

Opiniões de usuários para avaliação de uma nova proposta de recomendação de objetos de aprendizagem

Henrique L. dos Santos¹; Cristian Cechinel²

¹Universidade Federal de Pelotas – hldsantos@inf.ufpel.edu.br

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

1. INTRODUÇÃO

Objetos de aprendizagem (OAs) podem ser definidos como quaisquer recursos digitais que possam ser iterativamente utilizados a fim de melhorar o processo de aprendizagem, isto é, OAs são unidades de ensino que podem ser consideradas como sendo o núcleo da área de *Technology Enhanced Learning* (sigla TEL, área que se ocupa do estudo das tecnologias de ensino). Além disso, os OAs constituem uma grande porção dos recursos educacionais abertos disponíveis atualmente. Esses objetos tomam diferentes formas e podem ser reutilizados, combinados, atualizados e referenciados. OAs podem ser usados tanto por estudantes que desejam aprender determinado assunto quanto por professores que queiram produzir material para suas aulas. Os OAs são normalmente descritos por metadados, tais como título, descrição, tipo de material, disciplina, idioma, entre outros.

De maneira geral, objetos de aprendizagem são armazenados e disponibilizados à comunidade através de repositórios (ROAs). Estes repositórios fornecem recursos a diferentes comunidades de estudantes, educadores e outros interessados, o consumo desses recursos pode-se dar não apenas diretamente através do repositório, mas também indiretamente (quando o repositório funciona apenas como um indexador, também chamado de “referatório”). Os ROAs existentes diferem entre si de diversas formas, por exemplo, na área de estudos que abrangem, nos tipos de materiais disponibilizados e, especialmente, nos padrões de metadados utilizados (MCGREAL, 2008).

Alguns repositórios também permitem que seus usuários se registrem e façam comentários e avaliações acerca dos OAs por eles consumidos. Isso permite o uso dessas avaliações como forma de recomendação de recursos aos demais usuários. O processo de recomendação de OAs é cada vez mais importante tendo em vista o grande crescimento de recursos disponibilizados por cada repositório.

O trabalho aqui apresentado é uma continuação do trabalho exposto por (SANTOS & CECHINEL, 2015) no qual os autores avaliaram (de maneira *offline*) um novo método de pré-processamento de dados para recomendação de OAs utilizando filtragem colaborativa (FC). No presente trabalho, além da avaliação *offline* de outro algoritmo de similaridade no novo método, também serão expostos os resultados de uma avaliação preliminar, com usuários reais, das recomendações geradas pelo método proposto versus recomendações geradas pelo método tradicional de filtragem colaborativa (sem qualquer tipo de pré-processamento. (VERBERT *et al.*, 2011) afirma que experimentos desse tipo, onde o *feedback* de usuários reais são avaliados, são especialmente úteis se forem desenvolvidos utilizando grande quantidade de dados reais, como por exemplo os do repositório Merlot (www.merlot.org) aqui utilizados.

O objetivo do trabalho é verificar se a avaliação *offline* do novo método de recomendação encontra-se de acordo com a avaliação de usuários reais. Como exposto por (HERLOCKER *et al.*, 2004), a análise *offline* tem sua utilidade na

praticidade de se avaliar diversos métodos, porém fracassa na tentativa de avaliar a adequação das recomendações uma vez que não dispõe da opinião de usuários reais.

2. METODOLOGIA

2.1 Avaliação offline (erro MAE)

Foram coletadas 9910 avaliações (que variam num intervalo fechado de 1 a 5) feitas por 3659 usuários sobre 4968 diferentes OAs. Além disso, para possibilitar a aplicação das técnicas de pré-processamento, também foram coletadas algumas informações sobre os OAs avaliados. O método proposto consiste na clusterização desses objetos por meio da semelhança textual das suas descrições, para isso foi utilizado o algoritmo *k-means* que executou a clusterização em vetores TF-IDF que representavam cada objeto. O número de *clusters* gerados (*k*) variou de 1 a 9, onde 1 representa o método tradicional, uma vez que não há recorte no *dataset* original.

Após a clusterização, cada *cluster* foi tratado como uma base de dados individual e um algoritmo de filtragem colaborativa baseada em usuário foi responsável por gerar e avaliar as recomendações para cada *cluster*. Numa abordagem de treinamento-teste com 90% dos dados dedicados a treinamento, os parâmetros desse algoritmo de FC também foram experimentados, entre eles: tamanho de vizinhança (2 a 20) e algoritmo de similaridade entre usuários (*LogLikelihood Ratio* e Distância Euclideana). Finalmente, optou-se por utilizar a métrica de erro clássica MAE (erro médio absoluto) para verificar o quão preciso cada recomendador performou.

