

III Workshop de Dissertações de Mestrado do PPgSI (2016)

FICHA DA PESQUISAⁱ

DADOS GERAIS					
Título do projeto de pesquisa	Recomendação Baseada em Casos: uma abordagem híbrida para recomendação de notícias.				
Orientando	José Luiz Maturana Pagnossim				
Orientador(es)	Sarajane Marques Peres				
Momento atual	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input checked="" type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre	<input type="checkbox"/> 6º semestre
Qualificação	<input type="checkbox"/> Qualificação já realizada em: ____/____/____		<input checked="" type="checkbox"/> Qualificação planejada para: 11/10/2016		
Defesa	Prazo máximo para depósito: 02/02/2018		Depósito planejado para: 30/06/2017		
Linha e Área de pesquisa	<input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC		<input checked="" type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico		
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos <input type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Economia	<input type="checkbox"/> Educação <input type="checkbox"/> Educação a Distância <input checked="" type="checkbox"/> Internet <input type="checkbox"/> Jogos	<input type="checkbox"/> Linguagem Natural <input type="checkbox"/> Linguística <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Química	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input type="checkbox"/> [outro – escrever]	
Publicações	--				

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	Este trabalho é motivado pela possibilidade de melhoria na aplicação de técnicas de recomendação e pela expectativa de aumentar a captação e fidelização de leitores de notícias sob a perspectiva de portais de conteúdo.
Problema de pesquisa	Os recomendadores de notícias de diferentes portais de conteúdo recomendam diversas notícias irrelevantes ao leitor, limitação essa, causada principalmente pela baixa eficácia das técnicas de recomendação adotadas e das medidas de similaridade usadas nos algoritmos de recomendação.
Objetivo geral da pesquisa	O objetivo desta pesquisa é implementar uma estratégia de recomendação baseada em casos que considere diferentes aspectos de uma notícia dentro da medida de similaridade e da técnica de recuperação de casos, melhorando a qualidade das recomendações em comparação àquelas fornecidas por portais <i>baseline</i> .
Trabalhos relacionados	<i>Case-Based Recommendation</i> (Barry Smyth), <i>A Hybrid Similarity Concept for Browsing Semi-structured Product Items</i> (Markus Zanker), <i>Evaluating Recommendation Systems</i> (Guy Shani). Os trabalhos relacionados contribuem para a fundamentação teórica sobre recomendação baseadas em casos, fornecem conceitos de similaridade híbrida em recomendação e formas de avaliação em sistemas de recomendação. Há duas limitações associadas às referências: discutem os conceitos de forma isolada, sem combiná-los, como é o objetivo deste projeto; usam domínios de dados diferentes (produtos) do objetivado neste projeto (notícias).
Justificativa e relevância	Recomendar atendendo adequadamente às expectativas de um usuário sob diferentes perspectivas é um problema em aberto no qual a combinação de diferentes técnicas de recomendação e o uso de similaridade baseada em aspectos híbridos ainda precisa ser avaliada. E, se bem-sucedida, a solução apresentada deve despertar o interesse de portais de conteúdo, potencializando a relação com <i>stakeholders</i> , a captação de leitores, a fidelização de usuários, a escolha de notícias patrocinadas mais relevantes e a disponibilização de outras informações úteis.
Proposta para Solução	Combinar meta-informações sobre a notícias, dados advindos da interação com o usuário e resultados de recuperação de notícias obtidos a partir de análise de seu conteúdo (mineração de textos), em uma medida de similaridade híbrida para recuperação de notícias modeladas em uma base de casos.
Dados	O projeto prevê a coleta de notícias dos portais G1, Estadão e IG, para serem usadas pelas rotinas de mineração, recomendação, experimentos com usuários e avaliação dos resultados. Alguns dados foram obtidos já em formato estruturado, e outros em estado bruto, os quais foram tratados para o repositório de dados do projeto.
Validação	A solução será avaliada comparando a acurácia da recomendação proposta com recomendações aleatórias e recomendações realizadas por um portal <i>baseline</i> . Além disso, a qualidade da recomendação também deverá ser realizada a partir de indicadores coletados a partir de interações com usuários.
Limitações	A limitação deste trabalho está relacionada à avaliação extrínseca, a ser feita a partir de interações com usuários, a qual está sujeita à subjetividade inerente ao ser humano, a uma amostra possivelmente pequena de usuários e a problemas de efetividade de colaboração por parte dos usuários.

