

## Análise de sentimentos sobre temas de saúde em mídia social

Sentiment analysis on health issues in social media

Análisis de los sentimientos acerca de los problemas de salud en los medios sociales

Gabriela Denise de Araujo<sup>1</sup>, Fernando Sequeira Sousa<sup>2</sup>, Fabio Teixeira<sup>2</sup>, Felipe Mancini<sup>3</sup>, Edvane Birelo Lopes De Domenico<sup>4</sup>, Marcelo de Paiva Guimarães<sup>5</sup>, Ivan Torres Pisa<sup>6</sup>

### RESUMO

**Descritores:** Análise de sentimento; Doenças; Pacientes; Saúde

**Objetivo:** Identificar na literatura científica estudos sobre análise de sentimento por meio do tratamento computacional de opiniões, sentimentos e subjetividades em textos relacionados à saúde provenientes de mídias sociais. **Métodos:** Pesquisa de revisão narrativa de literatura com artigos publicados entre 2009-2011, nas bases de dados PubMed, ISI, ACM e IEEE. **Resultados:** Foram selecionados cinco artigos, segundo os critérios de inclusão. Há resultados positivos referentes à utilização da análise de sentimento como uma técnica auxiliar ou principal processo para diferentes temas. Embora o sentimento seja altamente variável e dependente do contexto, verificou-se que o método disponibiliza um conjunto de técnicas para extrair essa avaliação e lidar com os conteúdos subjetivos. **Conclusão:** Análise de sentimento tem sido pouco explorada em temas relacionados à saúde. Entretanto, investigações descrevem as potencialidades da aplicação destas técnicas na área da saúde como promissora para prover benefícios aos consumidores, serviços de saúde e governos.

### ABSTRACT

**Keywords:** Sentiment analysis; Disease; Patients; Health

**Objective:** Identify in the scientific literature studies on sentiment analysis using the computational analysis of opinions, sentiments and subjectivity in health texts presented in social media. **Methods:** A narrative literature review research of articles published between 2009-2011 in databases PubMed, ISI, ACM and IEEE. **Results:** Five articles were selected that met the inclusion criteria. All articles presented positive results regarding the use of sentiment analysis as a technical aid process for the different topics. Although the sentiment is highly variable and dependent on context, this method has a set of techniques to extract this assessment and dealing with subjectivity. **Conclusions:** Sentiment analysis has been underexplored in health issues. However investigations describe the potential of these techniques in healthcare as promising to provide benefits to consumers, health services and governments.

### RESUMEN

**Descriptores:** Análisis de sentimientos; Enfermedad; Pacientes; Salud

**Objetivo:** Identificar en la literatura científica, los estudios sobre análisis de los sentimientos con el tratamiento computacional de opiniones, sentimientos y la subjetividad en textos relacionados con la salud creada en los medios sociales. **Métodos:** La investigación consistió en la revisión narrativa de la literatura publicada entre 2009-2011, en la bases de datos PubMed, ISI, ACM y el IEEE. **Resultados:** cinco artículos fueron seleccionados de acuerdo a los criterios de inclusión. El conjunto mostró resultados positivos con respecto a la utilización de análisis de los sentimientos como proceso principal o auxiliar de diversos temas. Aunque los sentimientos son muy variable y depende del contexto, se encontró que el método proporciona una serie de técnicas para la extracción de la evaluación y tratar con los contenidos subjetivos. **Conclusión:** El análisis de los sentimientos ha sido poco explorado en temas de salud. Sin embargo, las investigaciones demuestran que el potencial de estas técnicas en el campo de la salud como una promesa de proporcionar beneficios a los consumidores, servicios de salud y los gobiernos.

<sup>1</sup> Bacharel em Informática Biomédica pela Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo – FFCLRP/FMRP- USP. Programa de Pós-graduação em Gestão e Informática em Saúde – PPGIS, Universidade Federal de São Paulo -UNIFESP, São Paulo, (SP), Brasil.