2.2 Avaliação online (feedback de usuários reais)

Foram recrutados 18 voluntários (provenientes de diversos países, tais como Brasil, Argentina, Espanha, Equador, etc) aos quais foi pedido que avaliassem cerca de 10 OAs (quaisquer) presentes no repositório Merlot. Foram obtidas 108 novas avaliações provenientes desses voluntários, essas avaliações foram aglutinadas ao *dataset* original descrito na subseção anterior. Para essa avaliação, escolheu-se confrontar dois métodos de recomendação: o novo método proposto com a configuração que obteve melhor desempenho na análise *offline* (doravante chamado: cluster) e o método tradicional de filtragem colaborativa pura (doravante chamado: geral). Para cada usuário voluntário, foram geradas no máximo 4 recomendações (2 de cada abordagem).

Cada recomendação foi apresentada ao usuário juntamente com o seguinte questionário:

- Q1 – Quão relevante você considera essa recomendação?
- Q2 – Quão difícil você considera o conteúdo desse recurso?
- Q3 – Qual a sua nota para esse recurso?

Todas as questões permitiam resposta via Escala Likert (1 a 5, ordem crescente de intensidade).

3. RESULTADOS E DISCUSSÃO

No que diz respeito à avaliação *offline*, o algoritmo de similaridade Distância Euclidiana apresentou um desempenho ligeiramente superior ao algoritmo *LogLikelihood Ratio*, como pode ser observado na Figura 1. Também observa-se

que a abordagem proposta somente obtém desempenho melhor do que a filtragem colaborativa tradicional ($k=1$) para valores altos de k , por exemplo, 6 e 8.

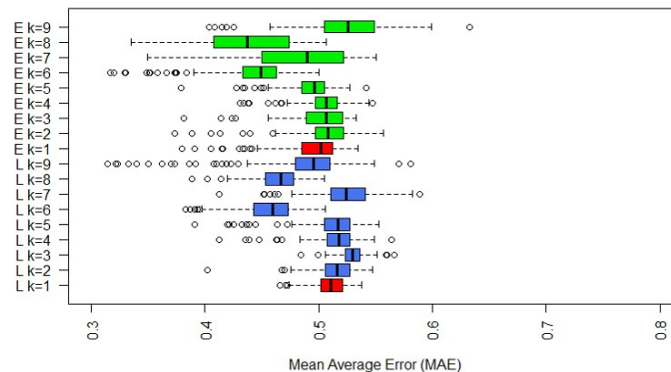


Figura 1. Boxplot para o MAE da abordagem de clusterização de usuários usando Distância Euclideana e *LogLikelihood Ratio*.

Outra medida importante na avaliação de recomendadores é a cobertura do espaço de usuários. (GE *et al.*, 2010) afirma que a cobertura está ligada ao nível sobre o qual os recursos são recomendados a todos usuários em potencial e também em relação à quantidade de recursos efetivamente recomendados. Para esse estudo, tomamos a métrica como sendo a porcentagem de usuários (em relação ao total de usuários do *dataset*) que receberam ao menos uma recomendação. Os resultados podem ser vistos na Figura 2 e, novamente, percebe-se um desempenho melhor do algoritmo de Distância Euclideana.

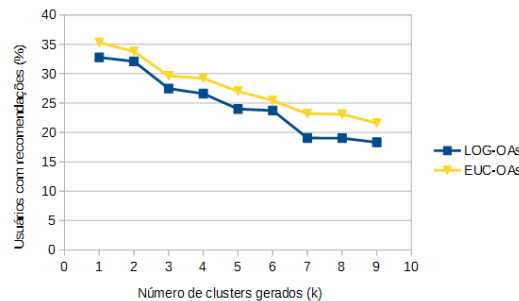


Figura 2. Cobertura do espaço de usuários para cada algoritmo.

