

**Descoberta de novos atributos de produtos com base nas
opiniões de usuários publicadas na Web**

**Orientador(a): Tiago Eugenio de Melo
Bolsista: Iranildo Santos Batalha**

1. Introdução

Atualmente, milhares de pessoas postam comentários sobre os mais variados tipos de produtos em sites de comércio eletrônico. Isto resulta em um grande número de opiniões e que podem ser uma valiosa fonte de informação para a tomada de decisão para fabricantes e consumidores. Os fabricantes podem obter um *feedback* imediato para melhorar a qualidade dos seus produtos e os consumidores podem usar esses comentários antes de efetuar a compra de um produto. Porém, o grande volume de opiniões é tamanha que excede a capacidade das pessoas de processarem todos os comentários e isto é uma relevante barreira para o seu uso efetivo [1].

Assuma, por exemplo, que um consumidor esteja interessado nas opiniões de outras pessoas sobre a tela de um telefone celular. Ler todas os comentários é inviável. Realizar uma pesquisa nos comentários usando o termo “tela” não é efetivo porque, comumente, as pessoas escrevem diferentes aspectos da tela, tais como resolução ou contraste, sem realmente usar o termo tela. Além disso, comentários postados em sites de comércio eletrônico são tipicamente escritos por usuários não-técnicos e, conseqüentemente, o texto nem sempre está correto e frequentemente contém erros ortográficos e/ou de tipografia. Assim, a pesquisa sobre como organizar opiniões de usuários é um desafio substancial tanto para a academia como para a indústria.

Recentemente, nós propusemos uma abordagem chamada *OpinionLink* [2] para lidar com o desafio descrito acima de organizar um grande volume de dados opinativos. Nós assumimos a seguinte hipótese: as mais importantes características dos produtos são representadas pelos atributos de catálogo de produtos, fornecidos pelos fabricantes, e comumente disponibilizados em sites de comércio eletrônico. De acordo com Fensel et al. [3], catálogos de produtos são projetados para as pessoas e a sua função é descrever produtos para potenciais consumidores. Portanto, nós argumentamos que o processo de organizar opiniões deveria ser guiado pelos atributos desses catálogos [2]. Para tanto, nós propusemos um método supervisionado baseado em classificadores para a tarefa de mapear as opiniões de usuários sobre produtos eletrônicos para os atributos do catálogo de produtos. O objetivo final desse trabalho foi enriquecer o catálogo de produtos com opiniões extraídas dos comentários de usuários. A Figura 1 mostra um exemplo onde a opinião “**touchscreen é muito bom**” do usuário é mapeada para o atributo `Tela` do produto Notebook Lenovo.

ANTONIO SANTOS
 ★★★★★ Conforme anunciado
 29 de abril de 2019
 Compra verificada

O touchscreen é muito bom, mas achei os teclados um pouco sensíveis na qualidade. Por enquanto, está atendendo minhas necessidades!

Útil | Comentar | Informar abuso

Notebook Lenovo Ideapad 330	
Marca	Lenovo
Tela	15.6 polegadas
Processador	Intel Core i5
Memória	8 GB
...	...

Figura 1: Associação de opiniões com atributos de produtos.

Apesar da boa performance alcançada pelo método *OpinionLink* na tarefa de associar opiniões a atributos de um catálogo de produtos, uma questão aberta que surgiu é como organizar as opiniões que se referem a atributos que não foram listados no catálogo de produtos. Por exemplo, de acordo com estudo que nós publicamos recentemente [5], teclado (“*keyboard*”) é o segundo mais frequente aspecto comentado por usuários na categoria de laptops, mas, tipicamente, não existe um atributo teclado no catálogo de produtos disponibilizado pelos fabricantes e pelos sites de comércio eletrônico. Esse é o exemplo mostrado na Figura 1, onde a opinião “teclados um pouco sensíveis na qualidade” não tem um atributo alvo, pois não existe o atributo teclado do produto Ideapad 330.

Diante desse cenário, pretende-se investigar nessa pesquisa proposta uma abordagem para organizar/mapear as opiniões de usuários relativas a características de produtos que não são representadas como atributos em catálogos. Uma ideia inicial seria agrupar esse tipo de opinião em subgrupos e que esses subgrupos poderiam ser transformados como novos atributos em catálogos de produtos, conforme a sua relevância. A descoberta de novos atributos a partir das opiniões dos usuários permitirá que os consumidores possam fazer uma comparação entre produtos de maneira mais detalhada. Além disso, a aplicação dessa abordagem para a descoberta de novos atributos a partir da opinião dos usuários permitirá que os fabricantes possam identificar os pontos fortes e fracos dos seus produtos relativos a novos atributos e, conseqüente, torná-los melhores.

