

LIGA DE ENSINO DO RIO GRANDE DO NORTE
CENTRO UNIVERSITÁRIO DO RIO GRANDE DO NORTE
CURSO DE SISTEMAS DE INFORMAÇÃO

JOHNATHAN MORONI SALES DO VALE SANTIAGO

SOFTWARES PARA ANÁLISE DE SENTIMENTOS

NATAL/RN
2015

JOHNATHAN MORONI SALES DO VALE SANTIAGO

SOFTWARES PARA ANÁLISE DE SENTIMENTOS

Trabalho de Conclusão de Curso
apresentado ao Centro Universitário do Rio
Grande do Norte como requisito final para a
obtenção do título de Bacharel em Sistemas
de Informação.

Orientador: Prof. Ytalo Rosendo do Amaral.

NATAL/RN
2015

Catalogação na Publicação – Biblioteca do UNI-RN
Setor de Processos Técnicos

Santiago, Johnathan Moroni Sales do Vale.

Softwares para análise de sentimentos. / Johnathan Moroni Sales do
Vale Santiago. – Natal, 2015.

23 f.

Orientador; Prof. Ytalo Rosendo do Amaral.

Monografia (Graduação em Sistemas de Informação) – Centro
Universitário do Rio Grande do Norte.

1. Sistemas de Informação – Monografia. 2. Mineração de Dados –
Monografia. 3. Odontologia – Monografia. I. Amaral, Ytalo Rosendo do. II.
Título.

RN/UNI-RN/BC

CDU 004.4

JOHNATHAN MORONI SALES DO VALE SANTIAGO

SOFTWARES PARA ANÁLISE DE SENTIMENTOS

Trabalho de Conclusão de Curso
apresentado ao Centro Universitário do Rio
Grande do Norte como requisito final para a
obtenção do título de Bacharel em Sistemas
de Informação.

Aprovado em: ____/____/____

BANCA EXAMINADORA

Yale Brendo L. Amorim

RESUMO

Análise de Sentimento (AS) ou também conhecida como Mineração de Opinião, extração de opinião, mineração de sentimento, análise de subjetividade, entre outras palavras é o campo de estudo que analisa as opiniões, sentimentos, avaliações, atitudes e emoções das pessoas a respeito de produtos, serviços entre outros meios. Com a globalização da internet as pessoas passaram a interagir e trocar informações de forma volumosa, então daí surge a necessidade de uma análise de opiniões nesse gigantesco mundo. A análise de sentimento explora o estudo computacional de opiniões expressas em fontes não estruturadas como textos, descobrindo qual a opinião expressa no documento e, classificar a sua polaridade. Para auxiliar a tomar suas decisões os usuários usam software que buscam experiências de diferentes indivíduos em *reviews*, fóruns e inclusive nas redes sociais. Para esta pesquisa bibliográfica e descritiva foram feitas buscas nas bases de dados: Scielo, Proquest e google acadêmico; nos idiomas português e inglês publicados no período de 2004 a 2014. Foram encontrados alguns softwares gratuitos com diferentes funcionalidades e foco de pesquisa, e foi escolhido um deles para analisar sua eficiência. Podemos ver que a análise de sentimentos não é 100% precisa no que se fez respeito aos resultados dos softwares com pôsteres, porém foi verificado que é boa o suficiente para ficar acima da média e obter um resultado significativo. Portanto há diversas ferramentas de análise de sentimento e o processo pode ser generalizado e aplicado a outros domínios e estudos onde militam organizações interessadas em monitorar opiniões dadas sobre produtos e serviços nas redes sociais.

Palavras-Chave: Mineração de Dados. Mineração de Opiniões. Análise de Sentimentos e Ferramentas da Análise de Sentimentos.

ABSTRACT

Sentiment analysis (SA) or also known as Opinion mining, extraction, mining, sentiment analysis of subjectivity, between other words is the field of study that examines the opinions, feelings, evaluations, attitudes and emotions of the people about products, services, among other means. With the globalization of the internet people have come to interact and exchange information so voluminous, so give the need arises for an analysis of opinions in this huge world. Sentiment analysis explores the computational study of opinions expressed in unstructured sources as texts, discovering what the view expressed in the document and classify their polarity. To help make your decisions users use software that seek experiences of different individuals in reviews, forums and social networks. For this bibliographical research and descriptive search were made in: Scielo databases, Proquest and google scholar; in Portuguese and English languages published in the period 2004 to 2014. Were found some free software with different functionality and focus of research, and was chosen one of them to analyze their efficiency. We can see that the analysis of feelings is not 100% accurate as it did about the results of the software with posters, but it was found that is good enough to stay ahead of the curve and get a meaningful result. Therefore there are several sentiment analysis tools and the process can be generalized and applied to other areas and studies where militate organizations interested in monitoring opinions given on products and services in social networks.

Keywords: Data mining. Sentiment analysis. Stages of Feelings and Analysis. Tools of Sentiment analysis.

