

UNIVERSIDADE FEDERAL DA PARAÍBA

CENTRO DE CIÊNCIAS APLICADAS A EDUCAÇÃO

DEPARTAMENTO DE CIÊNCIAS EXATAS

BACHARELADO EM SISTEMAS DE INFORMAÇÃO

**ANÁLISE DE SENTIMENTOS NO FACEBOOK UTILIZANDO
O SENTICNET**

GEORGE ALCÂNTARA PEREIRA DE LIMA

Orientador: Prof. Me. Yuri Malheiros

RIO TINTO - PB
2013

George Alcântara Pereira de Lima

**ANÁLISE DE SENTIMENTOS NO FACEBOOK UTILIZANDO
O SENTICNET**

Monografia apresentada para obtenção do título de Bacharel à banca examinadora no Curso de Bacharelado em Sistemas de Informação do Centro de Ciências Aplicadas e Educação (CCAEE), Campus IV da Universidade Federal da Paraíba.
Orientador: Prof. Me. Yuri Malheiros.

RIO TINTO - PB
2013

L732a Lima, George Alcântara Pereira de.

Análise de sentimentos no facebook utilizando o Senticnet / George Alcântara Pereira de Lima. – Rio Tinto: [s.n.], 2013.

64f.: il. –

Orientador: Yuri Malheiros.

Monografia (Graduação) – UFPB/CCAE.

1.Redes sociais – Aspectos pessoais. 2.Sentimento. 3.SenticNet. I.
Título.

UFPB/BS-CCAE

CDU: 316.3-059(043.2)

George Alcântara Pereira de Lima

ANÁLISE DE SENTIMENTOS NO FACEBOOK UTILIZANDO O SENTICNET

Trabalho de Conclusão de Curso submetido ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal da Paraíba, Campus IV, como parte dos requisitos necessários para obtenção do grau de BACHAREL EM SISTEMAS DE INFORMAÇÃO.

Assinatura do autor: _____

APROVADO POR:

Orientador: Prof. Me. Yuri Malheiros
Universidade Federal da Paraíba – Campus IV

Prof. Me. Marcus Williams Aquino de Carvalho
Universidade Federal da Paraíba – Campus IV

Prof^a. Dr. Flavia Veloso
Universidade Federal da Paraíba – Campus IV

RIO TINTO - PB
2013

À Deus, por está presente em todos os momentos da minha vida, que sempre me deu força para continuar até nos momentos mais Difíceis. Aos meus pais, Maria do Socorro Pereira de Lima e Jose Jorge de Lima Filho, por minha educação e apoio. Aos amigos, colegas e professores por me ajudarem a chegar nesta etapa.

AGRADECIMENTOS

Primeiramente a Deus, que iluminou o meu caminho durante esta etapa da minha vida. Com a ajuda Dele eu tive forças para chegar até o final com muita sabedoria. Ele me deu toda coragem que precisava para ir além dos meus limites e não me deixou faltar forças para ir até o final e quebrar as barreiras.

A meus pais Maria do Socorro e José Jorge pelo apoio em tudo que realizei durante minha graduação e compreensão em momentos de dificuldades.

A minha noiva amada Izabella Regina, que sempre me deu forças para que chegasse aonde cheguei até mesmo em momentos difíceis.

Ao meu orientador Yuri Malheiros, que desde o início me direcionou nas pesquisas, apresentando novas ideias e alternativas para o desenvolvimento do projeto.

Aos meus amigos dos quais puderam me ajudar diretamente em algumas atividades ocorridas no desenrolar do curso. Agradeço a Hermann Filho, Diego Lopes, Jefferson Araújo, Thiago Oliveira e todos os outros grandes amigos que fiz e contribuíram para a minha formação. Aqueles que estiveram comigo desde o início do curso, outros me acompanharam apenas no final dessa jornada, porém que também tiveram grande contribuição nessa caminhada.

Por fim e não menos importante, a todos os colaboradores do CCAE, desde professores aos auxiliares que contribuíram para a minha formação acadêmica e, principalmente, ética na área de TI.

“O ato da conquista de um grande sonho reflete o sucesso nas lutas diárias para alcançá-lo.”

Kléber Novartes

RESUMO

Saber a opinião das pessoas na Internet tem sido objeto de estudo de muitas pesquisas recentes, visto que, mais do que nunca, essas opiniões têm ganhado grande espaço nos blogs, fóruns e redes sociais. Empresas investem muito dinheiro para coletar esses dados importantes sobre as pessoas, para elaborar estratégias de marketing e vendas direcionadas. As redes sociais têm tido um crescimento significativo de usuários. As pessoas estão cada vez mais conectadas a essas redes e expressando suas opiniões, gostos, preferências, relacionamentos. O estudo de opiniões, sentimentos e emoções expressas em documentos é conhecido como Análise de Sentimento. Dentre algumas tarefas da Análise de Sentimento, destaca-se a classificação da polaridade em textos, ou seja, o quão negativa ou positiva são as opiniões expressas nele.

Este trabalho tem como objetivo o estudo, implementação e avaliação de um sistema de classificação de mensagens, coletadas através de pesquisas no *Facebook*, usando o SenticNet. O sistema é capaz de coletar mensagens e quebra-las por palavras, fazendo uma busca na base de dados do SenticNet e obtendo o valor da polaridade e sentimento da mensagem. Para avaliação da solução proposta, foi aplicado um questionário com algumas pessoas, a fim de realizar um comparativo entre os resultados coletados neste questionário e os apresentados pelo sistema desenvolvido e conseguir medir a eficiência do sistema. Os resultados se mostraram positivos, visto que o sistema conseguiu acertar 70,4% das mensagens e como trabalhos futuros realizar ajustes e conseguir melhorar esse valor para um valor mais próximo do ideal.

Palavras chave: redes sociais, análise de sentimento, SenticNet.

ABSTRACT

Know the opinion of people on the Internet has been object of study in many recent researches, seeing that, more than ever, these opinions have gained large space in blogs, forums and social networks. Companies invest a lot of money to collect this important data about people in order to develop marketing strategies and targeted sales. Social networking has been having a significant growth in users. People are increasingly connected to these networks to express their opinions, tastes, preferences and relationships. The study of opinions, feelings and emotions expressed in documents is known as sentiment analysis. Among some tasks of the sentiment analysis, stands out the polarity classification into texts, in other words, just how positive or negative are the opinions expressed therein.

This work aims the study, implementation and evaluation of a system of classification of posts, collected through surveys on Facebook, using the SenticNet. The system is able to collect messages and break them by words, by searching the database of SenticNet and obtaining the value and sense of the polarity of the message. To evaluate the proposed solution, a questionnaire was given to some people in order to make a comparison between the results collected in this questionnaire and submitted by the system developed and able to measure system efficiency. The results were positive, since the system managed to hit 70.4% of the messages and make adjustments as future work and get better this value to a value closer to the ideal.

Keywords: social network, sentiment analysis, SenticNet.

LISTA DE FIGURAS

Figura 1 – Sumarização dos resultados no sistema de Santos[6].	12
Figura 2 – Pagina inicial do Facebook	14
Figura 3 - Usuários ativos do Facebook.	16
Figura 4 - Modelo de um conceito do SenticNet no formato RDF/XML.	18
Figura 5 - Arquitetura do sistema desenvolvido.	21
Figura 6 - Graph API Explorer	22
Figura 7 - Arquivo JSON de uma busca realizada utilizando o API Graph.	23
Figura 8 - Exemplo de um recurso no SenticNet.	24
Figura 9 – Exemplo de saída do sistema	26
Figura 10 – Exemplo de classificação pelo sistema	27
Figura 11 – Exemplo de pergunta da avaliação.	28
Figura 12 – Resultado da classificação através da avaliação para uma mensagem	31
Figura 13 – Resultado da classificação através do sistema para uma mensagem	31
Figura 14 - Resultado da classificação através da avaliação para uma mensagem	32
Figura 15 - Resultado da classificação através do sistema para uma mensagem	32

LISTA DE TABELAS

Tabela 1 – Resultado da classificação realizada através do sistema.	29
Tabela 2 – Resultado da classificação realizada através da pesquisa.	29
Tabela 3 – Tabela comparativa dos resultados	30