<sup>2</sup> Mestre em Ciências. Programa de Pós-graduação em Gestão e Informática em Saúde – PPGIS, Universidade Federal de São Paulo -UNIFESP, São Paulo, (SP), Brasil.

<sup>3</sup> Doutor em Ciências - Gestão e Informática em Saúde. Instituto Federal de Educação Ciência e Tecnologia de São Paulo, Guarulhos, (SP), Brasil.

<sup>4</sup> Professora Adjunta do Departamento de Enfermagem Clínica e Cirúrgica, Escola Paulista de Enfermagem, Universidade Federal de São Paulo - EPE/ UNIFESP, São Paulo, (SP), Brasil.

<sup>5</sup> Professor Doutor. Secretaria de Educação a Distância, Universidade Federal de São Paulo -UNIFESP, São Paulo, (SP), Brasil.

<sup>6</sup> Professor Adjunto do Departamento de Informática em Saúde, Universidade Federal de São Paulo -UNIFESP, São Paulo, (SP), Brasil.

## INTRODUÇÃO

Nos últimos anos, o crescimento da quantidade de informação disponível na web modificou a forma como os indivíduos buscam conhecimento. Além dos conteúdos disponibilizados pela mídia, os usuários passaram a compartilhar na web seus conhecimentos, críticas e opiniões em blogs pessoais, redes sociais entre outros. Com isso, o uso da web para compartilhar dados tem gerado possibilidades para análises e mineração de dados<sup>(1)</sup>. Uma das possíveis aplicações da análise deste conteúdo compartilhado é na avaliação de questões na área da saúde, como por exemplo, estabelecer qual o “sentimento” que os usuários apresentam a respeito deste conteúdo, ou seja, a opinião positiva ou negativa expressa sobre o ele. Segundo Bing Liu<sup>(1)</sup>, as opiniões são importantes uma vez que indivíduos e organizações são influenciados pelas mesmas no momento de uma tomada de decisão. No caso das organizações, realizar pesquisas ou grupos de discussão com a finalidade de coletar opiniões dos consumidores sobre seus produtos e os de seus concorrentes é uma atividade cada vez menos comum, uma vez que atualmente existe uma abundância de informações e opinião dos indivíduos à disposição na web. Contudo, antes da existência da web, praticamente não existiam estudos computacionais a respeito de mineração de opinião, já que havia poucos textos sobre opinião de usuários disponíveis. Com essa expansão da participação dos usuários nos conteúdos da web expondo seus pensamentos, novos conceitos e metodologias surgiram para investigação desses conteúdos<sup>(2)</sup>. Assim, a Análise de Sentimento (AS) tornou-se uma área de pesquisa em Processamento de Linguagem Natural (PLN) e mineração de textos. A AS é o estudo computacional de como opiniões, atitudes, emoções e perspectivas são expressadas na linguagem natural<sup>(1)</sup>, podendo ser considerada uma disciplina de estudo multicêntrica, englobando áreas como psicologia, marketing e computação.

A AS pode ser utilizada para identificação e classificação do conteúdo emocional criado pelos usuários nas redes sociais, determinando opiniões positivas, negativas e neutras, fornecendo assim uma polaridade da opinião ou orientação do sentimento<sup>(3)</sup>.

Em uma visão geral, a AS faz o rastreamento de uma grande quantidade de mensagens sobre um tema pré-selecionado obtendo um relatório com a opinião de pessoas sobre este tema<sup>(4)</sup>. Os passos envolvidos nesse processo são coleta de dados, classificação e sumarização. A coleta de dados visa apenas buscar na web conteúdos relacionados ao tema e arquivá-los para análise e classificação. A etapa de classificação pode ser realizada por meio de técnicas de aprendizagem de máquina, seleção de palavras ou análise sintática. E por fim, na sumarização de resultados, as classificações das diversas opiniões devem ser resumidas e sintetizadas, com o intuito de facilitar o seu entendimento sobre as mesmas. Isto pode ser preparado em forma de texto ou gráfico<sup>(2)</sup>.