SUMÁRIO

1 INTRODUÇÃO	06
2 REFERÊNCIAL TEÓRICO	07
2.1 ANÁLISE DE SENTIMENTO.....	07
2.2 ETAPAS DA ANÁLISE DE SENTIMENTO.....	08
2.3 TRABALHOS RELACIONADOS.....	10
2.4 SOFTWARES E FERRAMENTAS PARA ANÁLISE DE SENTIMENTO.....	14
3 METODOLOGIA	16
3.1 CARACTERIZAÇÃO DDA PESQUISA.....	16
3.2 PROCEDIMENTOS.....	16
4 RESULTADOS	17
5 CONSIDERAÇÕES FINAIS.....	19
REFERÊNCIAS.....	20

1 INTRODUÇÃO

É extraordinário o crescimento do número de informações disponíveis na Web nos últimos anos. Páginas Web juntam opiniões e sentimentos sobre os mais variados objetos, tais como, análises e opiniões de produtos, pesquisas de mercado entre outros. Com o aparecimento da Web 2.0, além dos conteúdos disponibilizados pelas empresas, os usuários começaram a partilhar críticas e opiniões em blogs pessoais, sites de relacionamentos, microblogs, redes sociais dentre outros meios. Este artigo apresenta como objetivo principal a importância da informação que a análise de sentimentos tem para facilitar a gestão de uma empresa e também conhecer algumas das ferramentas de análise de sentimentos. Os primeiros resultados da pesquisa cujo objetivo específico é mostrar alguns softwares gratuitos ou de baixo custo, como funciona a metodologia de cada um deles e quais são suas diferenças e elaborar um teste de análise de sentimento de marca automobilística em blogs. O resultado é o desenvolvimento do teste capaz de monitorar o sentimento sobre as marcas de automóveis na Web, por exemplo, o que as pessoas andam dizendo sobre o Honda Civic ou Toyota Corola nas mídias sociais? Tais resultados podem ser aproveitados com o intento de aperfeiçoar os produtos ou direcionar táticas de marketing além de ser aplicativa em negócios e outras organizações com interesse em monitorar opiniões emitidas sobre seus produtos e serviços.

2 REFERÊNCIAL TEÓRICO

2.1 ANÁLISES DE SENTIMENTO

Análise de Sentimento também conhecida como Mineração de Opinião, é o campo de estudo que analisa as opiniões, sentimentos, avaliações, atitudes e emoções das pessoas a respeito de produtos, serviços, organizações, indivíduos, edições, eventos, tópicos e seus atributos. Ela representa uma boa parte do espaço do problema. Há também muitos nomes e funções um pouco diferente, por exemplo, mineração de opinião, extração de opinião, mineração sentimento, análise de subjetividade, análise de afeto, análise de emoção, avaliação de mineração (LIU, 2012).

A análise de sentimento explora o estudo computacional de opiniões expressas em fontes não estruturadas como textos (Boiy apud Cavalcanti *et al.*, 2011). O objetivo não é determinar sobre qual tópico ou tema o documento trata (como realizado em técnicas convencionais de classificação de texto), mas sim em descobrir qual a opinião expressa no documento e, classificar a sua polaridade (CAVALCANTI *et al.* apud PANG; VELOSO, 2012).

A pesquisa de opiniões a respeito de produtos e/ou serviços na internet tem se tornado comum. Para auxiliar a tomar suas decisões os usuários buscam experiências de diferentes indivíduos em *reviews*, fóruns e inclusive nas redes sociais. Opiniões falsas, ou intencionalmente tendenciosas a inverdade a respeito de algum produto, tem se tornado um obstáculo para que se analisem corretamente essas informações, e ainda com o enorme volume de conteúdos disponíveis na internet, torna-se complicada uma investigação manual desses dados (RICARTE NETO apud JINDAL e LIU, 2007). Para uma busca efetiva na Web de tais opiniões, experiências e sentimentos é preciso haver uma forma automática de retirar o sentimento do texto. Assim é a tarefa de análise de sentimento (THELWALL, 2014). Liu apud Porto *et al.* (2010) mostra alguns dos principais problemas existentes na área de pesquisa análise de sentimento, ressaltando a importância tanto a indivíduos físicos, quanto empresas em pesquisa de mercado. O texto traz importantes conceitos como *features* (propriedades ou partes de um objeto) e opinião (sentimento positivo, negativo ou neutro em relação a um objeto ou à uma *feature*).

Nos últimos anos se tem notado um rápido crescimento de interesse em análise de sentimentos ou mineração de textos, devido a disponibilidade do grande volume de informações, documentos e opiniões pessoais expressadas pelos usuários. E as emoções são importantes para que haja uma comunicação efetiva no envio de mensagens para que possamos compreendê-las. Estudos da comunicação necessitam de meios para detectar o sentimento, com o objetivo de investigar ou modelar os processos envolvidos.

(THELWALL, 2014; WANG, 2012)

2.2 ETAPAS DA ANÁLISE DE SENTIMENTO

Para que se alcance os melhores resultados, as pesquisas e análises feitas em um texto são divididas em tarefas que normalmente são sequenciais e complementares. Com tamanha complexidade existente em cada uma dessas tarefas, normalmente, as pesquisas atuais focam em apenas uma ou duas tarefas, permitindo que o restante seja realizado por intermédio do usuário ou fazendo-as de forma mais simples. (SIQUEIRA, 2010)

Essas Tarefas são constituídas em: Identificação de opiniões, extração de características, classificação de sentimentos e visualização e sumário, como mostra a figura 1.

a) **Identificação de opiniões:** Onde dado um texto é identificado se há sentimento e onde está localizado tal sentimento no texto.