SUMÁRIO

RESUMO	VIII
ABSTRACT	IX
LISTA DE FIGURAS	X
LISTA DE TABELAS.....	XI
SUMÁRIO.....	XII
1 INTRODUÇÃO.....	1
1.1 CONTEXTUALIZAÇÃO E MOTIVAÇÃO	2
1.2 OBJETIVOS, METODOLOGIA E ORGANIZAÇÃO DO TRABALHO	3
1.2.1 OBJETIVO GERAL.....	3
1.2.2 METODOLOGIA.....	4
1.2.3 ORGANIZAÇÃO DO TRABALHO.....	4
2 FUNDAMENTAÇÃO TEÓRICA.....	5
2.1 ANÁLISE DE SENTIMENTO.....	5
2.1.1 DIFICULDADES PARA A REALIZAÇÃO.....	7
2.1.2 APLICAÇÕES PRÁTICAS.....	8
2.1.3 ETAPAS DO PROCESSO E TÉCNICAS UTILIZADAS	9
2.2 REDES SOCIAIS	12
2.2.1 SOBRE O FACEBOOK.....	14
2.3 SENTICNET	17
3 DESENVOLVIMENTO	19
3.1 PROCEDIMENTOS METODOLÓGICOS	19
3.1.1 VISAO GERAL	19
3.1.2 O SISTEMA DESENVOLVIDO	20
4 EXPERIMENTOS E RESULTADOS	27
4.1 METODOLOGIA PARA EXPERIMENTO	27
4.2 RESULTADOS E AVALIAÇÃO	28
5 CONCLUSÃO E TRABALHOS FUTUROS	33
REFERÊNCIAS BIBLIOGRÁFICAS.....	34
APÊNDICE	37

1 INTRODUÇÃO

Com o crescimento da utilização das redes sociais, blogs, sites de opiniões, entre outras formas de comunicação *on-line*, é evidente o aumento da quantidade de informações disponíveis pela Internet, como mensagens, críticas, opiniões, classificações, recomendações. As pessoas estão cada vez mais dispostas a interagirem entre si, compartilhar seus conhecimentos, opinar sobre produtos ou serviços, observar qual a opinião de um grupo de pessoas sobre assuntos do seu interesse. Empresas também se interessam em medir a aceitação de um novo produto, monitorando as opiniões de um grupo em relação a ele.

A classificação automática de opinião é uma área de estudo muito desejada pelas empresas. Elas investem nesta área para direcionar seus trabalhos de marketing e vendas. Entretanto, a extração e análise do sentimento ainda costumam ser manual, pois, atualmente, existem poucas ferramentas capazes de realizar esse tipo de serviço de forma automatizada. Os poucos aplicativos disponíveis no mercado servem apenas para busca do conteúdo (SALUSTIANO, 2013).

A análise de sentimentos propõe identificar o sentimento que as pessoas expõem em documentos a respeito de alguma coisa de interesse (um produto específico, uma pessoa, uma empresa, um lugar, entre outros). Está associada à ideia de categorias (por exemplo, positiva, negativa, neutra) e busca, a partir de uma junção de algoritmos, desde técnicas tradicionais para processamento de texto até algoritmos que possibilitam o treinamento a partir de um conjunto de dados, entender o valor sentimental deles.

Para esta finalidade, este trabalho tem como objetivo o estudo, implementação e avaliação de um sistema de classificação de mensagens do *Facebook* baseado na influência das palavras utilizando o SenticNet, um recurso semântico e afetivo disponível publicamente para o conceito de mineração de opinião e análise de sentimentos.

Para tanto, em termos metodológicos, foi coletado e analisado as postagens feitas por usuários, utilizando de palavras-chave (temas) para buscas delas. Sobre as postagens coletadas, foi levado em consideração apenas o conteúdo da mensagem, desconsiderando qualquer outro indicador. Na análise, foi observado qual o sentimento exposto pelo autor a respeito daquela mensagem. Para isso, foi analisada cada palavra da postagem e utilizando o SenticNet, se tornou possível saber a polaridade do texto (métrica usada para saber o quanto positivo ou negativo aquela mensagem é sobre o assunto pesquisado).

Posteriormente, os dados foram cruzados com os coletados através de um questionário aplicado com pessoas, para saber o quanto eficiente é o sistema desenvolvido. Essa pesquisa

dispõe de um questionário com 27 mensagens sobre assuntos como, por exemplo, “obama”, “meteor”, “samsung”, “iphone”, “windows”, coletados através do sistema. Em cada mensagem, é pedido para as pessoas responderem se acham que a mesma fala sobre algo positivo ou algo negativo. Diante dessas informações, os resultados das classificações feitas pelas pessoas e pelo sistema serão comparados para que se possa medir a eficiência desse sistema de classificação de sentimento.

Esse estudo comparativo mostrou resultados interessantes, aonde apontou que o sistema desenvolvido obteve uma taxa de acerto considerável, acertando classificações com valores bem próximas do ideal, mas também errando em alguns textos. Nos últimos capítulos serão mostrados com mais detalhes esses resultados e avaliações.

1.1 CONTEXTUALIZAÇÃO E MOTIVAÇÃO

Antes do surgimento da Internet, as opiniões das pessoas eram obtidas com familiares, amigos e algumas mídias como: rádio, jornal televisão, entrevistas pessoais ou por correspondência. Já no lado comercial, esta tarefa era realizada por pesquisas qualitativas e quantitativas como: enquetes, grupos focais, entrevistas pessoais ou por correspondência, etc. Atualmente, porém, diante dos recursos oferecidos pelas Tecnologias de Informação e da Comunicação (TIC), e principalmente com suas mais recentes formas de uso da Web 2.0 (MOREIRA et al, 2009), surgiram inúmeras fontes de expressão permitindo pessoas ou grupos, organizações ou empresas expressarem suas opiniões em fóruns de discussão, blogs, redes sociais, dentre outras.

Diante disso, a obtenção e a análise dessas informações tornou-se uma área importante da computação. De acordo com Pang e Lee (PANG et al, 2002), empresas estão cada vez mais interessadas em reunir informações importantes a respeito de seus produtos, serviços, reputação e clientes. Além das empresas, as pessoas também se interessam em saber informações relevantes sobre produtos que possam vir a consumir, ou até mesmo o que está sendo falado sobre elas, como por exemplo, um candidato que quer saber se as pessoas estão falando bem ou mal a seu respeito.

Entretanto, como a velocidade em que essas informações são espalhadas na rede é tão grande, o trabalho de extração e análise acaba se tornando impraticável manualmente. Cada mensagem postada em um blog, rede social ou até mesmo em um *site*, tem suas especificidades e maneirismos, necessitando de abordagens distintas para a análise de seus

conteúdos. Diante desta realidade, faz-se necessária uma busca automática e específica de opinião e classificação destas quanto ao seu caráter (como positiva ou negativa).

Ainda segundo Pang e Lee (PANG et al, 2002), a análise de sentimento está preocupada em classificar textos não por tópicos, e sim pelo sentimento ou opinião contida neles. Geralmente associado à classificação binária entre sentimentos positivos e negativos, o termo é usado de uma forma mais abrangente para significar o tratamento computacional de opinião, sentimento e subjetividade em documentos, seja ele textos, imagens ou qualquer forma de expressão. Este trabalho lida com mensagens do *Facebook* e uma análise do seu sentimento, classificando a sua polaridade como positiva (se o valor for maior do que zero) ou negativa (se o valor for menor do que zero), numa escala de -1,0 a +1,0.

A principal motivação é obter um sistema capaz de classificar de forma automática corretamente o sentimento nas mensagens compartilhadas no Facebook. Além disso, não foi encontrado nenhum sistema ou pesquisa que trabalhe com análise de sentimentos em cima do Facebook. Outra motivação, como já mostrado no início deste tópico, é que existe um interesse enorme por parte de pessoas e empresas em saber a opinião das pessoas na Internet.

1.2 OBJETIVOS, METODOLOGIA E ORGANIZAÇÃO DO TRABALHO

1.2.1 OBJETIVO GERAL

Criar e avaliar um sistema que coleta mensagens do *Facebook* e, utilizando o SenticNet, classificá-las de acordo com a sua polaridade.

1.2.1.1 OBJETIVO ESPECÍFICO

- Pesquisar sobre análise de sentimento;
- Desenvolver uma aplicação para captura de mensagens do *Facebook* e classificação delas.
- Preparar e aplicar uma avaliação com as pessoas para coleta de opiniões.
- Realizar um estudo comparativo dos resultados obtidos pelo sistema e os coletados pelo questionário.

1.2.2 METODOLOGIA

De forma a alcançar os objetivos almejados, este trabalho utilizou de pesquisas de caráter bibliográfico e exploratório, que pudessem trazer fundamentação para o estudo e desenvolvimento do sistema. As etapas envolvidas foram:

- Levantamento bibliográfico sobre Redes Sociais e Análise de sentimentos;
- Estudo das tecnologias de apoio ao desenvolvimento do Software;
- Definição do modelo arquitetural do sistema;
- Metodologia para o processo de desenvolvimento;
- Desenvolvimento do sistema;
- Aplicação de questionário com usuários de redes sociais;
- Estudo comparativo dos resultados obtidos pelo sistema e questionário;

1.2.3 ORGANIZAÇÃO DO TRABALHO

Este trabalho está organizado da seguinte forma:

- No capítulo um é dada uma introdução sobre o tema abordado e o trabalho feito. Em seguida é mostrada uma contextualização do tema e a motivação do trabalho, bem como é fornecido os objetivos e sua organização.
- No capítulo dois são apresentados vários conceitos que serão abordados ao longo deste trabalho, os quais se fazem necessários conhecê-los e entendê-los para que se possa compreender as características e técnicas citadas no mesmo.
- O capítulo três ressalta, de forma detalhada, a solução, características e etapas do *software* desenvolvido para a análise de sentimento.
- No capítulo quatro são apontados resultados obtidos tanto pelo sistema desenvolvido como pelo coletado pela avaliação. Além disso, comparativos dos resultados entre os dois são apresentados.
- No último capítulo, são destacadas algumas características observadas durante o desenvolvimento do trabalho e pela a análise dos resultados obtidos. Também são apresentadas algumas propostas para trabalhos oriundos dessa pesquisa e sua contribuição no aspecto social e comercial.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 ANÁLISE DE SENTIMENTO

As pessoas procuram as opiniões das outras como forma de direcionarem suas compras e suas considerações em relação a um determinado assunto. Com a chegada da *web* como fonte de informações, grande parte dos usuários tem procurado nela textos que forneçam esse tipo de informação. Empresas estão interessadas em observar as opiniões de um grupo de pessoas sobre temas que lhe interessam, como por exemplo, estimar a aceitação de um novo produto, monitorando as opiniões em relação a ele. De acordo com Pang e Lee (PANG et al, 2002), a maioria das pessoas comenta que a pesquisa online sobre informações de produtos foram determinantes na decisão pela compra do mesmo.

