

V Workshop de Dissertações de Mestrado do PPgSI (2018)

FICHA DE PESQUISA

DADOS GERAIS				
Título do projeto de pesquisa	Ensemble de agrupamentos para sistemas de recomendação baseados em conteúdo			
Orientando	Fernando Henrique da Silva Costa			
Orientador(es)	Profa. Dra. Sarajane Marques Peres			
Semestre no curso, na data do workshop	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input type="checkbox"/> 4º semestre	<input checked="" type="checkbox"/> 5º semestre
Qualificação	<input checked="" type="checkbox"/> Qualificação já realizada em: 27/06/2017 <input type="checkbox"/> Realização da qualificação planejada para: dd/mm/aaaa			
Defesa	Prazo máximo para depósito: 29/08/2018 Realização da defesa planejada para: 10/10/2018			
Linha e Área de pesquisa	Gestão e Desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Engenharia de Software <input type="checkbox"/> Gestão de TI <input type="checkbox"/> IHC		Inteligência de Sistemas: <input checked="" type="checkbox"/> IA <input type="checkbox"/> Processamento Gráfico <input type="checkbox"/> Reconhecimento de Padrões	
Área de aplicação	<input type="checkbox"/> Ambientes corporativos <input type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Dispositivos móveis <input type="checkbox"/> Educação	<input type="checkbox"/> Educação a distância <input type="checkbox"/> Governo eletrônico <input type="checkbox"/> Internet <input type="checkbox"/> Jogos <input type="checkbox"/> Jogos sérios	<input type="checkbox"/> Língua Natural <input type="checkbox"/> Linguística <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Quimioinformática	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input checked="" type="checkbox"/> Sistemas de recomendação
Publicações associadas ao projeto de mestrado	<p>Diaz, A.K.R.; Lima, A.P.; Silva, A.M.; Costa, F.H.S.; Pagnossim, J.L.M. & Peres, S.M., Uma análise comparativa das ferramentas de pré-processamento de dados textuais: NLTK, PreText e R. Disponível em: http://ppgsi.each.usp.br/arquivos/RelTec/PPgSI-001_2018.pdf (publicado)</p> <p>Costa, F.H.S. & Peres, S.M., Ensemble de agrupamentos para recomendações serendipitosas baseadas em conteúdo. X Workshop de Teses e Dissertações em Sistemas de Informação, 2017. (publicado)</p> <p>Costa, F.H.S.; Silva, A.M. & Peres, S.M., Relevância, diversidade e serendipidade na recomendação de conteúdo usando agrupamentos. XV Encontro Nacional de Inteligência Artificial e Computacional, 2018. (submetido)</p>			

DESCRIÇÃO DO PROJETO DE PESQUISA

Contextualização / motivação	A pesquisa é referente a sistemas de recomendação baseados em conteúdo (SRbCs), os quais recomendam itens semelhantes aos itens apreciados pelo usuário. Por consequência, tais sistemas tornam-se bastante especialistas e recomendam itens com pouca surpresa ou novidade para o usuário.
Problema de pesquisa	Melhorar a qualidade do aspecto de serendipidade em sistemas de recomendação baseadas em conteúdo, uma vez que esses sistemas sofrem do problema de baixa serendipidade e superespecialização de recomendação.
Objetivo geral da pesquisa	Uma aplicação de estratégia de recomendação que seja capaz de recomendar itens serendipitosos por meio de agrupamento e similaridades parciais.
Trabalhos relacionados	Serendipidade em SRs: JENDERS et al. (2015), IAQUINTA et al. (2010), PIAO; WHITTLE (2011). Agrupamentos em SRs: LIAO; LEE (2016), GONG (2010), RAUSCH et al. (2014). <i>Ensemble</i> de agrupamentos em dados textuais: CÔRREA et al. (2015), WAHID; GAO; ANDREAE (2015)
Proposta para solução	Por meio da utilização de <i>ensemble</i> de agrupamentos, encontrar uma partição que seja superior aos algoritmos de agrupamentos isolados e, por meio dela, utilizar a estratégia de similaridades parciais encontradas nos dados com o propósito de construir uma base de recomendação, a qual seja possível recomendar itens serendipitosos.
Dados	Uma base de dados textuais referentes a 1157 notícias provenientes do portal de notícias EBC. Outra fonte de dados é o conjunto de respostas obtidas em um experimento com usuários para a avaliação de relevância, surpresa e serendipidade em recomendações.
Validação	Para qualidade de agrupamentos, índices de validação baseados em critérios externos e internos foram aplicados. Para qualidade das recomendações produzidas pela abordagem, um experimento com usuários foi realizado.
Limitações	A primeira limitação está relacionada a não existência de uma avaliação quantitativa relacionadas aos aspectos de recomendação. Uma outra limitação, relacionada ao experimento com usuários, é que o usuário não determina qual notícia ou qual caderno de notícias gostaria de ler, podendo ocasionar desinteresse pelo experimento.

