

Análise de sentimento de *tweets* com foco em notícias

Paula Nascimento, Rodrigo Aguas, Débora de Lima, Xiao Kong, Bruno Osiek,
Geraldo Xexéo, Jano de Souza

Programa de Engenharia de Sistemas e Computação
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia
Universidade Federal do Rio de Janeiro – Rio de Janeiro – Brasil

{pcn,rodrigoaguas,dlima,xiao,baosiek,xexeo,jano}@ufrj.br

***Abstract.** The curiosity for knowing what people think and how they feel about daily events has always existed. This study aims to satisfy this need and analyze if people react in a positive or negative way to news reported in the media. To this end, 3 distinct topics were selected and, for each one of them, information published on the microblogging service Twitter was collected, analyzed and had its polarity identified. The experiment was held using language classifiers and, besides verifying how people feel in relation to the selected news, it was able to identify between 3 different linguistic models which one has the best outcome classifying tweets.*

***Resumo.** A curiosidade por saber o que as pessoas pensam e como se sentem em relação aos acontecimentos do dia a dia sempre existiu. Este trabalho tem por objetivo satisfazer essa necessidade e analisar se as pessoas reagem de forma positiva ou negativa em relação às notícias divulgadas na mídia. Para isso, foram selecionados 3 tópicos e, para cada um deles, informações publicadas no serviço de microblogging Twitter foram coletadas, analisadas e tiveram sua polaridade identificada. O experimento realizado utilizou classificadores de linguagem e, além de verificar qual a opinião da população em relação às notícias selecionadas, foi possível identificar dentre 3 modelos linguísticos distintos qual deles obteve melhor resultado ao classificar tweets.*

1. Introdução

O conceito de *microblogging* surgiu em meados de 2006, quando uma nova iniciativa de rede social foi criada com uma proposta simples: permitir aos seus usuários divulgarem o que estão fazendo¹. Nascia então o Twitter, considerado, na época, apenas mais uma rede social. Com ele surgia também uma nova forma de comunicação na Internet, onde as pessoas podiam divulgar qualquer tipo de informação, em tempo real, para todos aqueles ligados à sua rede. No princípio, o limite de 140 caracteres parecia uma simples limitação da plataforma, mas foi determinante para o uso que as pessoas dariam e o modo como elas se expressariam nessa rede. Esse modelo de rede social permitiu que, mais tarde, o Twitter se tornasse uma rede baseada na troca de informações, onde o dado transmitido é a opinião dos usuários².

1 <http://www.140characters.com/2009/01/30/how-twitter-was-born/>

2 <https://twitter.com/about>

Dois fatores influenciaram fortemente para que o Twitter não fosse mais visto apenas como mais uma rede social. O primeiro deles é o alto nível de popularidade atingido por esta ferramenta. Em janeiro de 2011, segundo o próprio Twitter, a plataforma já possuía cerca de 200 milhões de usuários e gerava cerca de 110 milhões de *posts* por dia³. O segundo está relacionado ao uso dado a esta rede pelos seus usuários. Eles não somente a utilizam para divulgar informações sobre si, como previsto inicialmente, mas também a utilizam para compartilhar opiniões e informações sobre fatos e eventos em geral [Naaman and Boase 2010]. A união destas duas características torna o Twitter uma importante fonte de opiniões e sentimentos sobre eventos e acontecimentos, que podem ser analisados e, posteriormente, utilizados em diversas áreas, como política, social e até mesmo marketing empresarial, conforme verificado por [Li and Li 2011]. No estudo realizado, os autores puderam observar que, atualmente, existe uma necessidade de buscar a opinião do cliente e agir em relação a ela, ao invés de apenas reagir.

A importância de avaliar o sentimento das mensagens divulgadas no Twitter, os *tweets*, já foi identificada em diversos trabalhos. Dentre eles está [Kwak et al 2010], cujo objetivo é justamente avaliar o potencial do Twitter como uma rede de informação. Este estudo mostrou que mais de 85% dos *tweets* criados estão relacionados a manchetes de jornal e demais notícias divulgadas na mídia. Além disso, o fato de o Twitter permitir o compartilhamento em tempo real faz com que essa ferramenta permita captar o sentimento do usuário no momento em que ele soube da notícia em questão, o que o motiva a expressar sua emoção antes que outros fatores o influenciem e diminuam a intensidade do sentimento gerado [Sloman et al 2005].