2. Objetivo(s)

O objetivo deste estudo é investigar o problema de mapear/organizar opiniões de usuários relativas a características de produtos que não são listadas como atributos em catálogos de produtos.

Objetivos Específicos:

a) Estudar sobre trabalhos que façam o mapeamento de opiniões para atributos (alvos) preestabelecidos, tais como o *OpinionLink* mencionado na Introdução. Duração 2 (dois) meses.

b) Pesquisar sobre possíveis métodos que permitam agrupar as opiniões de usuários que não sejam associadas a atributos de produtos. Duração: 2 (dois) meses.

c) Implementar métodos que permitam agrupar as opiniões que não sejam associadas a atributos de produtos e transformar esse tipo de grupo de opiniões em um novo atributo do catálogo de produtos. Duração: 5 (cinco) meses.

d) Avaliar a performance dos métodos implementados e, eventualmente, adaptar esses métodos para obter resultados mais efetivos. Duração: 3 meses.

e) Desenvolver de uma ferramenta Web que permita visualizar os resultados obtidos com a implementação do método proposto. Duração: 2 meses.

3. Metodologia

O projeto de pesquisa será executado nas seguintes fases:

Fase 1:

Na primeira etapa, o aluno irá estudar o método *OpinionLink* e pesquisar sobre outros possíveis métodos que tratem do problema de associar opiniões com atributos de produtos. O professor-orientador proponente é um dos autores desse método e irá auxiliar o aluno na compreensão do método. O problema proposto nesse projeto é um dos trabalhos futuros estabelecidos no artigo publicado do *OpinionLink* [2].

Fase 2:

Na segunda etapa, o aluno fará uma pesquisa sobre possíveis métodos que permitam agrupar as opiniões de usuários que não foram mapeadas para os atributos de produtos. Essa etapa de investigação deverá envolver a implementação ou testes de possíveis métodos que possam ser usados como referências.

Fase 3:

Na terceira etapa, o aluno deverá implementar as possíveis técnicas investigadas na fase anterior.

Fase 4:

Na etapa final, o aluno deverá realizar testes para avaliar a performance dos métodos implementados e, eventualmente, comparar com possíveis *baselines*. Além disso, o aluno implementará um protótipo Web que permita visualizar o uso do método desenvolvido nessa pesquisa.

4. Resultados esperados

Espera-se como resultado do projeto de pesquisa proposto o desenvolvimento de novas tecnologias que tenham tanto relevância acadêmica, caracterizada por meio da publicação de trabalho científico de impacto, como também relevância econômica, caracterizada pelo desenvolvimento de protótipo de protótipo que demonstrem a viabilidade dos métodos propostos no projeto.

5. Cronograma de execução

Meta/Atividade	1	2	3	4	5	6	7	8	9	10	11	12
Revisão bibliográfica sobre o problema proposto.	■	■										
Pesquisa sobre possíveis métodos para tratamento do problema.			■	■								
Implementação dos métodos identificados na fase anterior.					■	■	■	■	■			
Realização de testes para avaliar a performance dos métodos e a eventual implementação de possíveis melhorias.									■	■	■	
Implementação de um protótipo que permita visualizar o funcionamento do método desenvolvido nessa pesquisa.										■	■	
Elaboração do relatório final e submissão de artigo científico para publicação com os resultados obtidos.											■	■

6. Referências

[1] Kwon, B. C., Kim, S.-H., Duket, T., Catalán, A., & Yi, J. S. (2015). *Do people really experience information overload while reading online reviews?* International Journal of Human-Computer Interaction, 31(12), 959-973.

[2] de Melo, Tiago, da Silva Altigran, de Moura Edleno, et al. *OpinionLink: Leveraging user opinions for product catalog enrichment.* Information Processing & Management 56.3 (2019): 823-843.

[3] Fensel, D., Ding, Y., Omelayenko, B., Schulten, E., Botquin, G., Brown, M., & Flett, A. (2001). *Product data integration in b2b e-commerce.* IEEE Intelligent Systems, 16(4), 54-59.

[4] de Melo T, da Silva Altigran, de Moura Edleno S. An aspect-driven method for enriching product catalogs with user opinions. *Journal of the Brazilian Computer Society*. 2018 Dec;24(1):15.

[5] Tiago de Melo (UEA); Altigran S. da Silva (UFAM); Edleno S. de Moura (UFAM); Pável Calado (Ulisboa). *The Importance of Canonical Product Attributes on User Opinions: an Empirical Investigation*. 10th Latin American Web Congress. 2019 (May), San Francisco.