Existem na literatura exemplos de artigos publicados que aplicaram essa técnica para diversos temas relacionados

à saúde, como estudos para saúde pública (rastreamento de tendências)<sup>(3)</sup>, viabilizar a confiabilidade de reclamações médicas<sup>(5)</sup>, sobre o mundo emocional das comunidades de saúde online examinando a força de vários tipos de emoções nos fóruns de discussões relacionados à saúde<sup>(6)</sup>, auxílio em pesquisas de artigos em bases de dados de saúde<sup>(7)</sup>, visualização social de comentários de saúde em mídias sociais<sup>(8)</sup>, entre outros.

Diante deste contexto, este artigo tem como objetivo apresentar resultados de uma revisão narrativa da literatura sobre técnicas computacionais utilizadas para o processo de AS de conteúdos relacionados à saúde em mídias sociais. Assim, é apresentada uma visão geral da importância e dos resultados que este tipo de estudo e análise de conteúdos em mídias sociais tem mostrado para diversos temas sobre saúde.

## MÉTODOS

Estudo de revisão narrativa de literatura realizado com artigos publicados no período de 2009 a 2011, coletados nas seguintes bases de dados: PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), ACM (<http://dl.acm.org>), ISI (<http://isiknowledge.com>) e IEEE (<http://ieeexplore.ieee.org/Xplore>). Como a busca por artigos científicos foi específica para pesquisas que fizessem o uso da análise de sentimento foi utilizado o nome da técnica *sentiment analysis* como palavra-chave relacionando-a com os seguintes descritores de saúde: *diseases*, *medicine*, *patients* e *health*, de forma singular e associada. Todos esses termos de saúde fazem parte do vocabulário Medical Subject Headings (MeSH). Para a base PubMed a busca foi realizada apenas utilizando o termo *sentiment analysis*, não associando-o com outros descritores, por se tratar de uma base sobre saúde. Já na ISI e na IEEEExplore foi relacionado o termo *sentiment analysis* com o *health* para a pesquisa. Por fim, na base ACM a busca relacionou todos os descritores, associando-os de forma que o termo principal, *sentiment analysis*, fosse combinado aos descritores de saúde *diseases*, *medicine*, *patients* e *health*.

Após executadas as buscas com a combinação dos termos, os resumos dos artigos retornados pelos buscadores foram lidos e analisados, segundo os seguintes critérios de inclusão: ter sido publicado no período 2009-2011; utilizar a técnica de análise de sentimento e estar em língua inglesa e portuguesa. Os critérios de exclusão considerados foram: excluir artigo que não apresentasse a versão completa para leitura; excluir artigo não relacionava o conceito à saúde.

## RESULTADOS

As buscas retornaram um total de 25 artigos. Desse total, apenas cinco foram selecionados conforme os critérios de inclusão e exclusão para a leitura completa. Nenhum artigo em português foi localizado para análise. Os artigos encontrados e selecionados para a leitura completa estão sistematizados no Quadro 1.

A seguir serão relatados os métodos e os resultados dos cinco artigos selecionados para o estudo.

Em 2009, um estudo<sup>(8)</sup> de temas de visualização social de sites com informações de pacientes para fornecer uma melhor exploração e uso do site foi realizado. Visualização social é uma forma de descrever os nossos ambientes online e gerar padrões de interação e salientes conexões<sup>(9)</sup>. Neste trabalho foi abordada uma série de visualizações de três temas como “clustering”, “história” e “sentimento” para extrair e preparar o dado. Apesar do tema saúde em redes sociais estar pouco presente, existem fóruns e comunidades online para reunir pessoas e discussões com foco em saúde. No entanto, essas comunidades podem ser difíceis de interagir, com isso o estudo fornece uma idéia para gerar um ambiente de tal forma que o público entenda com maior facilidade como participar neste espaço virtual, reconheça as pessoas envolvidas, quais tópicos abordados, entre outros. As visualizações apresentadas foram criadas a partir de dados de pacientes derivados do site CHFpatients (<http://www.chfpatients.com/>). Com o “clustering” como tema de visualização, o foco foi no agrupamento de informações similares, agrupando pessoas com medicações similares, ou apenas medicamentos. Essa visualização mostrou a informação completa do fórum para todo o seu período de existência. No tema “história”, a visualização pode ser feita em um determinado intervalo de tempo. Os internautas podem explorar discussões sobre medicamentos a partir do momento que a discussão se inicia.