A etapa inicial a ser executada é a percepção da subjetividade, distinguindo as sentenças objetivas das subjetivas. Pois apenas as sentenças subjetivas serão aplicadas na Análise de Sentimentos, por apresentarem sentimentos, opiniões ou pensamentos relacionados a algo. (PANG e LEE, 2008)

“Meu cunhado viajou para florida. **O tempo de viagem não é bom, mas ele vai na classe A. O lugar é lindo e aprender inglês não será tão difícil**”

No exemplo acima a identificação de opinião só classificou duas sentenças, descartando como na maioria dos trabalhos a sentença objetiva nessa primeira etapa. (LIMA, 2011)

b) **Extração de características:** Onde identifica-se as características relacionados ao objeto a ser analisado, essa é uma das etapas que se encaixam como parte da identificação de acordo com a figura 1.

É uma das etapas mais importantes da Análise de Sentimentos, pois será através de seus resultados que será realizada uma análise mais detalhada nas próximas etapas.

Quando encontramos uma opinião positiva não significa que o usuário concorda completamente com todos os aspectos do objeto. De maneira similar funciona com a opinião negativa. O foco desta tarefa é que as aplicações de análise de sentimentos descubram quais os atributos e aspectos que serão mencionados para possibilitar a melhoria do objeto ou a comparação entre diferentes marcas. Portanto é preciso captar as informações explícitas e implícitas a que cada opinião está se referindo.

“ O **motor** do carro não é bom, mas seu **design** é sofisticado. O **ensino** escola é bom e o prédio não é pequeno (**tamanho**). “

No exemplo acima vemos os atributos avaliados pelo portador da opinião, é possível perceber também que há atributos implícitos (no caso, tamanho) e que precisam ser referenciados por meio de um indicador de atributo. ” (LIMA, 2011)

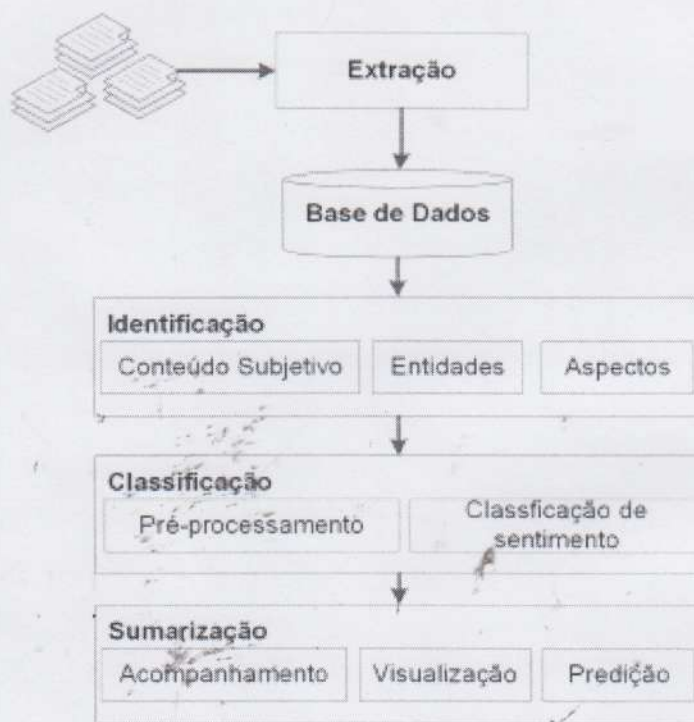
c) **Classificação de sentimentos:** Onde se diz se as polaridades das opiniões, ou seja, se é positiva, negativa ou neutra.

Essa classificação é repetidamente um problema de classificação conjugada, isto é, que rotula um texto em uma de duas classes: positivo ou negativo. Contudo, as adições de classes podem ser aceitas para que a análise seja mais vigorosa. Portanto, as classes podem ser estendidas em categorizações com diferentes graus de intensidade (muito positivo, moderadamente positivo), ou em intervalos numéricos representando um grau de intensidade. É importante considerar também a categoria neutra, que junta textos sem uma convergência clara quanto a sua polaridade. (BECKER e TUMITAN, 2013)

d) **Visualização e sumário:** Onde apresenta-se as informações obtidas ao usuário, facilitando assim as tomadas de decisão.

Opiniões são essencialmente subjetivas. E para identificar a opinião média, apenas um titular não é suficientemente para de tomar uma ação. Na maioria das aplicações, é necessário captar as opiniões de um número considerável de pessoas. (LIU, 2012). O objetivo desta etapa é dizer que são necessárias a criação de métricas e sumários que quantifiquem a distinção de opiniões encontradas a respeito um mesmo alvo, que representem o sentimento geral, as quais poderão ser mentalizadas ou servir de entrada para outras aplicações.

Figura 1: Etapas da Mineração de Opinião

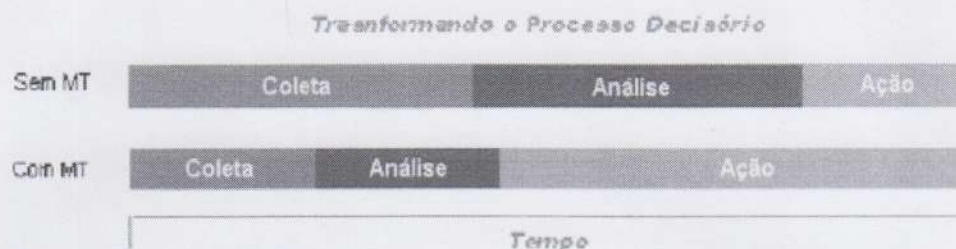


Fonte: (BECKER; TUMITAM, 2013)