Este trabalho visa avaliar a reação das pessoas em relação às notícias compartilhadas na mídia através da análise de publicações feitas no Twitter. O objetivo é concluir, a partir dos sentimentos expostos nos *tweets*, se a população achou um determinado fato positivo ou negativo. Foram escolhidas três categorias de assunto a serem investigadas e, para cada uma delas, foi verificado se as notícias daquele dado tópico são vistas de modo positivo ou negativo. A análise de sentimento dos *tweets* foi realizada através de experimento supervisionado utilizando classificadores de linguagem. Este trabalho também busca realizar uma comparação entre três diferentes classificadores de linguagem para verificar qual deles se adequa melhor às características das mensagens coletadas através do Twitter e consegue obter melhor resultado ao tratar textos em português brasileiro.

2. Trabalhos relacionados

Segundo [Pang and Lee 2008], descobrir o que as pessoas pensam sempre foi objeto de interesse. Com a popularização de plataformas que fornecem acesso à grande quantidade de dados subjetivos, a tarefa de identificar a polaridade e tentar classificar qual emoção um texto possui passou a ser o foco de diversas pesquisas. Esta área de estudo ficou conhecida como análise de sentimento e, aliado a ela, novos desafios no tratamento de texto foram apresentados. Isso se deve ao fato de que o enfoque está em extrair a opinião expressa em um determinado documento e não apenas descobrir sobre

3 <http://www.forbes.com/sites/oliverchiang/2011/01/19/twitter-hits-nearly-200m-users-110m-tweets-per-day-focuses-on-global-expansion/>

qual assunto ele trata. Esta nova perspectiva faz com que o objetivo da análise dos termos que compõem o texto seja verificar qual sentimento eles representam e decidir se o documento exprime uma opinião positiva, negativa, neutra ou, até mesmo, tentar classificar qual emoção está presente, como raiva, felicidade, tristeza, etc.

Neste contexto, o Twitter se mostrou uma rica fonte de informação devido ao seu formato de *microblogging*. Ao permitir que usuários compartilhem curtas mensagens frequentemente, esta plataforma deu origem a um *corpus* com grande carga subjetiva, porém com novos desafios a serem vencidos, como por exemplo o vocabulário utilizado, de caráter extremamente informal e inconsistente [Brew et al 2011]. Estas características permitem que abreviações e variações na escrita de uma palavra representem um mesmo significado [Jiang et al 2011], o que dificulta o tratamento dos textos coletados. Para auxiliar a tarefa de identificar o sentimento presente em um *tweet*, [Pak and Paroubek 2010; Davidov et al 2010; Li and Li 2011; Zhang et al 2011; Jiang et al 2011] fizeram uso de *emoticons* – ícones ou sequência de caracteres que transmitem o estado emotivo da mensagem que acompanham⁴. Segundo [Li and Li 2011], esta estratégia permite reduzir o esforço necessário para identificar a emoção expressa na publicação, já que cerca de 87% dos *tweets* coletados desta forma possuem o sentimento representado pelo *emoticon* presente no texto. No entanto, esta abordagem desconsidera uma grande quantidade de *tweets* que não carregam *emoticons* em suas frases, limitando os documentos que serão avaliados durante o estudo.

Diversas pesquisas buscando compreender as opiniões que os usuários do Twitter exprimem já foram conduzidas. Em 2009, [Bollen et al 2009] desenvolveu um dos primeiros trabalhos com enfoque de classificar emoções presentes em *tweets*. O objetivo do estudo realizado foi analisar a flutuação de sentimento dessas mensagens e buscar eventos sócio-econômicos e políticos que pudessem estar relacionados às oscilações identificadas. Para isso, os autores utilizaram um instrumento chamado POMS-ex, que classifica os *tweets* em 6 categorias diferentes de humor, a partir da comparação dos termos que compõem a mensagem com termos, definidos previamente, associados a cada uma das categorias. Em 2010, [Pak and Paroubek 2010] utilizou um classificador Naïve-Bayes para categorizar *tweets* em positivo ou negativo, com base em N-gramas e na classificação gramatical de partes do texto. Já [Davidov et al 2010] buscou categorizar os *tweets* utilizando mensagens com *hashtags* – palavras precedidas de # que representam o sentimento e/ou o assunto alvo daquela mensagem – e *emoticons* contidos no texto para treinar o classificador utilizado e categorizar *tweets* entre diversas emoções. Além disso, aspectos como pontuação e as palavras utilizadas também foram consideradas para a extração de sentimento.