Porém, o fórum do CHFpatients não discute apenas sintomas e drogas; há comentários sobre história pessoal, satisfação do atendimento e preocupações emotivas. Sendo assim, utilizou-se o terceiro tema para avaliar o sentimento agregado no fórum. A análise de sentimento foi utilizada para auxílio na determinação do gênero predominante nos comentários, que podem ser queixas,

busca de informações, mensagens de apoio e otimismo, e também avaliar a atitude dos usuários durante um período de tempo. Ainda, o estudo avaliou se as pessoas apresentam opiniões positivas ou negativas sobre uma droga com base em seus efeitos colaterais. Essa informação é útil para prever e melhorar condições de tratamento. Cada um dos temas discutidos de visualização tem vantagens e limitações. Contudo a pesquisa de Chee, Karrie G. Karahalios, Bruce Schatz argumenta que analisar comentários de saúde irá simular o que pode acontecer com o paciente, ou seja, pessoas podem gerir sua saúde baseadas em experiências reais de outras em situações semelhantes.

Ainda em 2009, outro estudo propôs uma abordagem que faz o uso de um passo adicional de pré-avaliação em sistemas de apoio a decisão que utilizam técnicas de aprendizado de máquinas simbólico para dar suporte a investigação de conhecimentos na área da saúde<sup>(7)</sup>. O sistema semi-automático proposto coleta um conjunto de dados a partir de repositórios. Posteriormente utiliza diferentes métodos de análise de dados como: mineração de dados, árvores de decisão, regras de associação, para gerar um conjunto de regras e padrões para análise dos dados. Considerando apenas estas etapas, um conjunto enorme de regras pode ser gerado, dificultando a avaliação feita por um especialista. Portanto, um passo de pré-avaliação foi desenvolvido. A pré-avaliação faz o uso de técnicas de mineração de textos e análise de sentimento para processar as regras já extraídas juntamente com outros documentos encontrados em repositórios online, fornecendo novos padrões mais concisos e um apoio maior ao especialista na avaliação dos documentos.

Já em 2010, foi realizado um estudo que faz a análise de conteúdo de mensagens divulgadas no Twitter, denominado *tweets*, durante o surto de H1N1 em 2009<sup>(3)</sup>.

**Quadro 1** - Artigos encontrados nas bases de dados PubMed, ISI, ACM e IEEE sobre Análise de Sentimento, seus objetivos e resultados.

Artigo/Veículo de Publicação	Base de Dados	Objetivo	Resultados
CHEE et al., 2009. Social Visualization of Health Messages. /System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on	IEEE	Estudar temas de visualização social de sites com informações de pacientes para uma melhor exploração e uso do site.	Favoreceram o desenvolvimento de ambientes online padronizados, com o objetivo de auxiliar os usuários na busca por informações de saúde.
ZORMAN; VERLIC, 2009. Explanatory approach for evaluation of machine learning-induced knowledge. / The Journal of International Medical Research	PubMed	Criar uma etapa de pré-avaliação para oferecer suporte à investigação de conhecimentos na área da saúde.	A pré-avaliação demonstrou ser útil para processar as regras extraídas, fornecendo novos padrões mais concisos e maior apoio ao especialista na avaliação dos documentos.
CHEW, EYSENBACH, 2010. Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. / PLoS ONE	ISI	Analisar conteúdos de mensagens divulgadas no Twitter, durante o surto de H1N1 em 2009.	Verificou que <i>tweets</i> podem ser usados para análise de conteúdos em tempo real e fontes de opiniões e experiências. Fornecendo assim um auxílio as autoridades de saúde pública a captarem com maior facilidade às preocupações do público.
VYDISWARAN et al., 2011. Gauging the Internet Doctor: Ranking Medical Claims based on Community Knowledge. / Workshop on Data Mining for Medicine and HealthCare	ACM	Verificar a viabilidade de avaliar automaticamente a confiabilidade de uma reclamação médica baseada no conhecimento da comunidade.	Ampliou a noção de fidelidade a um site a partir de uma pontuação de confiança. Mostraram que é viável pontuar a confiabilidade de reclamações baseadas nas opiniões expressadas pelos usuários em fóruns.
YU, 2011. The emotional world of health online communities. / iConference '11	PubMed	Medir o tipo de emoção expressada por membros de comunidades online de saúde.	A análise auxiliou a compreensão das comunicações emocionais nas comunidades de saúde, definindo os participantes e comparando seus status emocionais.