As três principais etapas envolvidas na mineração de texto são: a seleção, a indexação e a análise de textos. A seleção tem como objetivo constituir ou formar a base de dados onde será feita a análise; para possibilitar uma rápida pesquisa por documentos específicos entra a indexação onde em certas ocasiões poderá ser opcional; e, por fim, a análise de textos como objetivo extrair informações através de algoritmos inteligentes e interpretar o conhecimento adquirido em tal extração. (ARANHA, 2007)

Segundo Aranha (2007) a mineração de texto pode contribuir com a redução de tempo em coletar e analisar informações e assim desprender um tempo maior nas ações pretendidas como é esboçado na figura 2 a seguir:

Figura 2: Processo da Mineração de



Fonte: (ARANHA, 2007)

2.3 TRABALHOS RELACIONADOS

Mikalai Tsytsarau (2011) em seu estudo, analisou o desenvolvimento da Análise de sentimento e a Mineração de opinião durante os últimos anos, e também discutiu a evolução de uma relativa nova direção de investigação, ou seja, Análise Contradição. Dar uma visão geral dos métodos propostos e os avanços recentes nestas áreas, e tentou esquematizar as futuras direções de pesquisa no campo.

Reis *et al.* (2015) propôs uma abordagem para o uso de determinados métodos estado-da-arte para análise de sentimentos em 9 diferentes línguas. Para isto, foi utilizado bases de dados previamente rotuladas em cada idioma e uma simples tradução automática para o inglês e desenvolveu uma metodologia para comparar e validar os resultados. Os resultados demonstram o potencial desta abordagem para tornar a análise de sentimentos independente da língua inglesa.

Francisco Neto (2014) falou sobre o ASDP, um Processo de Análise de Sentimento em Debates Polarizados, com foco nos debates não ideológicos. A classificação da postura dos posts é feita com base na identificação de padrões linguísticos que foram observados em corpora de debates polarizados. Esses padrões recuperam triplas do tipo <produto, palavra opinativa, sentença>, que consideram o contexto de ocorrência dos termos para a atribuição da classe do post. Os resultados dos experimentos revelam taxas de 73,91% de acerto na classificação dos posts.

Porto *et al.* (2012) apresenta o resultado do projeto de pesquisa que objetiva criar um processo de análise de sentimento baseado em ontologias no domínio

automobilístico e desenvolver um protótipo. Tal processo busca analisar quais sentimentos e opiniões têm os usuários de mídias sociais na Web sobre marcas e partes de veículos de passeio. Os resultados podem ter várias serventias assim como auxiliar tomadas de decisão no intuito de aperfeiçoar os produtos e/ou direcionar táticas de marketing mais específicas.

Santos *et al.* (2010) realiza uma investigação sobre o Twitter, analisando-o como um meio de contato direto com os clientes de empresas. Para que se estime o quanto os usuários estão expressando suas opiniões q respeito de um determinado produto. O sistema operacional Windows foi escolhido além do método de aprendizado de máquina conhecido como SVM (Support Vector Machine). Os resultados indicam que a máxima parte do que é postado no Twitter não caracteriza a opinião do usuário.

Silas Junior (2015) realizou um trabalho de utilização de técnicas tanto de recuperação de informação, quanto de análise de sentimentos, para dar suporte à empresas e empreendedores a fazerem pesquisas de mercado mais precisas, levando em consideração as opiniões dos usuários da rede. Utilizando-se como base de pesquisa o Twitter, onde são postados diariamente uma quantidade considerável de opiniões, foi necessário a utilização de um conjunto de APIs e técnicas para classificar cada opinião expressa nas postagens. Concluiu-se que é possível utilizar o Twitter como fonte de pesquisa, para rotular tweets usando análise de sentimentos e auxiliar empreendedores em suas análises de mercado

Barcelos *et al.* (2014) realizou seu projeto com base em um estudo sobre o engajamento de viajantes com hotéis de um destino turístico através de *reviews online* em *websites* de planejamento de viagens. Obteve como resultado a categorização por temas e a visualização de dados em gráficos que podem condensar as informações disponíveis em uma rede social especializada de turismo (TripAdvisor).

A análise de sentimentos (AS) sobre dados encontrados na *World Wide Web* é útil para empresas ou instituições que preocupam-se com o controle de qualidade. No momento, obter o *feedback* do usuário ou cliente, quer dizer questioná-los sobre aspectos da empresa para saber a opinião do usuário ou cliente sobre os pontos fortes e fracos da empresa ou instituição em questão. (BOIY *et al.*, 2007)

Mane *et al.* (2014) disse que o *twitter* é um dos maiores sites de mídia social recebe os *tweets* em milhões todos os dias. E que esta enorme quantidade de dados

brutos pode ser usada para fins industriais ou de negócios através da organização de acordo com a nossa exigência e processamento. E que tal papel fornece uma forma de análise de sentimento usando *Hadoop* que vai processar a enorme quantidade de dados em um cluster *Hadoop* mais rápido em tempo real.

Mohammed (2013). Realizou alguns testes de precisão de acertos e identificação de polaridade com duas ferramentas de análise de sentimento on-line gratuitas: SocialMention e Twendz. Seus experimentos mostraram que a ferramenta de análise de sentimento SocialMention é mais eficaz do que o Twendz, as medidas de desempenho levaram em conta quatro atributos para o critério de avaliação: acertos, erros, precisão e o recall. Como mostra no quadro 1 e no quadro 2 abaixo:

Quadro 01: Resultados da avaliação de Naïve Bayes por SocialMention e Twendz.