Outra vertente da análise do conteúdo gerado no Twitter é a extração de informação a partir do estudo de *posts* publicados nessa rede. [Sakaki et al 2010] e [Achrekar et al 2011] promoveram pesquisas cujo objetivo era detectar acontecimentos de forma mais rápida através da análise das mensagens divulgadas. Nesses casos, a intenção era criar uma ferramenta capaz de detectar terremotos e epidemias de gripe, respectivamente, antes mesmo de tais eventos serem anunciados oficialmente. Já em [Li and Li 2011], o objetivo era utilizar as opiniões compartilhadas no Twitter para auxiliar

4 <http://pt.wikipedia.org/wiki/Emoticons>

empresas a tomarem decisões em suas campanhas de marketing. A ideia foi utilizar a análise de sentimento de *tweets* para tornar possível o acompanhamento da opinião dos clientes em relação a serviços e produtos e permitir que empresas tomassem decisões antes mesmo de os clientes chegarem até elas com reclamações e sugestões.

Dentre os trabalhos citados, podemos verificar que nenhum deles buscou analisar sentimento relacionado a fatos divulgados na mídia. Na verdade, a estratégia sempre foi inversa: analisar os sentimentos e buscar ocorrências que pudessem estar associadas a eles. Este estudo visa preencher esta lacuna e ser um primeiro passo para uma série de trabalhos onde o foco está em analisar sentimentos relacionados a acontecimentos que tiveram grande repercussão e foram amplamente divulgados por jornais e revistas. Vale ressaltar que este trabalho também está entre os poucos que trata *tweets* escritos em português brasileiro.

3. Metodologia

O objetivo deste trabalho é buscar um mecanismo que consiga lidar com as características presentes em textos oriundos de plataformas de microblog, ou seja, textos informais e ruidosos. Para isso, optou-se pelo uso de modelos de linguagem ou modelos N-grama. A escolha pelo uso desses modelos deu-se pela simplicidade e, ao mesmo tempo, poder de processamento desses algoritmos [Russell 2011].

De acordo com [Jurafsky and Martin 2008], estes modelos estatísticos permitem prever a probabilidade de um grupo de palavras aparecerem em uma determinada sequência a partir das N-1 palavras do N-grama – sequência de N palavras. Por funcionar de acordo com um paradigma conhecido por treinamento e teste, este modelo permite utilizar um *corpus* de treinamento que terá a função de ensinar ao classificador quais sequências de palavras estão associadas a uma determinada classificação [Jurafsky and Martin 2008]. Como o objetivo deste trabalho é a extração de sentimento, o conjunto de treinamento possui documentos que mostram ao classificador quais sentenças caracterizam uma opinião cuja classificação é positiva, por exemplo “Dilma fazendo história na ONU. #Orgulho #DilmaDay”, e quais são consideradas negativas, como no caso de “Emissoras de TV, principalmente a Globo, só falam do POP In RIO e ã passaram nem uma notinha da #MarchaContraCorrupcaoRJ Pq será?C liga Brasil”. Este conjunto de treinamento foi selecionado de forma aleatória entre o *corpus* total de *tweets* coletados, sendo a probabilidade de um *tweet* ser escolhido para fazer parte do conjunto de treinamento de 70%.

De acordo com este paradigma, ainda é necessário um outro grupo de documentos, o conjunto de teste [Jurafsky and Martin 2008]. Os documentos que compõem este conjunto são submetidos ao modelo treinado para calcular a probabilidade de cada um deles se encaixar em uma das classificações do problema. Neste trabalho, 30% dos *tweets* coletados fazem parte do *corpus* de teste.

Este processo permite reduzir o trabalho de limpeza dos textos a serem classificados e, ao mesmo tempo, alcança resultados bastante significativos quando se trata do processamento de textos com características similares às das mensagens tratadas neste estudo [Russell 2011].