Segundo Souza (SOUZA, 2011), opiniões são muito importantes quando precisamos tomar uma decisão, pois a partir delas criamos a nossa. Isso é fato tanto para pessoas como também para as empresas. Antigamente, antes de tomar alguma decisão, as pessoas recorriam a opiniões de amigos e familiares. As empresas apelavam para pesquisas de mercado quando queria saber mais sobre o seu público alvo. Nos dias de hoje, com as novas tendências sociais, o crescimento da Internet, as mudanças de comportamento das pessoas na web, essas pesquisas de opinião estão sendo mudadas. As pessoas encontram avaliações de produtos e serviços em sites especializados, fóruns e blogs, ou até mesmo em redes sociais. Elas podem expressar suas opiniões sobre os mais diversos temas para que outras pessoas leiam e compartilhem.

Com essa expansão da participação dos usuários nos conteúdos da web expondo seus pensamentos, novos conceitos e metodologias surgiram para investigação desses conteúdos (PANG et al, 2008). Assim, a análise de sentimentos tornou-se uma área de pesquisa em Processamento de Linguagem Natural e mineração de documentos. A análise de sentimentos é o estudo computacional de como opiniões, atitudes, emoções e perspectivas são expressas na linguagem natural (LIU, 2010), podendo ser considerada uma disciplina de estudo multicêntrica, englobando áreas como psicologia, marketing e computação.

A análise de sentimentos 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. Em uma visão geral, ela 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.

Para essa atividade são utilizadas ontologias, um conjunto de termos ordenados hierarquicamente para descrever um domínio que pode ser usado como esqueleto para uma base de conhecimentos. Em alguns casos, são usadas ontologias de tarefa. Este tipo de ontologia possui uma lista de palavras (simples e expressões) que servem como indicadores do sentimento. Além disto, são utilizados graus numéricos para indicar o quanto a palavra ou expressão indica o sentimento. Estes graus são gerados por processos probabilísticos, pois palavras de dois tipos podem estar presentes no texto (e mesmo um termo pode ser usado positiva ou negativamente). Assim, o processo probabilístico consegue identificar o sentimento predominante mesmo que palavras dos dois tipos estejam presentes (qual a probabilidade de cada sentimento estar presente no texto).

Com o objetivo de atingir melhores resultados, a análise da opinião de um texto é dividida em tarefas que normalmente são sequenciais e complementares. Tamanha é a complexidade de cada uma dessas tarefas que, geralmente, os trabalhos focam em uma tarefa ou duas, deixando o restante para ser realizado com a intervenção do usuário (SIQUEIRA, 2010).

Visando esta crescente procura por métodos que facilitem o processamento destes textos na web, a necessidade de *softwares* que realizem análise de sentimento tem aumentado. No entanto, a construção de um *software* que realize esse trabalho de maneira rápida e correta não é nem um pouco trivial. Ainda encontram-se vários estudos sobre o tema em andamento.

Uma parte das pesquisas existentes nesta área tem sido focada na mineração e recuperação de informações sobre fatos. Por exemplo, a busca na web, o agrupamento de textos, classificação de textos, e outros tipos de trabalho que trabalham com processamento de linguagem natural.

Até um bom tempo atrás, pouco trabalho tinha sido feito sobre a análise de opiniões realmente. Liu em (LIU, 2010) fala que uma das principais razões para a falta de estudo sobre opiniões é o fato de que havia pouco texto opinativo disponível antes da Web.

2.1.1 DIFICULDADES PARA A REALIZAÇÃO

Algumas dificuldades são encontradas durante o processo de análise de sentimento. Trabalhar com linguagem natural não é trivial nem para seres humanos, uma vez que diferentes opiniões influenciam o modo como cada um lê e entende um texto (SANTOS, 2010). Podem-se destacar algumas dificuldades para realizar esta tarefa (RODRIGUES et al, 2012):

- Textos com erros e sentenças sintaticamente mal formadas (o que é bastante comum nos blogs e redes sociais) dificultam a busca e classificação dos mesmos;
- Distinguir se um texto é uma opinião ou um fato;
- Textos podem conter sarcasmos e ironias, que são difíceis de serem identificados e podem impactar os resultados;
- Um texto pode referenciar mais de um item de interesse com opiniões diferentes sobre os itens, o que pode confundir a classificação;
- Uso de pronomes para referenciar itens pode dificultar a identificação de sentenças que mencionam o item de interesse;
- Uso de termos informais da Internet (por exemplo “blz”, “fds”) são bastante usados em redes sociais e dificultam a busca e classificação dos mesmos;

Com todos estes problemas citados, podemos perceber que realizar a análise de sentimentos é muito mais complexo do que se imagina, pois existem muitos fatores da linguagem natural difíceis de serem entendidos.

Algumas dificuldades podem ser contornadas no processo de classificação. Neste trabalho, as mensagens que representam fatos podem ser lidas e descartadas do processo. Textos com sarcasmos e ironias são muito difíceis de serem identificadas pela máquina, o que acaba sendo inviável para este sistema. Erros e sentenças mal formadas também são desconsiderados. Fica a critério do sistema apenas coletar a mensagem e classifica-la.

2.1.2 APLICAÇÕES PRÁTICAS

Segundo Santos (SANTOS, 2010), um sistema para análise de sentimento ou opinião pode ser construído para realizar diversas funções, cada uma necessitando de uma abordagem diferente, dependendo do que se deseja realizar.

São citadas por Rodrigues et. al. em (RODRIGUES et al, 2012) algumas aplicações praticas:

Sentimento sobre um produto: Esta análise pode ser usada por uma empresa para determinar a aceitação de certo produto lançado no mercado, contribuindo para alterações estratégicas nas áreas de marketings. Ainda é possível verificar o sentimento sobre um produto que foi anunciado, mas ainda não foi lançado, podendo ajudar em algumas alterações durante o desenvolvimento;

Empresas na bolsa de valores: Este cenário é usado para analisar empresas que contem ações na bolsa de valores. O sentimento sobre determinada empresa pode ser obtida através da opinião de especialistas e sites de notícias. O analisador ajuda na tomada de decisão sobre compras e vendas de ações, mas não deve ser um fator decisivo;

Análise sobre pessoas: O minerador de opinião pode ser usado para se obter o sentimento sobre determinada pessoas, especialmente daqueles que aparecem muito na mídia. É possível saber, por exemplo, a aceitação ou rejeição de um determinado político durante a época de eleições, dando alternativa para que o político altere sua estratégia de campanha;

Outras aplicações para o uso da análise de sentimentos também podem ser citadas:

Análise de lugares: uma pessoa que vai viajar pode utilizar as opiniões de outras sobre o lugar pretendido para planejar o roteiro da viagem, evitando passeios desinteressantes;

Análise de filmes e jogos: também é possível analisar a opinião das pessoas sobre filmes e jogos eletrônicos;

Os dados que são necessários para executar esta tarefa podem ser obtidos de diversas fontes oferecidas pela Internet, antigamente restrito a blogs e fóruns, hoje se expande as redes sociais. Podemos citar dois meios de obtenção de dados:

Usando mecanismos de busca: É possível usar mecanismos de busca como Google ou Bing para obtenção de opiniões na Internet, principalmente em fóruns e

blogs, mas podendo estender a qualquer site. Este método de obtenção de opinião é variável, conforme a ferramenta de busca altera seus resultados, e não é muito acurado, sendo que páginas favorecidas pelo mecanismo de busca podem alterar o resultado;

Facebook: O *Facebook* possui uma grande quantidade de usuários (1 bilhão de usuários ativos em outubro de 2012)¹. Os usuários podem demonstrar suas opiniões através de mensagens, compartilhamentos e “likes”. A obtenção dessas informações pode não ser tão trivial em comparação a outras redes sociais como o *Twitter*, visto que o *Facebook* possui mais recursos, o que acaba gerando um número maior de dados. O *Facebook* possui uma ferramenta em suas *fan pages* chamada “Facebook Insight”. Ela não funciona exatamente como um minerador de opiniões, mas possui informações sobre a página como *feedbacks* negativos;

2.1.3 ETAPAS DO PROCESSO E TÉCNICAS UTILIZADAS

O objetivo da análise de sentimentos é entender como o leitor pode interpretar uma emoção ou um sentimento em um texto e com isto desenvolver *softwares* que executem esta tarefa.