Naïve Bayes Evaluation Results for SocialMention and Twendz

Performance Measures	Twendz	Socialmention
Accuracy	45.3%	66.2%,
Error	54.7%	33.7%
Precision	0.45	0.66
Recall	0.45	0.66

Fonte: (MOHAMMED, 2013)

Quadro 02: SVM os resultados da avaliação para Socialmention e Twendz.

SVM Evaluation Results for SocialMention and Twendz.

Performance Measures	Twendz	Socialmention
Accuracy	43.3%	65.4%
Error	56.7%	34.6%
Precision	0.47	0.65
Recall	0.47	0.65

Fonte: (MOHAMMED, 2013)

2.4 SOFTWARES E FERRAMENTAS PARA ANÁLISE DE SENTIMENTOS

LIWC (Linguistic Inquiry and Word Count). É um programa de análise de texto transparente que conta palavras em categorias psicologicamente significativas. Além de detectar a sua polaridade. Os resultados empíricos usando LIWC demonstram a sua capacidade para detectar o significado em uma ampla variedade de configurações experimentais, incluindo a mostrar foco, emotividade, as relações sociais, o pensamento, estilos, e diferenças individuais. A ferramenta é unicamente comercial seu dicionário padrão é para o idioma inglês. O software encontra-se disponível em: <http://www.liwc.net/>. (TAUSCZIK; PENNEBAKER, 2010)

SentiStrength. Aprimorado por aprendizagem de máquina esse sistema utiliza um dicionário qualificado por humanos. Assim como uma lista de sentimentos léxicos e emoções seus recursos são armazenados como arquivos de texto e o sistema quando iniciado aponta para esse mesmo local. (TAUSCZIK e PENNEBAKER, 2010) O referido método classifica uma determinada mensagem numa escala de [-5, 5], onde -5 e 5 representam as pontuações mais negativas e positivas respectivamente. A versão do Windows pode ser baixada gratuitamente no site: <http://sentistrength.wlv.ac.uk/> (REIS *et al.* 2015)

NRC Emotion Lexicon. É uma ferramenta que utiliza um método léxico que rotula textos em 8 categorias, ao qual foi definida no livro '*A general psychoevolutionary theory of emotion*' de Plutchik em 1980 que são: alegria, tristeza, raiva, medo, confiança, desgosto, antecipação e surpresa. (MOHAMMAD, 2012)

OpinionLexicon. Seu outro nome é *Sentiment Lexicon*, incide de uma lista com aproximadamente 6.800 palavras classificadas como positivas e 6.800 palavras classificadas como negativas, abrangendo as gírias e abreviações no idioma Inglês. (HU; LIU, 2004)

SocialMention. É uma plataforma de pesquisa e análise de mídias sociais que agrega o conteúdo gerado pelo usuário de todo o universo em um único fluxo de informações. Agrega e fornece conteúdo de 'user-generated' de mais de 100 redes sociais diferentes, com base em uma pesquisa de palavra-chave. (LUDWIG; REUTER, 2014).

Oneall. Disponibiliza uma API que acopla os meios de comunicação sociais que são mais de 25 serviços e algumas funções sociais em uma única solução. As funcionalidades básicas são gratuitas; Tal API tem a capacidade de acessar apenas

as mensagens de mídias sociais, que também foram publicados pelo oneall. (LUDWIG; REUTER, 2014).

Topsy. É uma empresa de análise fundada em 2006, coletando tweets, links, fotos e vídeos de acordo com sua polaridade. No que se diz respeito aos sentimentos, ele raciona-os em dois valores: positivo e negativo. (LI, 2015)

Mention. É uma empresa que lida com a extração de marca, palavra-chave e análise de sentimento. Permite também processar o texto através de muitos recursos como: Blog, Twitter, Facebook. (LI, 2015)

3 METODOLOGIA

3.1 CARACTERIZAÇÃO DA PESQUISA

Para esta pesquisa bibliográfica e descritiva foram feitas buscas nas bases de dados: Scielo, Proquest e google acadêmico; nos idiomas português e inglês publicados no período de 2004 a 2015. A pesquisa é do tipo bibliográfica e descritiva