Este estudo teve como objetivo monitorar o uso dos termos relacionados à H1N1 e gripe suína ao longo do tempo. Realizaram uma análise de conteúdo dos *tweets* e por fim, houve uma investigação do uso do Twitter como uma ferramenta de apoio ao rastreamento de tendências e epidemias em tempo real. Foram arquivados mais de dois milhões de mensagens do Twitter que continham palavras-chaves como “swineflu” e “H1N1” entre o dia primeiro de maio e 31 de dezembro de 2009. Vários picos de atividade do Twitter coincidiram com notícias veiculadas na web mais importantes sobre o tema. Os resultados apresentaram alta correlação com os dados de incidência H1N1. O estudo mostrou que os *tweets* relacionados à H1N1 foram utilizados principalmente para disseminar informação de fontes credíveis, mas também foram fontes de opiniões e experiências. Verificou-se que esses *tweets* podem ser usados para análise em tempo real de conteúdo, auxiliando as autoridades de saúde pública a captarem com maior facilidade às preocupações do público.

Em 2011, um estudo para medir o tipo de emoção expressada por membros de comunidades online de saúde foi realizado<sup>(6)</sup>. No estudo foi possível encontrar padrões em expressões dos membros de diferentes perfis, como: questionadores, respondentes, homem, mulher, médico, paciente, cuidadores, entre outros, utilizando técnicas de processamento de linguagem natural e de análise de sentimento. No trabalho foram coletados comentários de usuários em três fóruns: grupo de apoio ao câncer de mama, grupo de apoio ao câncer de próstata e o grupo de tratamento de câncer. Em seguida, o programa de análise de texto *linguistic inquiry and word count* (LIWC, <http://www.liwc.net>), construído por psicólogos para medir a emoção de expressões em textos, foi utilizado para analisar os tipos de sentimentos revelados. Outra ferramenta, Open NLP (<http://openlp.sourceforge.net>), foi utilizada para extrair pronomes e substantivos nos comentários com a finalidade de prever quem fez o comentário. O estudo de Yu relatou que a identificação precisa do autor dos comentários é um desafio assim como medir o impacto do perfil identificado na expressão da emoção. Porém, a pesquisa conseguiu realizar uma análise que melhorou a compreensão das comunicações emocionais nas comunidades de saúde, definindo os participantes e comparando seus status emocionais.

Ainda em 2011, foi publicado um estudo da viabilidade de avaliar automaticamente a confiabilidade de uma reclamação médica baseada no conhecimento da comunidade<sup>(5)</sup>. A proposta baseou-se em gerar uma pontuação (*scores*) de confiabilidade sobre a opinião de usuários expressa em comunidades online. Assim, as reclamações publicadas foram usadas para classificar novas reclamações quanto à confiabilidade depois de concebido a pontuação. Com isso, pôde-se ampliar a noção de fidelidade a um site a partir de uma pontuação de confiança. Os autores mostraram que é viável pontuar a confiabilidade de reclamações baseadas nas opiniões expressadas pelos usuários em fóruns e que pela formulação da pesquisa como uma consulta estruturada, foram capazes de coletar resultados mais representativos

que por sua vez ajudaram a calcular melhor o apoio para uma reclamação médica.

