****

**Autor:** Paulo Henrique S. L. Floro de Oliveira {phslfo@cin.ufpe.br}

**Orientador:** Ricardo Bastos Cavalcante Prudêncio {rbcp@cin.ufpe.br}

Recife, Abril de 2012

**Universidade Federal de Pernambuco**

Graduação em Ciência da Computação

Centro de Informática

2012.1

Proposta de Trabalho de Graduação

**Análise de Sentimentos no Twitter:**

Um abordagem utilizando o Classificador Naive Bayes

Conteúdo

[1. Contexto 2](#_Toc321947429)

[2. Objetivos 3](#_Toc321947430)

[3. Cronograma 4](#_Toc321947431)

[4. Possíveis Avaliadores 5](#_Toc321947432)

[5. Referências Bibliográficas 6](#_Toc321947433)

[6. Assinaturas 7](#_Toc321947434)

1. Contexto

A quantidade de dados presente na web tem crescido desenfreadamente com a popularização de redes sociais como Twitter[[1]](#footnote-1) e Facebook². A quantidade de dados já existente na web a cada dia ultrapassa barreiras históricas e cálculos indicam que até 2015 essa quantidade irá quadruplicar [1].

Além disso, tais redes tem se mostrado bastante acessíveis para disseminação de comentários e opiniões sobre produtos e/ou serviços. Tanto que algumas empresas de TI, como a empresa local E-life³, já atuam exclusivamente na área de monitoração e análise de dados sociais.

Essas informações públicas são usadas como fonte para extração de características e opiniões sobre produtos, empresas e/ou pessoas. Grandes empresas mundiais e personalidades da mídia tem o interesse de acompanhar a sua popularidade e a relevância de sua marca. Empresas de marketing precisam analisar os resultados de campanhas publicitárias. Os próprios consumidores tem o interesse de saber se um determinado produto está sendo bem comentado antes de compra-lo.

Dai que surge o conceito de Análise de Sentimento, que se propõe a identificar características de um texto de acordo com um assunto específico. Analisando e classificando-os em subgrupos pré-definidos, como sentimentos e emoções (feliz, triste), polaridade (positivo, negativo) ou em qualquer outro contexto previamente definido. [2]

Vários métodos e algoritmos podem ser implementados e a proposta desse trabalho é analisar e comparar algumas delas, como o Classificador Naivy Baies [3].

Hoje em dia, já é bastante comum encontrar sistemas cujo objetivo é classificar automaticamente esses dados. Tornando mais simples o acesso a informações através de gráficos e resumos de dados pré-processados.

1. Objetivos

O objetivo deste trabalho é desenvolver uma aplicação que automatize o processo de análise de sentimento. Serão implementados algumas técnicas diferentes para comparação e sumarização de resultados, dentre elas o Classificador Naive Bayes. E, além disso, os dados analisados serão capturados em tempo real.

3. Cronograma

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Atividade / Mês** | Março | Abril | Maio | Junho |
| Levantamento bibliográfico e estudo do material |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Definição da abordagem e técnicas a serem utilizadas |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Implementação |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Elaboração do relatório |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Preparação da apresentação |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

4. Possíveis Avaliadores

Os possíveis avaliadores são:

* Flávia de Almeida Barros {fab@cin.ufpe.br}
* Patrícia Cabral de Azevedo Restelli Tedesco {pcart@cin.ufpe.br}

5. Referências Bibliográficas

[1] Olhar Digital. Web. 05 de abril de 2012. <http://olhardigital.uol.com.br/produtos/digital\_news/noticias/966\_exabytes\_esta\_e\_a\_quantidade\_de\_dados\_que\_trafegarao\_na\_internet\_em\_2015>

[2] Sentiment Analysis. Wikipedia, thre Free Encyclopedia. Web 05 de abril de 2012.

<http://en.wikipedia.org/wiki/Sentiment\_analysis>

[3] Naive Bayes. Wikipedia, thre Free Encyclopedia. Web 05 de abril de 2012.

<http://en.wikipedia.org/wiki/Naive_bayes>

[4] Pang, Bo and Lee, Lillian. Opinion mining and sentiment analysis. Vol. 2, 1–135 (2008)

[5] Das, S.; Chen, M.; Yahoo! for Amazon: Extracting market sentiment from stock message boards. 2001.

[6] Liu, Bing. Handbook of Natural Language Processing – Analysis Sentiment and Subjectivity. Departament of Computer Science, University of Illinois at Chicago, 2010.

6. Assinaturas

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ricardo Bastos Cavalcante Prudêncio

**Orientador**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Paulo Henrique S. L. Floro de Oliveira

**Aluno**

1. http://www.twitter.com

² http://www.facebook.com

³http://www.elife.com.br [↑](#footnote-ref-1)