3.2 PROCEDIMENTOS

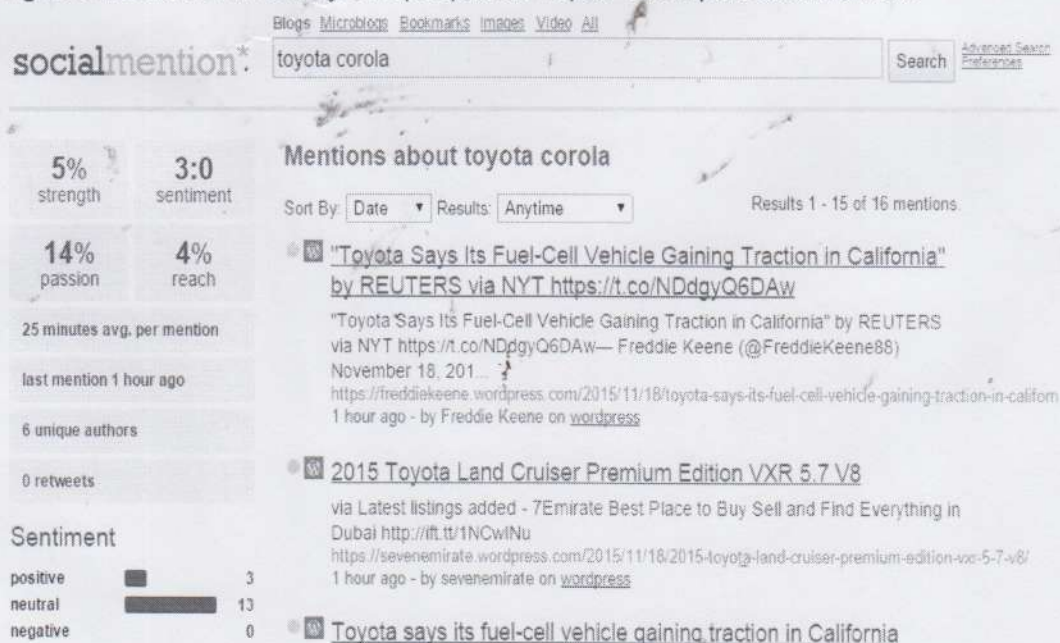
Foram pesquisados artigos científicos e livros sobre a análise de sentimento através das bases de dados: Scielo, Google Acadêmico, e Proquest; com os seguintes descritores: "Mineração de Dados". "Análise de Sentimentos". "Etapas da Análise de Sentimentos" e "Ferramentas da Análise de Sentimentos".

4 RESULTADOS

Dentre os softwares encontrados escolheu-se a ferramenta Socialmention pelo fato de ser multiplataforma, gratuita, ter uma boa usabilidade e não haver necessidade de realizar qualquer instalação ou configuração. Pois ao entrar no site qualquer usuário já pode começar a realizar suas pesquisas. Segundo Li (2015) O Socialmention não é capaz de analisar símbolos, e palavras de polaridades indiretas como no inglês "not bad" que no português seria algo como 'nada mal' o software resulta com um alto índice negativo onde deveria ser positivo ou até neutro dependendo da frase. Apesar de suas limitações o socialmention foi bem-conceituado em algumas análises feitas por Mohammed (2013).

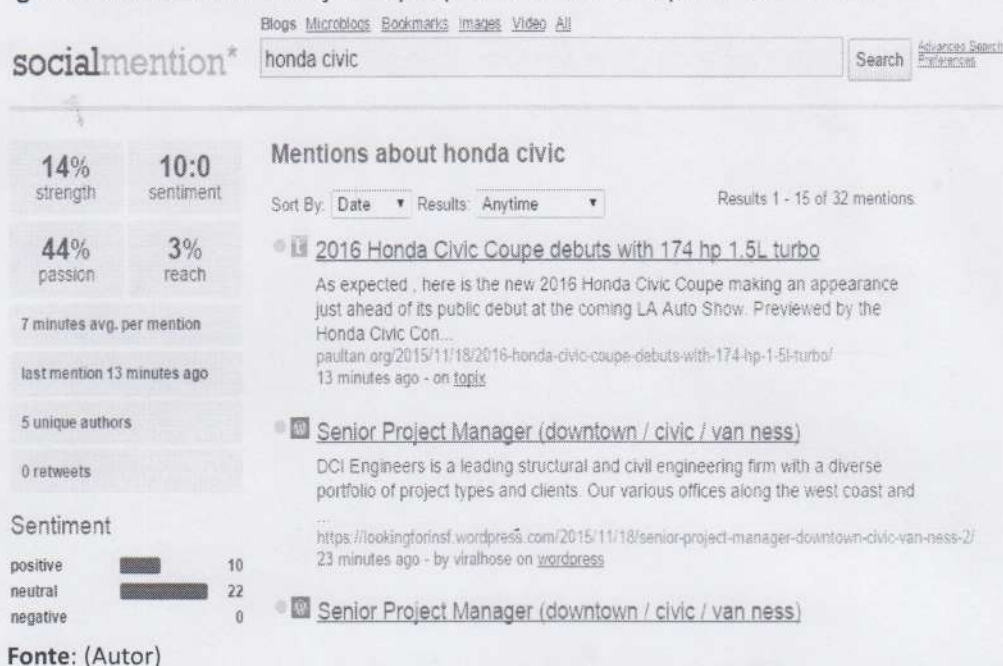
Ao realizar uma comparação entre o Toyota Corola e o Honda Civic no Socialmention apenas nos blogs os primeiros resultados obtidos para o Corola foram: 16 dados encontrados onde 3 eram positivos, 13 neutros e 0 negativos. E os campos de análise da página foram *strength* (força) 5%, *sentiment* (sentimento) 3:0, *passion* (paixão) 14% e *reach* (alcance) 4%. No caso do Civic foram: 33 dados encontrados na pesquisa onde 9 eram positivos, 24 neutros e 0 negativos. E os campos de análise da página foram *strength* (força) 14%, *sentiment* (sentimento) 10:0, *passion* (paixão) 44% e *reach* (alcance) 3%. Como mostram as figuras 3 e 4.

Figura 3: Resultados da avaliação de pesquisa do Toyota Corola pelo SocialMention.



Fonte: (Autor)

Figura 4: Resultados da avaliação de pesquisa do Honda Civic pelo SocialMention



Onde força é a probabilidade ou a intensidade de sua marca nos meios de comunicação sociais. Para realizar esse cálculo o sistema faz: frase mencionada nas últimas 24 horas, dividido pelo número total possível da frase.