Como um dos objetivos deste trabalho é também comparar o desempenho de implementações diferentes de modelos N-grama, foram utilizados 3 classificadores distintos. São eles:

- UNIGRAMA: Este classificador segue o modelo N-grama e calcula a probabilidade de cada texto ser positivo ou negativo, analisando cada palavra separadamente.
- OCTOGRAMA: Este classificador segue o modelo N-grama e calcula a probabilidade de cada texto ser positivo ou negativo, analisando sequências de 8 palavras.
- NAIVE_BAYES: Classificador onde os textos são representados como *bag of words*, ou seja, suas posições exatas são ignoradas, e o classificador é montado com base em um modelo probabilístico baseado no teorema de Bayes, assumindo independência entre as variáveis [Jurafsky and Martin 2008]. Também calcula a probabilidade do texto ser positivo ou negativo.

4. Experimento

4.1. Criação do *corpus*

A etapa de criação do *corpus* pode ser analisada em 2 momentos.

4.1.1. Seleção de notícias

O primeiro passo foi a escolha das notícias a serem consideradas no estudo. Este processo foi realizado manualmente e contou com o apoio de uma coluna da seção de Tecnologia e Games do jornal online G1 onde, para cada dia da semana, os *Trending Topics* – principais assuntos falados no Twitter – brasileiros são apresentados, explicados e associados às notícias correspondentes⁵. Após analisar esses dados por um período de cerca de 2 meses, observamos que dentre os assuntos mais populares 3 categorias se destacam: Policial, Política e Entretenimento. Sendo assim, optou-se por classificar *tweets* que estivessem relacionados a esses 3 tópicos com a intenção de maximizar a ocorrência de textos subjetivos, ou seja, que “expressam a avaliação, emoção, julgamento, incertezas, crenças e outras atitudes e afetividades” [Wiebe 1990] sobre o assunto debatido.

Durante a janela de tempo considerada, foram selecionadas notícias que não apenas tiveram grande repercussão entre os usuários do microblog, mas que também foram bastante divulgadas através de jornais e revistas escritos e televisionados. Para cada categoria escolhida, foram selecionadas entre 2 e 3 notícias, apresentadas na tabela abaixo e ordenadas pela data em que foram divulgadas.

⁵ <http://g1.globo.com/tecnologia/noticia/2012/04/assuntos-no-twitter-segunda-feira-23042012.html>

Tabela 1. Notícias selecionadas para o experimento

Notícia	Categoria	Data
Exames confirmam que ator Reynaldo Gianecchini tem câncer	Entretenimento	10/08/2011
Juíza é baleada e morta em Niterói	Policial	12/08/2011
Perícia diz que sobrecarga provocou desabamento de camarote em SP	Entretenimento	21/08/2011
Público protesta contra corrupção e impunidade no centro do Rio	Política	20/09/2011
Dilma defende Estado palestino e exalta papel da mulher na política	Política	21/09/2011
Câmara derruba criação de novo imposto para a saúde	Política	21/09/2011
Rock in Rio	Entretenimento	08/10/2011
Condenado por chefiar tráfico no Rio, Polegar é preso no Paraguai	Policial	19/10/2011

4.1.2. Coleta e filtragem de *tweets*

Após selecionar as notícias a serem consideradas no experimento, o passo seguinte foi coletar *tweets* relacionados a elas e filtrar essas mensagens de forma a obter apenas *tweets* subjetivos para serem utilizados na etapa de classificação automática.

A coleta das mensagens foi realizada através da interface de busca oferecida pelo próprio Twitter⁶. Para cada uma das notícias, palavras-chave foram utilizadas de forma que os *tweets* retornados estivessem relacionados a elas. Por exemplo, para a notícia “Dilma defende Estado palestino e exalta papel da mulher na política” a busca realizada foi “dilma onu” e para a notícia “Juíza é baleada e morta em Niterói” os termos da busca foram “juíza patricia”. Dessa forma, para cada notícia, foram coletados cerca de 400 *tweets* e o fator tempo se mostrou extremamente importante nesta fase do experimento. Quanto mais distante da data de divulgação da notícia o processo de coleta fosse realizado, menos subjetivos e menos relacionados aos fatos eram os textos das mensagens.