Segundo Souza em (SOUZA, 2011), as fases envolvidas no processo de análise de sentimentos são:

- 1) **Coleta de dados:** Fase onde são coletados os dados sobre um determinado objeto de pesquisa. Ela visa apenas buscar na *web* conteúdos relacionados ao tema e arquivá-los para análise e classificação.
- 2) **Classificação:** A etapa de classificação pode ser realizada por meio de técnicas de aprendizagem de máquina, seleção de palavras ou análise sintática. É nesta fase onde é calculado o quão positivo ou negativo os textos são em relação um tema.
- 3) **Sumarização dos resultados:** É na sumarização de resultados onde 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.

¹ <http://g1.globo.com/tecnologia/noticia/2012/10/facebook-atinge-1-bilhao-de-usuarios-ativos-mensais.html>

2.1.3.1 Coleta de dados

Nesta fase é feita uma busca em diversas fontes, tais como artigos em sites, comentários, postagens em redes sociais, blogs, anúncios, documentos dentre outras. É importante a utilização de técnicas avançadas de busca, procurando identificar se o conteúdo encontrado trata-se de uma opinião ou um fato. Dependendo de como for feita a etapa de análise, fatos por si só devem ser descartados, porém opiniões expressas em fatos devem ser mantidas.

É na *web* onde existe a maior coleção de textos existente atualmente. Os problemas enfrentados residem em sua grande dimensão e nas diferenças entre os vários tipos de páginas, ou seja, “a heterogeneidade é o desafio predominante”, segundo Carrilho Junior (CARRILHO JUNIOR et al, 2007). Assim, é possível imaginar como é custoso percorrer a *web* em busca de informações sobre um determinado tema e realizar o seu arquivamento de maneira manual. Para isso então, foram criadas ferramentas que executam todo o trabalho de navegar na Internet e buscar o que se deseja.

São por esses e outros motivos que a tarefa de busca do conteúdo e classificação em fato ou opinião não é uma atividade trivial. Ainda hoje existe uma grande demanda por pesquisas na área, e várias delas encontram-se em andamento.

2.1.3.2 Classificação

A fase de classificação é uma das principais no processo, pois é nessa etapa que ocorre a identificação da polaridade do texto, que é o principal objetivo da análise de sentimentos. Esta etapa busca determinar se uma opinião é positiva, neutra ou ainda negativa.

Na classificação, são as palavras opinativas que tem maior importância. Através delas, é possível determinar a polaridade expressa pelo texto. Exemplos de palavras opinativas: bom, legal, ótimo, ruim, péssimo etc.

Para alcançar o objetivo da classificação, existem diversas técnicas, dentre as quais:

- Aprendizagem de máquina supervisionada;
- Análise sintática;
- Seleção de palavras;

A técnica de aprendizagem de máquina supervisionada requer uma classificação prévia de um conjunto de conteúdos para ser utilizado como base de treinamento do modelo. A partir desse conjunto de treinamento, novos conteúdos são classificados. Essa técnica já é bastante utilizada e consolidada no ramo da inteligência artificial, o que constitui uma vantagem. Entretanto, a necessidade de diferentes conjuntos de treinamento para diferentes temas constitui uma desvantagem. Por exemplo, um modelo de classificação de carro não é adequado para classificar opiniões referentes a lugares (SOUZA, 2011).

A técnica de análise sintática é feita no conteúdo do texto, identificando os adjetivos e advérbios que possam apontar a opinião contida no texto. Essa técnica tem a vantagem de não ser necessário classificar anteriormente um conjunto de textos, como na aprendizagem de máquina. Entretanto, um analisador sintático eficiente é necessário, assim como que as sentenças devem estar sintaticamente corretas (SOUZA, 2011).

A técnica da seleção de palavras faz uso de um dicionário de palavras previamente selecionadas e com valores atribuídos a elas para classificar as palavras compreendidas no texto que está sendo analisado. Essa abordagem classifica o conteúdo do texto em positivo, negativo ou neutro. Entre suas vantagens está a simplicidade do processo e não ter que utilizar um conjunto de treinamento. Entretanto, podemos apontar como desvantagem o fato de ter que constituir previamente uma base de palavras associadas a valores. Por exemplo, a palavra grande, no contexto de avaliação de um carro, pode representar tanto uma opinião positiva, quanto uma opinião negativa (SOUZA, 2011).

Neste trabalho será utilizada esta última abordagem, onde utilizaremos a base de palavras do SenticNet. Essa base de dados será detalhada mais na frente.

2.1.3.3 Sumarização

Esta fase é focada na apresentação dos resultados, que pode ser de forma textual, ou gráfica. Os resultados gerados pelo sistema muitas vezes são complexos, cheios de informações ou difíceis de serem compreendidos. Assim, é preciso organizar essas informações para que o usuário consiga compreendê-las. A forma mais simples de sumarizar os resultados é usando textos e/ou números, aonde é mostrado algumas informações textuais e o valor para a classificação. Gráficos ou tabelas também podem ser usados, onde os resultados são sumarizados em totais e dados estatísticos, facilitando sua apresentação. Santos (SANTOS, 2010) utiliza em seu trabalho uma representação

gráfica dos resultados apresentados pelo seu sistema. A figura 1 ilustra a forma como ele exibe os resultados.

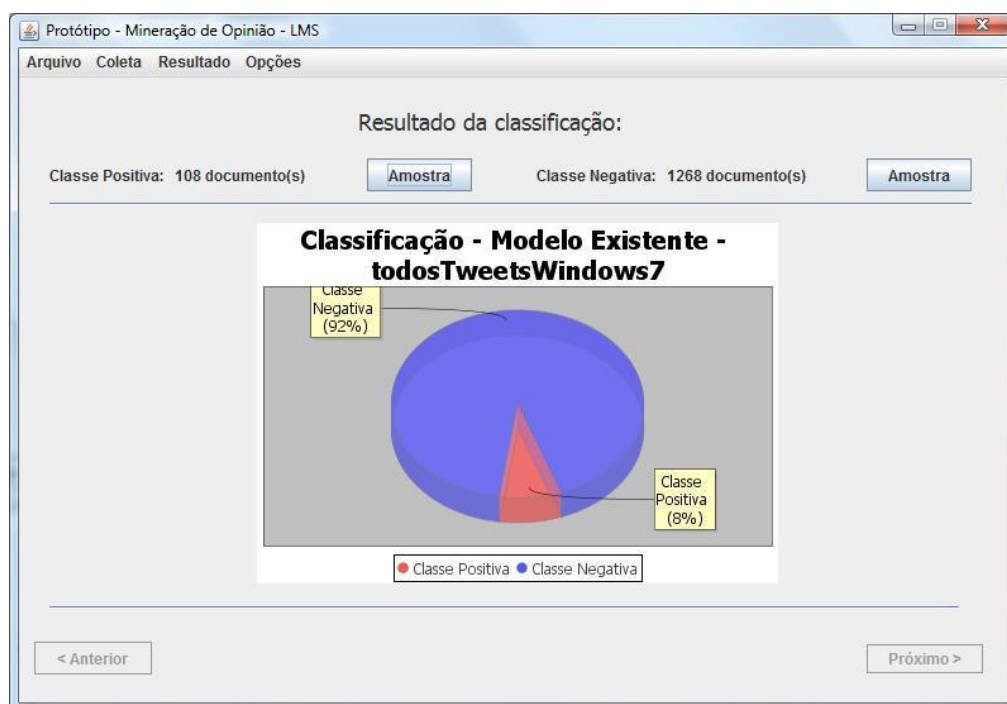


Figura 1 – Sumarização dos resultados no sistema de Santos (SANTOS, 2010).

2.2 REDES SOCIAIS

Segundo Oléques (OLÉQUES, 2010), enquanto seres sociais, as pessoas procuram constantemente interagir uns com os outros, nas mais diversas ocasiões. Esse segmento que proporciona a ampla informação a ser compartilhada por todos, sem canais reservados e fornecendo a formação de uma cultura de participação, é possível, graças ao desenvolvimento das tecnologias de comunicação e da informática, à globalização, à evolução da cidadania, à evolução do conhecimento científico sobre a vida, etc. As redes sociais virtuais são um resultado dessa necessidade de se comunicar do ser humano, aplicado às redes *web*.