Para a análise *online* utilizou-se, no método proposto, a configuração que teve melhor performance na avaliação anterior, seja ela: algoritmo de similaridade Distância Euclideana, número de clusters igual a 6 e tamanho de vizinhança igual a 10. Essa configuração gerou recomendações que foram disponibilizadas aos usuários juntamente com recomendações geradas pelo método tradicional, sem clusterização. Os resultados do *feedback* podem ser vistos na Figura 3. Na imagem, percebe-se levemente um melhor desempenho das recomendações geradas pelo método tradicional, especialmente no que diz respeito à nota e a relevância, uma vez que no método tradicional a concentração de avaliações se dá especialmente nos valores mais altos enquanto que no método de *clusters* a concentração maior de opiniões concentra-se em valores baixos (2 e 3).

Entretanto, a fim de validar o resultado obtido junto aos usuários, foi executado um Teste de medianas de Mann-Whitney-Wilcoxon para verificar se houve diferença significativa entre os dois grupos de recomendação, considerando a relevância, a nota e a dificuldade avaliadas. Os resultados, obtidos com um nível de confiança de 95%, mostraram que não é possível atestar

a diferença significativa entre os dois grupos para qualquer uma das variáveis avaliadas.

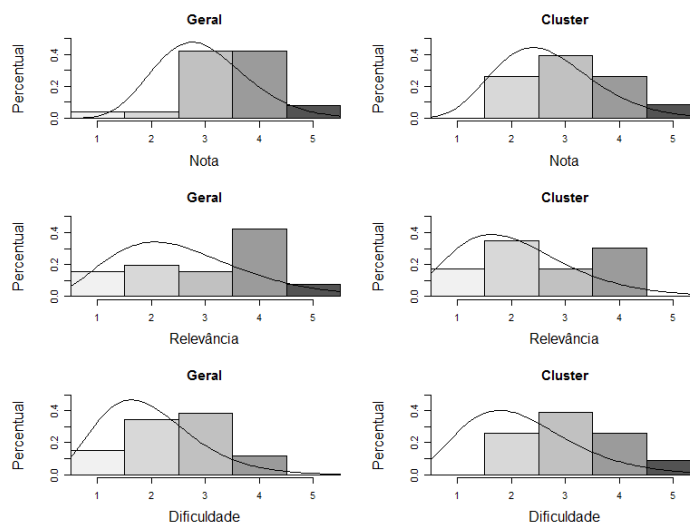


Figura 3. Avaliações dos usuários em relação às recomendações geradas.

4. CONCLUSÕES

A avaliação de usuários reais aqui apresentada não permitiu definir qual método de fato tem um melhor desempenho nas recomendações. Tal fato indica que ainda é necessária uma futura reavaliação do método proposto para efetivar sua qualidade. Uma possibilidade é a introdução da ideia de recomendador baseado em crítica, utilizando os *feedbacks* fornecidos pelos usuários como forma de refinamento do processo de recomendação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- MCGREAL, R. A typology of learning object repositories. In: ADELSBERGER, H.; PAWLOWSKI, J.; SAMPSON, D. **Handbook on Information Technologies for Education and Training**. Heidelberg: Springer Berlin, 2008. p.5-28.
- HERLOCKER, J.; KONSTAN, J.; TERVEEN, L. e RIEDL, T. Evaluating collaborative filtering recommender systems. **ACM Trans. Inf. Syst.**, vol. 22, no. 1, p. 5–53, 2004.
- GE, M.; DELGADO-BATTENFELD, C.; JANNACH, D. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. **Proceedings of the Fourth ACM Conference on Recommender Systems**. Nova Iorque: ACM. p. 257–260, 2010.
- VERBERT, K.; DRACHSLER, H.; MANOUSELIS, N.; WOLPERS, M.; VUORIKARI, R. e DUVAL, E. Dataset-driven research for improving recommender systems for learning. In: **1st International Conference on Learning Analytics and Knowledge**. Nova Iorque, 2011. Proceedings of the 1st International Conference on Learning Analytics and Knowledge. Nova Iorque: ACM, 2011, pp. 44–53.
- SANTOS, H. L.; CECHINEL, C. Avaliação dos impactos de abordagens de pré-processamento em técnicas de filtragem colaborativa no repositório educacional Merlot. In: **XXIV CONGRESSO DE INICIAÇÃO CIENTÍFICA DA UNIVERSIDADE FEDERAL DE PELOTAS**, Pelotas, 2015, **Anais do XXIV Congresso de Iniciação Científica da Universidade Federal de Pelotas** Pelotas: Pró-reitoria de Pós-graduação e Pesquisa, 2015.