Sentimento é a proporção de menções que são positivas para aqueles que são geralmente negativas.

Paixão é uma medida da probabilidade ou intensidade de indivíduos falando sobre algo repetidamente. Por exemplo, se você tem um pequeno grupo de defensores muito apaixonados, que falam sobre seus produtos ou marca o tempo todo, você terá uma maior pontuação de paixão. Por outro lado, se cada menção é escrita por um autor diferente você terá uma pontuação mais baixa.

Alcance é uma medida do intervalo de influência, para obter esse valor você pega o número de autores originais, referenciando sua marca dividida pelo número total de menções.

5 CONSIDERAÇÕES FINAIS

O TCC apresenta os resultados da pesquisa na área de análise de sentimentos e mineração de opiniões. Foram encontrados 8 softwares e ferramentas foram eles: Liwc, SentiStrength, NRC Emotion Lexicon, OpinionLexicon, SocialMention, Oneall, Topsy, Mention. E foi desenvolvido um processo de análise de sentimento, utilizando o SocialMention aplicado ao domínio automobilístico e que inclusive poderá ser generalizado e aplicado a outros domínios. Preliminarmente, apenas a começo de exemplo do tipo de informação que se pode obter do sistema no domínio de veículos, pôde-se concluir com a análise como mostra o resultado das Figuras 5 e 6, que existem mais pessoas falando bem do automóvel Honda Civic do que o Toyota Corola. Além disso, percebeu-se uma quantidade maior de informações relacionadas ao Civic do que ao Corola. A opinião intensiva dos usuários falando repetidamente sobre os carros foi de aproximadamente três vezes maior no Civic do que no Corola. Trata-se de um resultado automático construído pelas opiniões que circulam nas redes sociais sobre tais veículos. Entender como um produto é visto por clientes é o primeiro passo para melhorar sua imagem. Existem muitas ferramentas eficientes de análise de sentimentos e o socialmention se mostrou capaz de trazer resultados consideráveis.

REFERÊNCIAS

- ARANHA, Christian Nunes. **Uma Abordagem de Pré- Processamento Automático para Mineração de Textos em Português: Sob o Enfoque da Inteligência Computacional**. 2007. 144 f. Tese (Doutoradp em Engenharia Elétrica) - Departamento de Engenharia Elétrica da Puc-rio., Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007. Disponível em: <<http://livros01.livrosgratis.com.br/cp064589.pdf>>. Acesso em: 30 nov. 2015.
- BARCELOS, Yussif Tadeu et al. Ferramenta de visualização de dados e processamento de texto: análise de reviews de viajantes no tripadvisor. **Revista Eletrônica do Alto Vale do Itajaí**, Itajaí, v. 3, n. 4, p.25-39, dez. 2014. Disponível em: <<http://www.revistas.udesc.br/index.php/reavi/article/download/5051/4198>>. Acesso em: 28 out. 2015.
- BEKER, Karin; TUMITAN, Diego. **Introdução à Mineração de Opiniões: Conceitos, Aplicações e Desafios**. 2013. (Pós-graduação) - Programa de Pós-Graduação em Ciência da Computação, Instituto de Informática Universidade Federal do Rio Grande do Sul. Porto Alegre, 12 f. [201-]. Disponível em: <http://www.inf.ufrgs.br/~kbecker/lib/exe/fetch.php?media=minicursosbbd_versaosubmetida.pdf>. Acesso em: 25 set. 2015.
- BOIY, Erik; et al. Automatic Sentiment Analysis in On-line Tex. In: CONFERENCE ON ELECTRONIC PUBLISHING, 2007, Vienna, Austria. **Anais eletrônicos...** Vienna, 2007. Disponível em: <http://elpub.scix.net/data/works/att/138_elpub2007.content.pdf>. Acesso em: 05 out. 2015.
- CAVALCANTI, Diana C. et al. **Análise de Sentimento em Citações Científicas para Definição de Fatores de Impacto Positivo**. Universidade Federal de Pernambuco, UFPE, Recife. 10f. 2012. Disponível em: <<http://www.labic.icmc.usp.br/wti2012/artigos/105238.pdf>>. Acesso em: 20 out. 2015.
- HU, Minqing; LIU, Bing. **Mining and Summarizing Customer Reviews**. Department of Computer Science University of Illinois. 10f. 2004. Disponível em: <<https://www.cs.uic.edu/~liub/publications/kdd04-revSummary.pdf>>. Acesso em: 22 out. 2015.
- LI, Qin. **Examining the accuracy of sentiment analysis by brand monitoring companies**. University of Twente, Faculty of Management and Governance, Enschede, 2015. Disponível em: <http://essay.utwente.nl/67414/1/Li_BA_MB.pdf>. Acesso em: 27 out. 2015.

LIMA, Diego Carlos Lucena de Albuquerque. **PairExtractor**: Extração de Pares Livre de Domínio para Análise de Sentimentos. Universidade Federal de Pernambuco, Recife, 2011. Disponível em: <<http://www.cin.ufpe.br/~tg/2011-2/dclal.pdf>>. Acesso em: 29 out. 2015.

LIU, Bing. **Sentiment Analysis and Opinion Mining**. University of Illinois, department of computer Science, Chicago. abr. 2012. Disponível em: <<https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>>. Acesso em: 30 out. 2015.

