

Universidade Federal de Pernambuco
Graduação em Ciência da Computação
Centro de Informática

Analise de Sentimento em Redes Sociais
Utilizando Influência das Palavras

Proposta de Trabalho de Graduação

Autor: Emanuel de Barros Albuquerque Ferreira
Orientador: Sérgio Ricardo de Melo Queiroz

31 de Agosto de 2010

Sumário

<u>SUMÁRIO.....</u>	<u>2</u>
<u>1. CONTEXTO.....</u>	<u>3</u>
<u>2. OBJETIVO.....</u>	<u>4</u>
<u>3. CRONOGRAMA.....</u>	<u>5</u>
<u>REFERÊNCIAS.....</u>	<u>6</u>
<u>ASSINATURA.....</u>	<u>8</u>

1. Contexto

Muitas opiniões são diariamente expressas na internet em revisões de produtos, *blogs* pessoais, redes sociais, etc. Cada vez mais pessoas têm participado deste cenário deixando suas opiniões *online*. Por isso, muitas pesquisas interessantes e desafiadoras têm surgido, gerando tópicos como detecção de subjetividade, classificação da orientação semântica (positiva ou negativa), e classificação do documento em questão.

Além disso, muitas dessas pesquisas facilmente têm se adequado a aplicações práticas em diversos setores como detecção de opiniões sobre conteúdo políticos [4], extração automática de prós e contras em documentos de revisões *online* [5] e até detecção de potenciais grupos radicais na *web* [2].

Existem muitas tarefas de classificação de polaridade de sentimentos. Três características importantes de tais tarefas são identificadas como: as classes, os níveis de classificação e as assunções sobre a fonte e destino do sentimento. Destacando as classes, existem diversos problemas, os com duas classes comumente envolvem classificar sentimentos como: positivos ou negativos. Esta polaridade de sentimentos pode ser conduzida em níveis de documentos, sentenças ou frases [1].

Em muitas técnicas desenvolvidas para avaliação da polaridade dos sentimentos de documentos ou sentenças, são levadas em consideração características sintáticas e estilísticas dos mesmos. Apesar de bons resultados já poderem ser obtidos em diversas dessas técnicas, poucas utilizam características estilísticas. Em estudos desenvolvidos [1], constatou-se que tais características, ao serem combinadas às sintáticas, revela melhores resultados que quando utilizadas apenas uma das duas.

2. Objetivo

Este trabalho de graduação tem como objetivo a investigação e o desenvolvimento de possíveis otimizações ao *framework* desenvolvido por ABBASI, A., et. al., porém aplicado a documentos extraídos por redes sociais (comentários em tópicos de comunidades do *Orkut* e *tweet* do *Twitter*) e sobretudo em língua portuguesa. A escolha do *framework* de ABBASI et al. dá-se pela sua flexibilidade e independência de língua.

Para tais objetivos, serão realizados estudos sobre o estado da arte da análise de sentimentos, bem como aplicações em redes sociais.

Uma otimização planejada consiste na utilização de um conjunto de palavras com um grau semântico associado a elas, obtidos por BRADLEY [3], no domínio da psicologia. Esse conjunto de palavras será utilizado para gerar novas características para o classificador desenvolvido nesse trabalho. No entanto, como o dicionário de palavras com grau semântico associado foi obtido para o inglês, serão experimentadas alternativas com tradução automática das palavras, com ou sem o uso de sinônimos, utilizando-se a *Wordnet* [9]. Também será avaliado o uso da *SentiWordNet* [8] para otimização dos resultados.

Para os testes, os documentos extraídos das redes sociais serão manualmente classificados. E, ao final deste trabalho, apresentando um comparativo entre as soluções obtidas pelo *framework* original e o mesmo com as otimizações desenvolvidas.

3. Cronograma

O cronograma abaixo apresenta alguma das datas para as principais atividades do processo de desenvolvimento do trabalho de graduação, compreendendo o período entre 04/08/2010 a 30/11/2010.

Atividade	Ago	Set	Out	Nov
Pesquisa e levantamento bibliográfico.	■			
Implementação do Framework baseado.		■		
Implementação das Otimizações.		■		
Recolhimento de documentos das redes sociais.		■		
Análise das soluções.			■	
Escrita do Relatório.		■	■	
Elaboração da Apresentação.				■

Referências

- [1] ABBASI, Ahmed., et. al. **Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums**. ACM Transactions on Information Systems, Vol. 26, No. 3, Article 12. 2008.
- [2] BERMINGHAM, Adam., et al. **Combining Social Network Analysis and Sentiment Analysis to Explore the Potential for Online Radicalisation..** Science Foundation Ireland. 2008.
- [3] BRADLEY, Margaret M.; LANG, Peter J. **Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings**. NIMH Center for the Study of Emotion and Attention. University of Florida.1999
- [4] GRYC, Wojciech.; MOILANEN, Karo. **Leveraging Textual Sentiment Analysis with Social Network Modelling: Sentiment Analysis of Political Blogs in the 2008 U.S. Presidential Election**. T2PP Workshop. 2010.
- [5] KIM, Soo-Min; HOVY, Eduard. Automatic Identification of Pro and Con Reasons in Online Reviews. In **Proceedings of the COLIG/ACL Main Conference Poster Sessions**. 2006 Ages 483-490.
- [6] LIU, Bing. Sentiment Analysis and Subjectivity. In. **Handbook of Natural Language Processing**. Second Edition. 2010.
- [7] SENTIWORDNET. Disponível em: <<http://sentiwordnet.isti.cnr.it/>>. Acesso em Agosto 2010.
- [8] SU, Fangzhong.; MARKERT, Katja. **From Words to Senses: A Case Study of Subjectivity Recognition. Creative Commons Attribution-Noncommercial-Share Alike 3.0**. 2008.
- [9] WORDNET. Princeton University. Disponível em: <http://wordnet.princeton.edu/>. Acesso em Agosto 2010.

Assinatura

Sérgio Ricardo de Melo Queiroz

Orientador

Emanuel de Barros Albuquerque Ferreira

Aluno

31 de Agosto 2010