Essas observações foram possíveis de serem feitas devido ao passo seguinte do processo de coleta e filtragem de *tweets*. Cada *tweet* foi considerado um documento e, para cada um deles, 3 dos pesquisadores envolvidos neste trabalho emitiram suas opiniões sobre o sentimento presente nas mensagens. Este processo foi realizado para que fosse possível cruzar os votos de cada um dos pesquisadores e eleger qual sentimento classificaria cada *tweet*. O objetivo desta análise foi classificar as mensagens entre positivo e negativo para que depois esta categorização manual fosse utilizada

⁶ <http://twitter.com/#!/search>

como base para a ferramenta de classificação automática. Optou-se, assim como em [Pang et al 2002], por não se considerar *tweets* neutros, já que não foi possível encontrar na literatura um consenso sobre quais seriam as características típicas de textos classificados desta forma. A intenção da filtragem era obter um *corpus* anotado composto apenas por textos que expressassem sentimentos e, após este processo, foi criada uma base com cerca de 850 documentos, divididos em cerca de 50% positivos e 50% negativos.

Vale ressaltar que o processo de classificação manual não considerou apenas os significados isolados das palavras utilizadas nos *tweets*. Casos como ironias e abreviações foram considerados para determinar a polaridade do texto analisado. O objetivo era verificar se os classificadores utilizados conseguiriam classificar corretamente esses documentos.

4.2. Ferramenta de classificação automática

Este experimento contou com a criação de uma ferramenta de classificação automática capaz de estabelecer, para um dado *tweet*, a polaridade da opinião contida no texto analisado, conforme mostrado na Figura 1. Para apoiar o processo de classificação automática, foi utilizada “uma biblioteca para processamento de texto usando linguística computacional”⁷, chamada LingPipe. Escolheu-se usar este recurso na ferramenta construída por viabilizar a criação de classificadores baseados em modelos N-grama e permitir extrair resultados estatísticos das classificações realizadas⁷. Além disso, esta é uma biblioteca bastante estável e utilizada em diversos trabalhos de pesquisa⁸.