De acordo com Recuero (RECUERO, 2009), rede social “é um conjunto que envolve dois elementos, os atores (pessoas, instituições ou grupos) e suas conexões”. Ou seja, na visão da autora, as pessoas se interligam através de um laço social em um determinado *site* composto por relações sociais que são formadas por interações sociais. Recuero ainda complementa:

Rede social é gente, é interação, é troca social. É um grupo de pessoas, compreendido através de uma metáfora de estrutura, a estrutura de rede. Os nós da rede representam cada indivíduo e suas conexões, os laços sociais que compõem os grupos. Esses laços são ampliados, complexificados e modificados a cada nova pessoa que conhecemos e interagimos. (RECUERO, 2009)

Rede social é uma estrutura social composta por pessoas e/ou organizações, conectadas por um ou vários tipos de relações, que dividem valores e objetivos comuns. Uma das características fundamentais na definição das redes é a possibilidade de relacionamentos horizontais e não hierárquicos entre os participantes.

É possível observar nos últimos anos, que as redes sociais tem feito um grande sucesso entre os usuários de Internet. Isso se dá pelo motivo de haver uma interação entre as pessoas, onde podem se expressar e contribuir com suas opiniões sobre o que acham sobre determinado assunto ou situação. Esse foi um diferencial que as redes sociais trouxeram, e que as tornou tão populares e acessadas (DALLACORTE, 2010).

Deste modo, com as redes sociais proporcionando interação e liberdade de expressão na Internet, as pessoas podem se conectar não somente a sites ou portais, mas também a outras pessoas. Com isso podem transferir ou agregar conhecimento sobre os mais diversos assuntos.

Com o advento da Web 2.0, e a crescente interação dos usuários com os sites, uma nova forma de cooperar e interagir surgiu. E assim como surgiu, teve uma disseminação muito rápida, juntando um número surpreendente de usuários. As denominadas redes sociais vieram para ficar, e cada vez mais atrai usuários, independente de faixa etária ou qualquer outra classificação. No ambiente Web, vêm se tornando unanimidade (DALLACORTE, 2010).

A sua importância nos dias atuais é muito grande. Com o ritmo movimentado das pessoas, o usuário acaba sendo informado por diversos assuntos, podendo conversar e entrar em grupos de debates, tudo pelas redes sociais. Além de poder ver vídeos, postar fotos e manter contatos com os amigos, sem perder o ritmo dos dias atuais.

Conforme Bouças (BOUÇAS, 2013), “quem utiliza redes sociais são pessoas que tem o hábito de acompanhar o que é publicado a seu respeito ou a respeito de alguma ideia defendida em seu espaço virtual”.

Neste ambiente, os usuários encontram espaço para criticar e elogiar pessoas, produtos, serviços e ideias com audiências que muitas vezes ultrapassam milhões de usuários. Um estudo desenvolvido pela In Press Porter Novelli e pela empresa de análise de mídia E.Life, indica diversas motivações que fazem com que um usuário adote uma rede social para interagir (BOUÇAS, 2013).

Depois de elencadas as principais motivações dos internautas para acessar uma rede social, será feita uma breve apresentação sobre o Facebook, rede social escolhida para servir como fonte de informações para este trabalho. Com isso, pretende-se analisar os fatores que levaram ao sucesso da mesma e escolha dela neste trabalho.

2.2.1 SOBRE O FACEBOOK

O Facebook é uma rede social lançada em 4 de fevereiro de 2004. Ele é apresentado como uma rede social que reúne pessoas a seus amigos e àqueles com quem trabalham, estudam e convivem. As pessoas participam do Facebook para manter contato com seus amigos, carregar um número ilimitado de fotos, compartilhar links e vídeos e aprender mais sobre as pessoas que conhecem (FACEBOOK, 2013).

Sua estrutura é diferente das outras redes de relacionamento, por isso seu layout é uma mistura de mural de escola e um caderno universitário. Logo ao acessar a página do Facebook (Fig. 2) são percebidas algumas características. A página mostra as últimas atualizações de seus amigos. Nela, qualquer coisa escrita na caixa “atualizar status” é compartilhada com os amigos. Há um bate papo que ajuda as pessoas a se comunicarem com seus amigos, familiares e colegas de trabalho com mais eficiência. O usuário participa de redes que reflitam as comunidades da sua vida real para saber mais sobre as pessoas que trabalham, vivem, ou estudam perto dele. É possível escrever uma frase, compartilhar ideias, imagens, vídeos ou até criar uma enquete.



Figura 2 – Página inicial do Facebook

Ainda, essas atualizações afetam de maneira direta qualquer organização. Se uma informação negativa sobre um mau atendimento ou funcionamento de um produto é escrita na atualização de um usuário e logo em seguida repassada a frente, uma empresa pode estar sofrendo uma perda de valor direta sem nem ao menos tomar conhecimento. Sendo assim, é extremamente necessária a atuação e monitoramento de qualquer assunto relacionado ao produto/ serviço para que se possa rapidamente sanar as dúvidas e retirar uma impressão incorreta (ROSSI, 12).

Desde sua criação, o Facebook tinha o objetivo de oferecer um espaço em que as pessoas pudessem manter laços, encontrar umas às outras, dividir opiniões e compartilhar fotografias. Como foi criada em um ambiente universitário o intuito inicial era fazer com que os acadêmicos mantivessem um relacionamento online, comentassem o que acharam da última festa, convidassem alguém para sair, terem um espaço virtual para interagir com os amigos e conhecer novas pessoas.

No Brasil, principalmente nos últimos anos houve um crescimento significativo no Facebook, onde numa pesquisa feita pelo site IG (Disponível em <http://tecnologia.ig.com.br/2012-08-03/com-146-mais-usuarios-em-um-ano-brasil-lidera-crescimento-do-facebook.html>), o país lidera o crescimento do uso da rede social Facebook em todo o mundo, com alta de 146% no total de usuários no último ano. Em junho de 2012, o País alcançou 54 milhões de usuários ou 5,6% do total de usuários da rede social em todo o mundo.

A pesquisa mostra ainda que em junho de 2012, a rede social possuía 955 milhões de usuários ativos por mês em qualquer plataforma (web e dispositivos móveis), 29% a mais que no mesmo período do ano passado (Fig. 3). Além do Brasil, Índia e Indonésia aparecem no ranking de países que apresentaram maior crescimento em número de usuários da rede social, com altas de 84% e 24%, respectivamente.

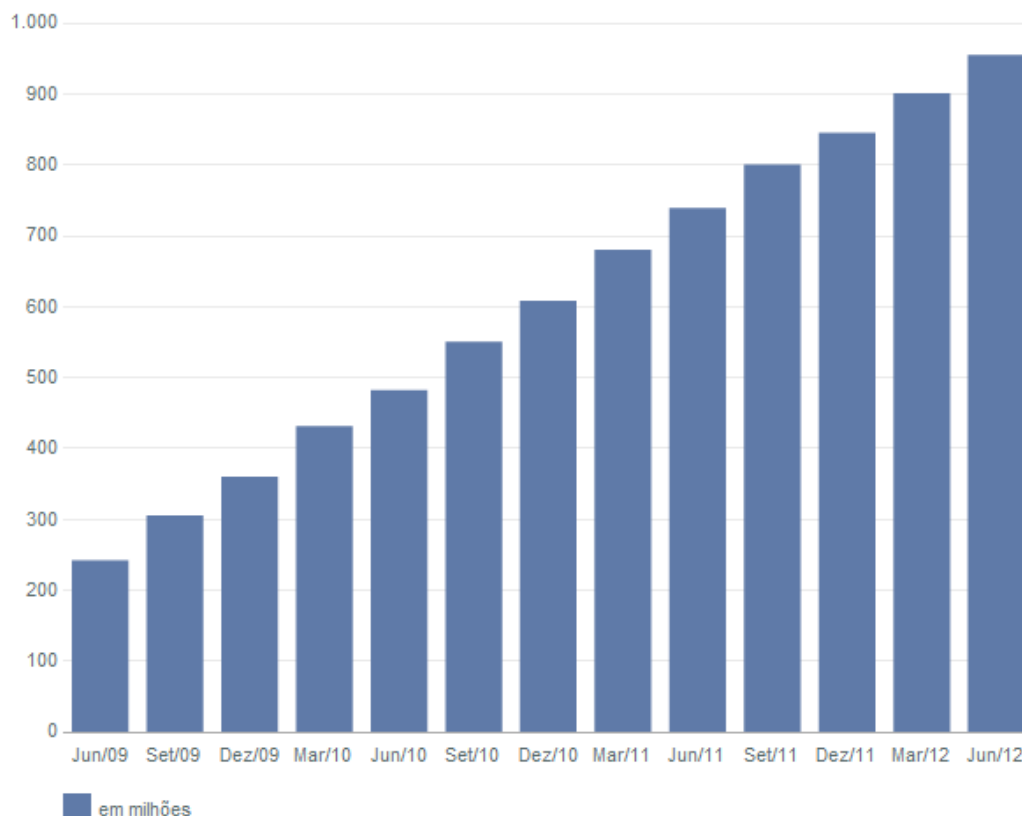


Figura 3 - Usuários ativos do Facebook.

Em pouco tempo, a rede social popularizou-se entre os jovens de todo o mundo, passou a receber investimentos de terceiros, sua página foi sofrendo mudanças e recebendo novos aplicativos que só ajudaram a enriquecê-la.