## DISCUSSÃO

Um dos desafios da área da inteligência artificial, em especial na subárea de processamento de linguagem natural, é tratar computacionalmente opiniões, sentimentos e a subjetividade em texto. A análise de sentimento é uma das técnicas que vem ganhando força cada vez mais neste tratamento de reconhecimento da afetividade inferida a partir da escrita<sup>(2)</sup>. Tratar opiniões e sentimentos relacionados à saúde é mais um grande desafio. Atualmente, buscar informações sobre saúde na internet está se tornando comum. Pacientes e familiares usam a internet para compreender o processo saúde-doença, conhecer abordagens terapêuticas, sanar dúvidas sobre o uso de medicamentos, como também, para discutir os sintomas e as preocupações com outros pacientes em fóruns e redes sociais.

Nos artigos encontrados e descritos nesta pesquisa foi possível notar os diversos temas de saúde em que a técnica de AS foi inserida para o tratamento das opiniões, obtendo-se os resultados esperados para cada estudo. Alguns trabalhos apresentam um foco maior em redes sociais<sup>(3,8)</sup>. Entre esses, um sistema de rastreamento de comentários sobre a gripe suína apenas utilizando a rede Twitter foi desenvolvido<sup>(3)</sup>, armazenando em um banco de dados seus resultados e após isso fazem consultas estruturadas para classificar o sentimento nos comentários. O outro trabalho relata que havia pouca informação disponível sobre o uso de redes sociais para obter informações de saúde na época, sendo assim, utilizou um fórum de discussão específico sobre a doença *congestive heart failure* (CHF) para criar seus visualizadores sociais<sup>(8)</sup>. O estudo que fez uso de um fórum de discussão mostrou que a extração dos dados foi facilitada pela criação de estruturas de dados baseadas em disco e algoritmos eficientes que podem se adaptar a muitas mensagens. Depois de feita a coleta dos dados, algumas modificações foram realizadas para indexar palavras (ngramas) para então ser realizada a classificação do sentimento. Após esta classificação mostraram a possibilidade de desenvolver ambientes online padronizados de forma auxiliar os usuários na busca pelas informações de saúde. Outro artigo também mantém o foco sobre as comunidades online de saúde, porém não conduz a investigação em apenas um fórum específico de uma doença<sup>(6)</sup>. O estudo valeu-se de três fóruns de saúde relacionados ao câncer de mama e ao de próstata, e o diferencial deste estudo foi o fato de agrupar as sentenças pelos papéis dos autores (médicos, paciente, cuidadores, questionadores) e fazer uma comparação da força da emoção dos diferentes grupos.

Além desses um artigo trabalhou com a questão da confiabilidade das reclamações médicas<sup>(5)</sup>. Analisar as informações e distinguir o que é confiável e o que não é, pode ser uma tarefa complexa já que comentários podem ser subjetivos e variáveis muito dependentes do contexto inserido, porém este trabalho conseguiu mostrar que é

viável pontuar a confiabilidade de reclamações baseadas nas opiniões expressadas pelos usuários.

Por fim, diferentemente de todos os assuntos e propostas discutidas nos artigos descritos acima, um estudo apresentou uma abordagem para auxiliar especialistas na avaliação de documentos relacionados à saúde arquivados em repositórios online<sup>(7)</sup>. Essa abordagem presa em adicionar um novo passo de pré-avaliação utilizando técnicas de AS e mineração de texto para filtrar ainda mais as regras e padrões desenvolvidos.

Por fim, diferentemente de todos os assuntos e propostas discutidas nos artigos descritos acima, o estudo<sup>(7)</sup> apresentou uma abordagem para auxiliar especialistas na avaliação de documentos relacionados à saúde arquivados em repositórios online. Essa abordagem presa em adicionar um novo passo de pré-avaliação utilizando técnicas de AS e mineração de texto para filtrar ainda mais as regras e padrões desenvolvidos.