7 <http://alias-i.com/lingpipe/>

8 <http://alias-i.com/lingpipe/web/citations.html>

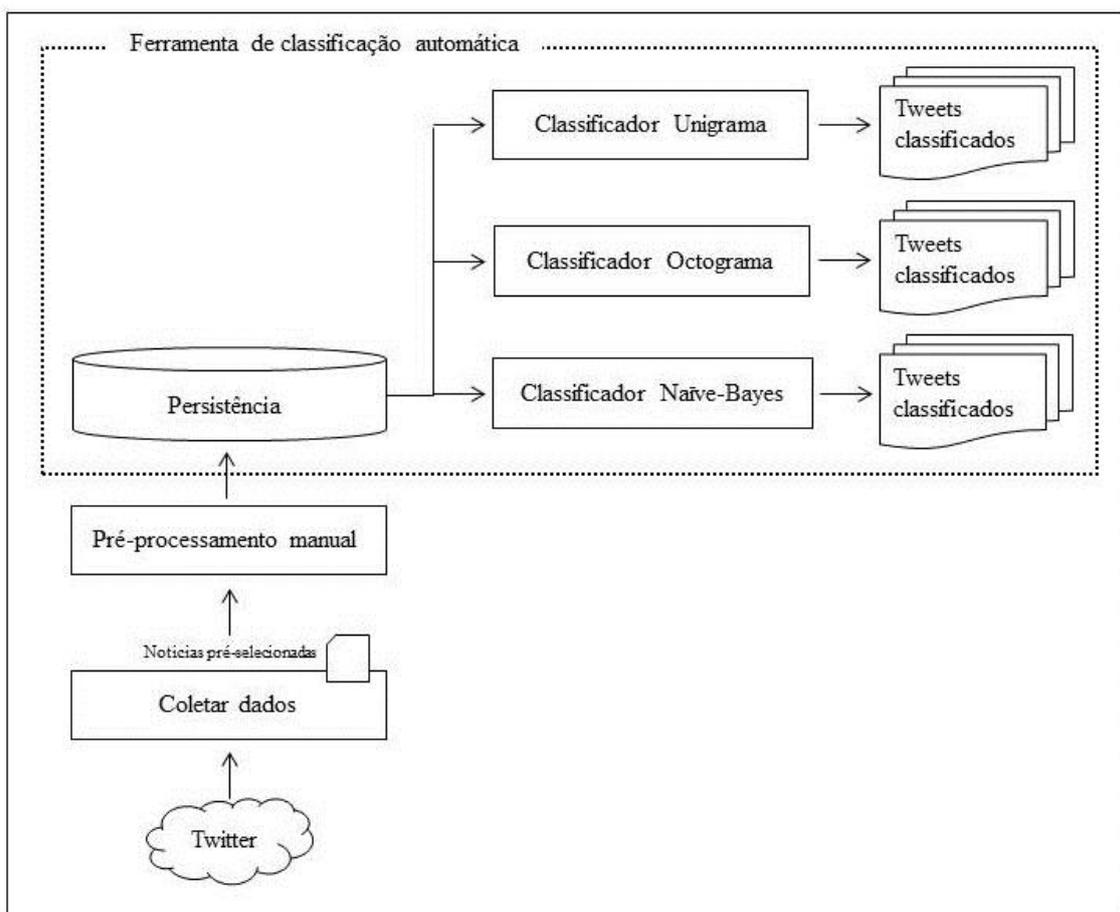


Figura 1. Processo de classificação de tweets

5. Resultados

Nesta seção apresentamos os resultados obtidos após a execução de 10 rodadas de treinamento e testes realizadas com cada um dos classificadores considerados. Como é possível observar na Tabela 2, a diferença entre os resultados obtidos pelos classificadores dentro de uma mesma categoria não justifica a escolha absoluta por um deles, já que a performance alcançada pelos 3 foi muito similar. No entanto, podemos concluir que, para o conjunto de dados utilizados no experimento, o classificador que segue o algoritmo Naïve-Bayes foi o que obteve melhor resultado ao categorizar textos com características de microblog. Outro ponto que podemos observar é que, conforme previsto por [Jurafsky and Martin 2008], ao se trabalhar com N-gramas de tamanho 8, o resultado é melhor do que quando se trabalha com unigramas. Essa conclusão é possível de ser feita através da análise da acurácia dos classificadores para cada uma das categorias de notícia, conforme mostrado na Tabela 2. Esse valor foi calculado de acordo com a seguinte fórmula, também utilizada por [Pak and Paroubek 2010]:

$$\text{Acur} = \frac{\text{Número de classificações corretas}}{\text{Total de classificações}} \quad (1)$$

Tabela 2. Acurácia dos classificadores para cada uma das categorias testadas

	UNIGRAMA		OCTOGRAMA		NAIVE BAYES	
	Acur.	Var.	Acur.	Var.	Acur.	Var.
Entretenimento	0.74	0.0018	0.74	0.0007	0.75	0.0011
Policial	0.76	0.0021	0.77	0.0029	0.81	0.0038
Política	0.68	0.0018	0.71	0.0008	0.69	0.0018

Os resultados obtidos nas categorias Entretenimento e Policial são bastante satisfatórios se considerarmos que a capacidade humana de avaliar corretamente a subjetividade de um texto varia de 72% [Wiebe et al 2006] a 85% [Golden 2011]. Considerando essa margem como o objetivo a ser alcançado com o experimento, podemos verificar que, para as duas categorias citadas, todos os classificadores utilizados obtiveram resultados dentro dessa margem, o que demonstra que o método escolhido é eficaz na classificação de *tweets*.

Já para o caso da categoria Política, acredita-se que a razão para os resultados obtidos estarem abaixo dos 72%, considerado o mínimo esperado de uma classificação humana, está no fato de os textos relacionados a esta categoria possuírem construções mais difíceis de serem analisadas automaticamente, como é o caso da ironia e do sarcasmo, conforme mostrado na Tabela 3. Além disso, algumas dessas mensagens estão bastante ligadas ao contexto a que estão relacionadas, fator que dificulta ainda mais a classificação, tanto automática como manual. Esta última característica do texto não fez parte do objetivo de estudo deste trabalho.

Tabela 3. Casos de classificação incorreta de tweets

Tweet	Classificação manual	Classificação automática
Ufa!! Os parlamentares em Brasilia, não aprovarão mais um Imposto, para pagar o doente sistema de saúde no Brasil	POSITIVO	NEGATIVO
só achei que faltou jogar a caneta na cabeça do obama: "seu incompetente!" #dilmaday	POSITIVO	NEGATIVO
Contra a corrupção:desde o cafezinho do guarda até os mensalões milionários #MarchaContraCorrupcaoRJ	POSITIVO	NEGATIVO
Prestem atenção, pois um novo imposto pode estar a caminho. Os deputados estão pensando em criar uma taxa destinada à saúde! Isso é ABSURDO	NEGATIVO	POSITIVO
Brasil é assim: SE #RedeGlobo tivesse divulgado,Se ela tivesse mostrado a #MarchaContraCorrupcaoRJ teriam 100/200.000 pessoas.Infelizmente!	NEGATIVO	POSITIVO

notícia desejada. Neste estudo, a coleta e classificação de *tweets* foi feita de forma manual, tornando possível assegurar que todos os *tweets* coletados são realmente relacionados à notícia em questão. Futuramente, pretende-se desenvolver um módulo de busca e identificação de tópicos para que a coleta e associação de *tweets* e notícias sejam realizadas automaticamente.