Segundo o site de estatística MonkeyBusiness (Disponível em <http://www.slideshare.net/mkbusiness/facebook-stats2011>), atualmente as pessoas gastam cerca de 700 bilhões de minutos por mês acessando o Facebook. A página se tornou o maior acervo de fotos, seu crescimento maior é por conta do acesso por meio de dispositivos móveis. Mais de 350 milhões de pessoas acessam o Facebook regularmente por meio de celulares, smartphones e tablets.

Esses foram os motivos que me levaram a escolher o Facebook como fonte de informações para o sistema desenvolvido. Essa vasta gama de informações disponíveis nesta rede social e o grande número de usuários que acessa ela favorecem na qualidade final do sistema.

2.3 SENTICNET

Segundo Cambria (CAMBRIA et al, 2010), SenticNet é um recurso semântico e afetivo disponível publicamente para o conceito de mineração de opinião e análise de sentimentos. SenticNet é construído por meio de *Sentic Computing*, um paradigma que explora tanto a IA e técnicas de Web Semântica para melhor reconhecer, interpretar e processar opiniões de linguagem natural através da web.

O SenticNet é uma iniciativa concebida no *MIT Media Laboratory*, em 2010 dentro de uma indústria de projeto de pesquisas *Cooperative Awards in Science and Engineering* (CASE), financiado pelo *UK Engineering and Physical Sciences Research Council* (EPSRC) e nasceu de uma colaboração entre a Universidade de Stirling , o *Media Lab*, e *Labs Sitekit* (PORIA et al, 2012).

Desde então, o SenticNet tem sido desenvolvido e aplicado para o projeto de emoção com reconhecimento de aplicações inteligentes em áreas que vão desde a mineração de dados até interação homem-máquina.

O principal objetivo do SenticNet é fazer com que as informações conceituais e afetiva transmitida por linguagem natural (destinado ao consumo humano) estejam mais facilmente acessíveis às máquinas. Ela busca deduzir a polaridade das palavras (chamada por ele de conceitos) associada a conceitos de senso comum e representá-la em um formato semântico. Esses conceitos podem ser compostos por uma ou várias palavras juntas como, por exemplo, o conceito “afraid” e “afraid death”. Tem sido explorado para o desenvolvimento de aplicações em campos como a mídia social, marketing e interação humano-computador. A questão básica de saber se os clientes ou eleitores estão satisfeitos com um produto ou ações podem ser respondidas através da detecção da polaridade média de seus comentários.

Com cerca de 5.700 expressões e disponível apenas em versão inglês, a versão 1.0 do SenticNet fornece uma pontuação para a polaridade na faixa entre -1 (ruim) a 1 (bom), com valores neutros igual a zero. Conceitos com valor da polaridade menores que zero, indicam que o sentimento expresso é negativo. Ao contrario, conceitos com o valor da polaridade maior que zero são aqueles que expressam algo positivo. Por exemplo, *aggravation*: -0,925 (ruim), *accomplish goal*: 0,967 (bom) e *december*: 0,111 (neutro).

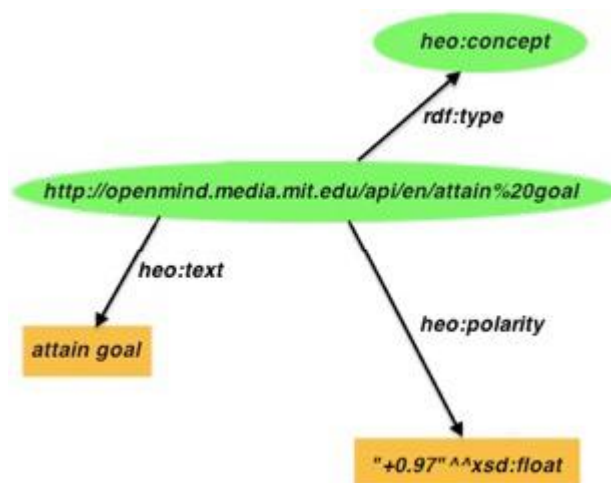


Figura 4 - Modelo de um conceito do SenticNet no formato RDF/XML.

O SenticNet é acessível em formato RDF/XML. A figura 4 mostra um conceito, modelado como um arquivo RDF. Podemos observar que para o conceito, temos o texto, que informa o nome do conceito e a polaridade, que diz o quanto aquele conceito é positivo ou negativo. O SenticNet 1.0 simplesmente associa pontuações de polaridade com um subconjunto de conceitos. O recurso pode ser baixado tanto como um todo (por exemplo, para ser incorporado em grandes sistemas *off-line*) ou seletivamente acessado online através de uma API.

O arquivo do SenticNet é estruturado da seguinte forma:

- Cada conceito possui uma *tag* RDF, que contém atributos para definir o tipo do recurso, o texto do conceito e sua polaridade.

O desenvolvimento do SenticNet foi inspirado no SentiWordNet, um recurso lexical em que cada *synset*² WordNet está associada a três valores numéricos que descrevem como são o objetivo, sentimento positivo e negativo dos termos contidos no *synset*. Cada uma das três pontuações vai de 0.0 a 1.0 e a sua soma é de 1,0 para cada *synset*. Isto significa que um *synset* pode ter pontuações diferentes de zero para todos os três categorias, o que indica que os termos correspondentes têm em cada um das três opiniões, propriedades relacionadas apenas até certo grau (ESULI et al, 2006).

² O termo *synset* constitui um conceito, e contém um conjunto de palavras, onde cada uma delas tem um sentido relacionado a um conceito (e cada uma dessas palavras é portanto sinônimo uma da outra no *Synset*).

3 DESENVOLVIMENTO

Para realização deste trabalho, foi desenvolvido um *software* para coletar e analisar o sentimento ou opinião contida em documentos da *web*, usando como fonte de informações mensagens do *Facebook*. Quanto aos objetivos, este trabalho faz uma pesquisa descritiva com a finalidade de conhecer e entender como é feita a análise de sentimentos sobre mensagens expressas por usuários na Internet e construir um *software* que, usando o SenticNet, seja capaz de expor esse sentimento através de mensagens coletadas no *Facebook*. Este *software* tem a finalidade de realizar as classificações de polaridade das mensagens coletadas. A última etapa desse trabalho é a aplicação de um formulário online com as pessoas, aonde serão comparados os resultados obtidos como o *software* e assim poder avaliar sua eficiência.

3.1 PROCEDIMENTOS METODOLÓGICOS

Primeiramente, será dada uma visão geral do que foi feito e posteriormente, a apresentação detalhada da construção e das funções que o *software* em questão possui.

3.1.1 VISÃO GERAL

Foi implementado utilizando a linguagem orientada a objetos Java, um *software* para coletar postagens da rede social *Facebook* que contenham uma palavra-chave fornecida pelo usuário, realizar todo o processamento para obtenção do sentimento expresso no texto usando o SenticNet e sumarização dos resultados obtidos, para uma melhor compreensão do usuário.

Este sistema coleta as mensagens do Facebook utilizando a API Graph, a principal API da plataforma que permite realizar leitura e escrita de dados no e a partir do *Facebook*. É passada uma palavra-chave como tema para buscar das mensagens e o sistema busca no Facebook postagens que contenham aquela palavra.

A classificação do sentimento das mensagens é feita utilizando o SenticNet v1.0, que possui informações suficientes para a realização da análise de sentimentos neste trabalho. A mensagem é quebrada em partes, onde cada palavra é analisada individualmente para que possa ser calculada a polaridade da mensagem completa.

Após coletado e classificado as mensagens, o sistema exibe para o usuário as informações geradas. A saída do sistema contém a mensagem coletada, o valor da polaridade calculada e a classificação sentimental feita com base no valor da polaridade.

Nas próximas seções serão mostrados os detalhes referentes à construção do sistema e seu uso.

3.1.2 O SISTEMA DESENVOLVIDO

Para construção do sistema, foi preciso definir alguns fatores importantes como: requisitos, arquitetura e tecnologias. Foram levantadas as funcionalidades do sistema junto ao orientador deste trabalho. Discutimos o que faria o sistema, quais os passos de cada funcionalidade e como eles seriam feitos. Daí então surgiram as funcionalidades do sistema e sua arquitetura.

O sistema possui as seguintes funcionalidades:

- Coletar mensagens do Facebook;
- Processar as mensagens para a classificação;
- Atribuir o valor da polaridade das mensagens;
- Classificar as mensagens perante o sentimento (positivo, negativo);
- Exibir uma saída para o usuário contendo o conteúdo da mensagem, valor da polaridade e o valor sentimental;

Definido todas as funcionalidades, construímos então a arquitetura do sistema. Ele é dividido em três componentes: Coleta de Dados, Classificação de sentimento e Sumarização dos resultados. A coleta de dados é responsável por toda a parte de obtenção das mensagens da rede social Facebook. É nela que acontece toda a comunicação entre a aplicação e o servidor do Facebook. Na classificação de sentimento ocorre todo o processo para classificação da mensagem, desde o processamento das mensagens até a atribuição da polaridade utilizando o SenticNet e classificação quanto ao sentimento. A sumarização é responsável por organizar as informações geradas pelo sistema e exibi-las ao usuário. A figura 5 mostra o modelo arquitetural do sistema.