Além disso, técnicas de AS também podem ser aplicadas em um buscador específico da área da saúde<sup>(10)</sup> para auxiliar os usuários em suas buscas na web, informando se um texto encontrado que provém de uma mídia social contém informações positivas, negativas ou neutras, melhorando a experiência do usuário em suas buscas na web<sup>(11)</sup>.

A partir da análise destas publicações pôde-se constatar

que obter opiniões em mídias sociais e monitorá-las ainda é uma tarefa árdua devido ao grande número de sites existentes (redes sociais, blogs, fóruns) e estes possuem um grande volume de textos opinativos que podem ainda conter opiniões veladas em longos comentários, com elevado conteúdo de subjetividade, entre outros fatores limitantes.

## CONCLUSÃO

Com base nos resultados encontrados na literatura, descreveu-se as potencialidades em aplicar a análise de sentimento na área da saúde, técnica atualmente pouco explorada. Entretanto, os achados científicos apresentados nos artigos selecionados e discutidos sinalizam para os benefícios da aplicação dessa técnica em diferentes setores vinculados à saúde, como serviços especializados, órgãos públicos, bem como para os próprios consumidores.

No campo da saúde coletiva, a mineração de opiniões aplicada aos fóruns dos quais consumidores (cidadãos) participam, pode gerar mapeamentos de ocorrência de doenças, além de conteúdos de diálogos fecundos sobre respostas emocionais humanas e opiniões ante a diferentes informações veiculadas, podendo influenciar significativamente a prospecção e marketing de temas em saúde coletiva.

## REFERÊNCIAS

1. Liu B. Sentiment analysis: A multi-faceted problem. *IEEE Intell Syst.* 2010;25:1-5.
2. Pang B, Lee L. Opinion mining and sentiment analysis. *Found Trends Inf Retr.* 2008;2(1-2):1-135.
3. Chew C, Eysenbach G. Pandemics in the Age of Twitter: Content analysis of tweets during the 2009 H1N1 Outbreak. *PLoS ONE.* 2010 29;5(11):e14118.
4. Ohana B, Tierney B. Sentiment classification of reviews using SentiWordNet. In: *Proceedings of the 9th IT & T Conference.* 2009 oct 22-3; Dublin, Ireland.
5. Vydiswaran VGV, Zhai CX, Roth D. Gauging the internet doctor: ranking medical claims based on community knowledge. In: *Proceedings of the 2011 workshop on Data mining for medicine and healthcare.* 2011 aug 21-4; San Diego (CA), USA.
6. Yu B. The emotional world of health online communities. *Proceedings of the 2011 iConference [Internet].* New York, NY, USA: ACM; 2011 [cited 2012 feb 15]. p. 806-7. Available from: <http://doi.acm.org/10.1145/1940761.1940914>
7. Zorman M, Verlic M. Explanatory approach for evaluation of machine learning-induced knowledge. *J Int Med Res.* 2009;37(5):1543-51.
8. Chee B, Karahalios KG, Schatz B. Social Visualization of health messages. In: *Proceedings of the Hawaii International Conference on System Sciences.* IEEE Computer Society. Los Alamitos (CA) USA; 2009 jan 5-8. p.1-10.
9. Karahalios KG, Viégas FB. Social visualization: exploring text, audio, and video interaction. *CHI '06 extended abstracts on Human factors in computing systems [Internet].* New York, NY, USA: ACM; 2006 [cited 2012 mar 1]. p. 1667-70. Available from: <http://doi.acm.org/10.1145/1125451.1125758>
10. Mancini F, Sousa FS, Teixeira FO, Falcão AEJ, Hummel AD, da Costa TM, et al. Use of medical subject headings (MeSH) in portuguese for categorizing web-based healthcare content. *J Biomed Inform.* 2010;44(2):299-309.
11. Sousa F, Mancini F, Teixeira F, Falcão A, Hummel A, Nunes F, Sigulem D, Pisa I. Categorização automática de conteúdos web de saúde em português brasileiro com classificador bayesiano. *J. Health Inform.* 2012; 4(1):10-6.