LUDWIG, Thomas; REUTER, Christian. **Identification of Social Network Providers and API Design**. Siegen, Set. 2014. Disponível em: <http://www.fp7-emergent.eu/wp-content/uploads/2014/05/20140930_D5.1_IdentificationofSocialNetworkProvidersandAPIDesign_EmerGent.pdf>. Acesso em: 03 nov. 2015.

MANE, Sunil B et al. Real Time Sentiment Analysis of Twitter Data Using Hadoop. **International Journal of Computer Science and Information Technologies**, v. 5, n. 3, Pune, 2014. Disponível em: <<http://www.ijcsit.com/docs/Volume%205/vol5issue03/ijcsit2014050393.pdf>>. Acesso em: 15 set. 2015.

TSYTSARAU, Mikalai; PALMANAS, Themis. **Survey on mining subjective data on the web**. University of Trento, 24:478–514, 2012. Disponível em: <<http://disi.unitn.it/~themis/publications/dmkd11.pdf>>. Acesso em: 17 set. 2015.

MOHAMMAD, Saif M. #Emotional Tweets, **Emerging Technologies National, Research Council**, Montreal, p.246–255, jun.2012. Disponível em: <<http://www.aclweb.org/anthology/S12-1033>>. Acesso em: 21 set. 2015.

AL KABI, M. et al. Arabic / English Sentiment Analysis: An Empirical Study. **Conference Paper**, 7p. abr. 2013. Irbid, Jordan. Retrieved on Nov. 2015. Disponível em: <http://www.researchgate.net/profile/Mohammed_Al-Kabi/publication/236833290_Arabic__English_Sentiment_Analysis_An_Empirical_Study/links/0c960539aee7fac825000000.pdf>. Acesso em: 23 set.2015.

PANG, Bo; LEE, Lillian. Opinion mining and sentiment analysis. **Foundations and Trends in Information Retrieval**, v. 2, n. 1-2, 135 p. 2008. Disponível em: <<http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>>. Acesso em: 29 set. 2015.

PORTO, Renata Maria Abrantes Baracho et al. Análise de sentimento sobre veículos em redes sociais. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA

INFORMAÇÃO, 8., 2012. Belo Horizonte: Universidade Federal de Minas Gerais. 20p. 2012. Disponível em:
<http://enancib.ibict.br/index.php/enancib/xiiienancib/paper/viewFile/3863/2986>.
 Acesso em: 30 out. 2015.

REIS, Julio C. S. et al. **Uma Abordagem Multilíngue para Análise de Sentimentos**. 12 f. 2015. Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, 2015. Disponível em: <http://homepages.dcc.ufmg.br/~fabricio/download/brasnam15-multi.pdf>. Acesso em: 16 out. 2015.

RICARTE NETO, Francisco Assis. **ASDP: Um Processo de Análise de Sentimento em Debates Polarizados não Ideológicos**. 123 f. 2014. (Dissertação de Mestrado) – Universidade Federal de Pernambuco, Recife, 2014. Disponível em: <http://www.repositorio.ufpe.br/bitstream/handle/123456789/11415/DISSERTA%C3%87%C3%83O%20Francisco%20Assis%20Ricarte%20Neto.pdf?sequence=1&isAlloved=y>. Acesso em: 25 nov. 15.

SANTOS et al. **Twitter, análise de sentimento e desenvolvimento de produtos: Quanto os usuários estão expressando suas opiniões?** 12 f. 2010. Disponível em: <http://revistas.ua.pt/index.php/prismacom/article/view/790/722>. Acesso em: 02 set. 2015.

SILAS, José da Silva Junior. **Aplicação de Recuperação de Informação e Análise de Sentimentos para Suporte à Pesquisas de Mercado**. 64 f. 2015. Trabalho de Conclusão de Curso (Graduação em Sistema de Informação) - Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, 2015. Disponível em: <http://200.17.137.109:8081/novobsi/psi-na-ufpe-recife/trabalhos-de-conclusao-de-curso/2015.1/Silas%20Jose%20da%20Silva%20Junior.pdf>. Acesso em: 07 out. 2015.

SIQUEIRA, Henrique Borges Alencar. **WhatMatter: Extração e visualização de características em opiniões sobre serviços**. 109 f. 2010. (Mestrado em Ciência da Computação) - Universidade Federal de Pernambuco em Recife, Recife, 2010. Disponível em: http://www.repositorio.ufpe.br/bitstream/handle/123456789/2421/arquivo3250_1.pdf?sequence=1. Acesso em: 29 set. 2015.

TAUSCZIK, Yla R.; PENNEBAKER, James W. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. 31 f. 2010. **Jornal da Linguagem e Psicologia Social**, Austin, 2010. Disponível em: <http://homepage.psy.utexas.edu/HomePage/Faculty/Pennebaker/Reprints/Tausczik&Pennebaker2010.pdf>. Acesso em: 10 out. 2015.

THELWALL, Mike. **Heart and Soul: Sentiment Strength Detection in the Social Web with SentiStrength**. 14 f. 2013. Statistical Cybermetrics Grupo de Pesquisa da Faculdade de Tecnologia da Universidade de Wolverhampton, Wolverhampton. 2013. Disponível em: <<http://sentistrength.wlv.ac.uk/documentation/SentiStrengthChapter.pdf>>. Acesso em: 23 set. 2015.

WANG, Hao. et al. **A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle**. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, p 115–120, Jeju, Republic of Korea. 8-14 July 2012. Disponível em: <<http://www.aclweb.org/anthology/P12-3020>>. Acesso em: 04 nov.2015.