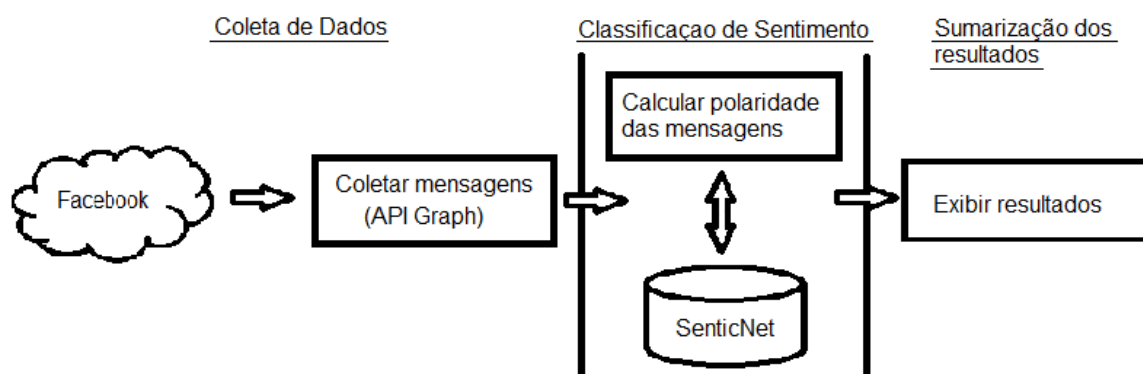


Figura 5 - Arquitetura do sistema desenvolvido.

Após todo esse processo, tiveram que ser definidas as tecnologias utilizadas no desenvolvimento. Essa etapa é muito importante para a qualidade final do sistema. E para isso, foi feito um estudo dos componentes do sistema, para que essa decisão não venha afetar o desenvolvimento e objetivos deste trabalho. Foram levados em consideração fatores como desempenho, flexibilidade, robustez.

Para a linguagem de desenvolvimento foram analisados fatores como experiência com a mesma, flexibilidade e robustez. Essas características me levaram a tomar a decisão por JAVA. Foram definidas outras tecnologias: REST, para comunicação com a API do Facebook, Apache Jena para manipulação do arquivo RDF/XML do SenticNet. Escolhemos desenvolver o sistema em desktop, pelo seu nível de complexidade baixa.

Definido todos esses pontos, partimos para o desenvolvimento do sistema. As próximas seções mostram em detalhes o desenvolvimento de cada componente.

3.1.2.1 Coleta de Dados

A aplicação tem como fonte de informações mensagens da rede social *Facebook*. Para coletar essas mensagens, é usada a API Graph³, a principal API da plataforma que permite realizar leitura e escrita de dados no e a partir do *Facebook*. É uma API com base em HTTP, que você pode usar para consulta de dados, postar novas histórias, criar *check-ins* ou qualquer uma das outras atividades que uma aplicação possa precisar fazer. A figura 6 mostra uma visao geral da arquitetura do sistema e nela é ilustrada como é feita essa fase no sistema.

A API Graph é conduzida por solicitações HTTP. Métodos HTTP tendem a mapear diretamente para ações no gráfico. *GET* para ler, *POST* para modificar e adicionar e *DELETE* para remover nós. A maneira mais fácil de entender a API Graph é

³ Disponível em <http://developers.facebook.com/docs/reference/api/>.

explora-la com o Graph API Explorer⁴. A Graph API Explorer é uma ferramenta de baixo nível que você pode usar para consultar, adicionar e remover dados. É um recurso muito útil quando se está construindo uma aplicação que utilize o *Facebook*. No Explorer, você pode selecionar o método à esquerda (GET, POST ou DELETE), o caminho na área de texto principal (/ ID) e os modificadores ao pedido (campos = id, nome), como mostra a figura 7. Estes mesmos métodos são o que você usa de qualquer cliente HTTP, incluindo o seu navegador.

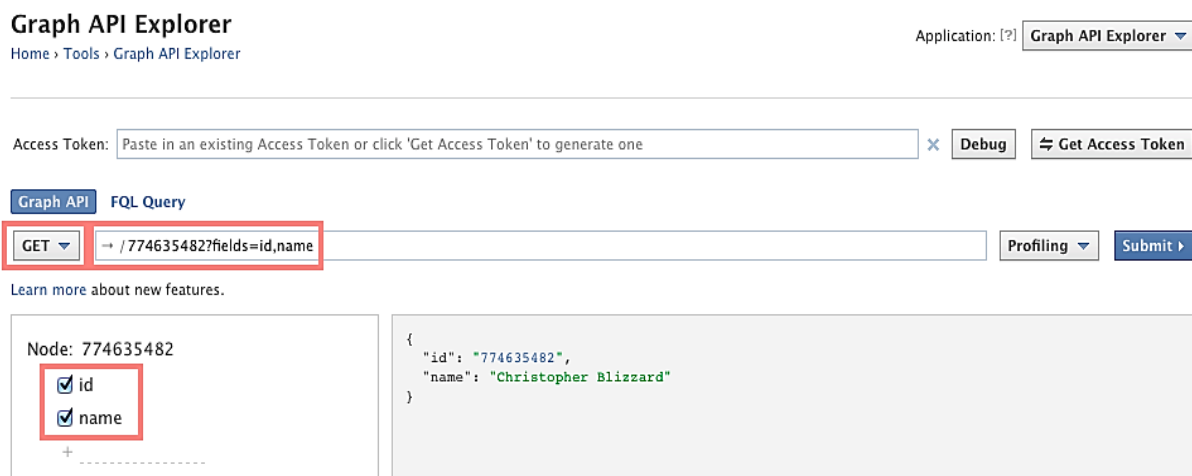


Figura 6 - Graph API Explorer

Com esta API disponibilizada pelo *Facebook* para desenvolvedores, é possível acessar qualquer recurso da rede social. Para este sistema, precisaremos apenas realizar buscas por postagens dos usuários, a fim de encontrarmos o que cada um pensa. E para realizar buscas pelo Facebook, esta API dispõe de um recurso chamado *Search*, onde é possível pesquisar sobre todos os objetos públicos no gráfico social. Aos objetos da API Graph é atribuído um ID exclusivo e eles são facilmente acessíveis usando uma URL que pode ser qualificada mais adiante para abordar um objeto/conexão específico. A estrutura geral de uma URL de objeto é a seguinte: *https://graph.facebook.com/OBJECT_ID/CONNECTION_TYPE* em que *OBJECT_ID* é o ID exclusivo e *CONNECTION_TYPE* é um dos tipos de conexão suportados pelo objeto. Por exemplo, uma página suporta as seguintes conexões: alimentação/wall, fotos, notas, postagens, membros, e assim por diante.

⁴ Disponível em <http://developers.facebook.com/tools/explorer>.

Para realizar a busca das postagens para alimentar esta aplicação, a URL utilizada foi a https://graph.facebook.com/search?q=QUERY&type=OBJECT_TYPE, onde QUERY deve ser substituído pela palavra-chave e OBJECT_TYPE pelo tipo de objeto a ser procurado, neste caso postagens. Dentro da aplicação, foi desenvolvido um cliente REST para que fosse possível realizar uma chamada GET na *url* acima e receber as mensagens como retorno. Ao executarmos essa *url* em algum navegador ou a chamamos pela aplicação, obtemos um objeto JSON⁵. Este objeto JSON retornado contém todas as informações necessárias para a esta aplicação. A representação de um JSON para uma busca no *Facebook* encontra-se abaixo (Fig. 8.):

```
"data": [
  {
    "id": "100002144190629_413932078688262",
    "from": {
      "name": "Diego Vieira Dos Santos",
      "id": "100002144190629"
    },
    "message": "Esta tragedia no Sul do Brasil, abalou o mundo inteiro, a noticia esta na primeiras paginas dos principais",
    "privacy": {
      "value": ""
    },
    "type": "status",
    "created_time": "2013-01-27T21:19:36+0000",
    "updated_time": "2013-01-27T21:19:36+0000"
  },
  {
    "id": "100002169856684_411571378925187",
    "from": {
      "name": "Joao Helio Rocha",
      "id": "100002169856684"
    },
    "message": "Todos reclamam que a Sa\u00fade vai mal, insatisfa\u00e7\u00e3o generalizada, filas intermin\u00e1veis, dif",
    "privacy": {
      "value": ""
    },
    "type": "status",
    "created_time": "2013-01-27T21:19:36+0000",
    "updated_time": "2013-01-27T21:19:36+0000"
  },
]
```

Figura 7 - Arquivo JSON de uma busca realizada utilizando o API Graph.