Referências

- Achrekar, H., Gandhe, A., Lazarus, R., Yu, S., Liu, B. (2011). “Predicting Flu Trends using Twitter data”, Proceedings of 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 702 – 707.
- Bollen, J., Pepe, A., and Mao, H. (2009). “Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena.”, Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011), 17-21 July 2011, Barcelona, Spain.
- Brew, A., Greene, D., Archambault, D., and Cunningham, P. (2011). “Deriving Insights from National Happiness Indices.”, 2011 IEEE 11th International Conference On Data Mining Workshops (ICDMW), pp. 53 –60.
- Davidov, D., Tsur, O., and Rappoport, A. (2010). “Enhanced sentiment learning using Twitter hashtags and smileys.”, Proceedings of the 23rd International Conference on Computational Linguistics: Posters, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 241–249.
- Golden, P. (2011). “Write here, write now.”. Disponível em <<http://www.research-live.com/features/write-here-write-now/4005303.article>>.
- Jiang, L., Yu, M., Zhou, M., Liu, X., and Zhao, T. (2011). “Target-dependent Twitter sentiment classification.”, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 151–160.
- Jurafsky, D., and Martin, J. H. (2009). Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition, Prentice Hall, 2nd Edition.
- Kwak, H., Lee, C., Park, H., and Moon, S. (2010). “What is Twitter, a social network or a news media?”, Proceedings of the 19th International Conference on World Wide Web, (New York, NY, USA: ACM), pp. 591–600.
- Li, Y.-M., and Li, T.-Y. (2011). “Deriving Marketing Intelligence over Microblogs.”, Proceedings of 44th Hawaii International Conference On System Sciences (HICSS), pp. 1 –10.
- Naaman, C.-H. L. Mor., and Boase, J. (2010). “Is it all About Me? User Content in Social Awareness Streams”, Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, 2010.

- Pak, A., and Paroubek, P. (2010). "Twitter as a corpus for sentiment analysis and opinion mining.", Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC'10).
- Pang, B., Lee, L., and Vaithyanathan, S. (2002). "Thumbs up?: sentiment classification using machine learning techniques.", Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 79–86.
- Pang, B., and Lee, L. (2008). "Opinion Mining and Sentiment Analysis", Foundation and Trends in Information Retrieval, Vol. 2 (1-2), pp. 1-135.
- Russell, M. A. (2011). Mining the Social Web, O'Reilly Media, Inc, 1st Edition.
- Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). "Earthquake shakes Twitter users: real-time event detection by social sensors.", Proceedings of the 19th International Conference on World Wide Web, (New York, NY, USA: ACM), pp. 851–860.
- Sloman, A., Chrisley, R., Scheutz, M. (2005). "The architectural basis of affective states and processes.", Who Needs Emotions?: The Brain Meets the Machine, v. 3, pp. 203–244.
- Waltinger, U. (2009). "Polarity reinforcement: Sentiment polarity identification by means of social semantics.", AFRICON, 2009, pp. 1 – 6.
- Wiebe, J.M. (1990). "Identifying subjective characters in narrative.", Proceedings of the 13th Conference on Computational Linguistics - Volume 2, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 401–406.
- Wiebe, J., Wilson, T., Cardie, C. (2006). "Annotating Expressions of Opinions and Emotions in Language", Language Resources and Evaluation, v. 39, n. 2-3, pp. 165 – 210.
- Zhang, K., Cheng, Y., Xie, Y., Honbo, D., Agrawal, A., Palsetia, D., Lee, K., Liao, W., Choudhary, A. (2011). "SES: Sentiment Elicitation System for Social Media Data.", Proceedings of 11th International Conference on Data Mining Workshops (ICDMW), pp. 129 – 136.