outros algoritmos do weka para a criação dos modelos. Outro trabalho que pode ser feito é a inclusão de outras variáveis na filtragem dos subconjuntos para avaliação como, por exemplo, a linguagem do material.

5. REFERÊNCIAS BIBLIOGRÁFICAS

IEEE LTSC. **IEEE Standard for Learning Object Metadata**, nov. 2002. Acessado em: 19 jun. de 2014. Online. Disponível em: <http://ltsc.ieee.org/wg12/>

OCHOA, X.; DUVAL, E. Quantitative Analysis of Learning Object Repositories. **Learning Technologies, IEEE Transactions on**, v.2, n.3, p.226–238, 2009.

CECHINEL, C.; SANCHEZ-ALONSO, S.; SICILIA, M.-A. Avaliação Automática da Qualidade de Objetos de Aprendizagem dentro de Repositórios. **Revista Brasileira de Informática na Educação**, v.20, n.3, p.43-59, 2012.

IVORI, M. Y.; HEARST, M. A. Statistical profiles of highly-rated web sites. In: **SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI '02)**, Minneapolis, 2002, **Proceedings...** Nova York: ACM, 2002. p.367–374.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. São Paulo: Saraiva, 2004.

SHEKIN, D. J. **Handbook of Parametric and Nonparametric Statistical Procedures**. Flórida: CRC Press, 2003.