Podemos observar que o arquivo retorna todas as informações sobre o que foi coletado, como quem postou a mensagem (“from”), o conteúdo da mensagem (“message”), quando ela foi criada (“created_time”). Para o nosso caso, é preciso obter o valor do campo “message”, que é onde fica conteúdo da mensagem, para que a partir dela possamos calcular o valor da sua polaridade. O objeto JSON dispõe de alguns métodos para manipular as informações contidas nele. Por exemplo, o método GET é usado para recuperar algum atributo seu. Usando a figura 7 como exemplo, ao chamarmos o `objetoJSON.get(“data”)`, recuperamos a lista de postagens, que são representadas por esses elementos entre {}, logo após o “data”. De posse agora apenas

⁵ **JSON** é um formato leve para intercâmbio de dados computacionais. Para seres humanos, é fácil de ler e escrever. Para máquinas, é fácil de interpretar e gerar.

da lista de mensagens em si, iteramos entre elas e novamente utilizamos o *get* para recuperar desta vez o atributo “message” do arquivo. Fazendo isso, temos apenas as mensagens recuperadas na busca. Já com as mensagens, passamos para etapa de classificação, mostrada na próxima seção.

3.1.2.1 Classificação das mensagens

O processo de classificação das mensagens consiste em ler as informações do arquivo RDF/XML do SenticNet e calcular o valor da polaridade de cada mensagem. O sistema calcula o valor da polaridade individual de cada palavra para obter a classificação da mensagem inteira. A classificação se deu avaliando as mensagens coletadas como positivas, negativas ou não classificáveis.

As mensagens consideradas positivas foram aqueles cujos autores apresentavam um posicionamento claramente favorável ao tema ou entidade pesquisada. O mesmo para mensagens negativas, porém com posicionamento contra. Já as não classificáveis foram aquelas onde nenhuma das palavras da mensagem foi encontrada no dicionário do SenticNet.

O sistema quebra as mensagens por palavras e realiza uma busca no arquivo RDF/XML do SenticNet afim de procurar o valor da polaridade daquela palavra. Cada conceito no dicionário do SenticNet está vinculado ao valor que corresponde a sua polaridade. A figura 9 ilustra um termo do SenticNet expresso dentro do arquivo RDF/XML.

```
<rdf:Description rdf:about="http://openmind.media.mit.edu/api/en/concept/unlove">
  <rdf:type rdf:resource="http://sentic.net/api/concept"/>
  <text xmlns="http://sentic.net/api/">unlove</text>
  <polarity xmlns="http://sentic.net/api/"
    rdf:datatype="http://www.w3.org/2001/XMLSchema#float">-0.916</polarity>
</rdf:Description>
```

Figura 8 - Exemplo de um recurso no SenticNet

Nesse exemplo temos valores para o conceito “unlove”. Podemos perceber que ele significa algo muito ruim, pois o valor da polaridade é negativo = -0.916.

Observando a figura 8, é possível definir a estrutura completa do RDF/XML do SenticNet:

- 1) Na primeira linha existe uma *tag* que identifica que ele é um arquivo baseado na estrutura *xml*. (`<?xml version="1.0" encoding="UTF-8"?>`)

- 2) A segunda linha contém a *tag* que identifica que ele é um arquivo RDF. (Ex.: `<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">`)
- 3) A terceira linha contém uma visão geral do recurso descrito. Serve apenas para ajudar na identificação. (Ex.: `<rdf:Description
rdf:about="http://openmind.media.mit.edu/api/en/concept/unlove">`)
- 4) A quarta linha contém a definição do recurso a ser descrito. No caso do SenticNet, o recurso é sempre o conceito, que é a palavra ou expressão que procuramos. (Ex.: `<rdf:type rdf:resource="http://sentic.net/api/concept"/>`)
- 5) A quinta linha contém uma das propriedades do conceito, que é nome do recurso e o seu valor. (Ex.: `<text xmlns="http://sentic.net/api/">unlove</text>`)
- 6) A sexta linha contém a outra propriedade do conceito. Neste caso, a polaridade do conceito e o seu valor. (Ex.: `<polarity xmlns="http://sentic.net/api/"
rdf:datatype="http://www.w3.org/2001/XMLSchema#float">-0.916</polarity>`)
- 7) As linhas seguintes do arquivo são repetições das linhas 3, 4, 5 e 6, respectivamente, cada uma para um conceito diferente.

Para manipular esse arquivo RDF/XML do SenticNet, utilizamos o Apache Jena (versão 2.7.4), um *framework* Java de código aberto para criar, armazenar, manipular e consultar em arquivos RDF/XML. O Jena oferece suporte para manipulação de arquivos RDF. Com ele é possível dentro do software desenvolvido, ler o arquivo RDF do SenticNet e buscar cada termo contido e o valor da sua polaridade. Desta forma, é possível a partir de uma mensagem capturada do *Facebook*, classificar qual a sua polaridade.

De posse das mensagens coletadas, a classificação consiste em quebrar a mensagem em palavras e realizar uma busca palavra por palavra na base do SenticNet afim de encontrar o valor da polaridade de cada. Ao terminar a varredura nas palavras é feito um somatório dos valores das polaridades das encontradas e calculado uma média, aonde essa média diz qual a polaridade da mensagem.

O seguinte algoritmo foi utilizado no sistema para calcular a polaridade da mensagem:

$Polarity = \sum_1^n P_i$, onde P_i é a polaridade individual de cada palavra e n é o tamanho do conjunto que contém as polaridades individuais.

$$Polarity\ final = \frac{Polarity}{n}$$

A partir desse valor, classificamos as mensagens da seguinte forma:

- 1) Para Score final > 0,00, a mensagem expressa algo “Positivo”.
- 2) Para Score final < 0,00, a mensagem expressa algo “Negativo”.
- 3) Para Score final = NaM, a mensagem não pode ser classificada.

3.1.2.2 Sumarização dos resultados

Essa etapa é responsável por exibir ao usuário o resultado de todo o processo de captura e análise do sentimento das mensagens feito pelo sistema. Estudamos algumas formas para exibir esses dados, como por exemplo, textos, gráficos ou tabelas. Foi analisado o grau de complexidade das informações geradas para decidir como elas seriam mostradas.

Quando levantadas as funcionalidades no início da construção do sistema, o objetivo era mostrar para o usuário o conteúdo da mensagem capturada, o valor da polaridade calculada e o valor sentimental da mensagem. Assim, organizamos essas informações em modo textual mesmo, sem nada complexo ou embaralhado. Foi definido um padrão para exibir a saída do sistema, que pode ser visualizado abaixo na figura 9. A primeira linha contém o conteúdo da mensagem. Já na segunda é exibida uma mensagem “A polaridade da mensagem é de:” seguido do valor calculado pelo sistema. E por fim a mensagem “A mensagem expressa algo:” seguido da classificação sentimental da mensagem.

In America we should always look to promote opportunity for one to move up. Obama is that kind of guy. He became president from a apartment in Chicago and from a college loan. He gets it. He's a product of it.

A polaridade da mensagem é de: 0.49949995

A mensagem expressa algo: positivo

He did what i never knew would have been possible in my life so, i have every reason to praise Him and He's worthy to be praised

A polaridade da mensagem é de: 0.35842854

A mensagem expressa algo: positivo

Figura 9 – Exemplo de saída do sistema

4 EXPERIMENTOS E RESULTADOS

Nesse capítulo serão mostrados os resultados obtidos, a metodologia utilizada para realização dos experimentos e uma breve análise dos resultados.

4.1 METODOLOGIA PARA EXPERIMENTO

Como foi dito anteriormente, a base de dados usada para os experimentos foi as postagens de usuários do Facebook coletadas pelo sistema, para que fossem realizadas comparações com os resultados obtidos pela avaliação aplicada com as pessoas.

A avaliação consistiu num questionário com 27 mensagens coletadas através do sistema desenvolvido. Esse questionário foi criado e disponibilizado para alunos e professores do Campus IV da Universidade Federal da Paraíba de forma *online*, através de um documento criado pelo Google Drive. Para cada uma dessas mensagens, foi pedido para que as pessoas respondessem se achavam que ela se tratava de algo bom(positivo) ou ruim(negativo). As mensagens que receberam mais votos positivos foram consideradas como positivas e as que receberam mais votos negativos foram considerados negativos. O resultado da classificação das mensagens classificadas pelo sistema foi armazenado em um documento para que fosse possível realizar as comparações com os resultados obtidos através do questionário. O documento consiste de todas as mensagens, com suas respectivas polaridades. Cada classificação no documento é expressa da seguinte forma (Fig. 10):

did anybody else hear about the meteor that hit russia yesterday?

A polaridade da mensagem é de: -0.103

A mensagem expressa algo: negativo

Figura 10 – Exemplo de classificação pelo sistema

O primeiro parágrafo identifica o conteúdo da postagem, ou seja, a mensagem. O segundo identifica o valor da polaridade da mensagem atribuída pelo sistema. O terceiro representa a classificação sentimental da mensagem, que é feita com base no valor da polaridade. Todos esses critérios são fundamentais para realização do estudo comparativo.

Por limitações do SenticNet, o sistema consegue classificar mensagens apenas no idioma inglês. Então foi disponibilizada no questionário uma tradução feita através do

Google Tradutor para ajudar na classificação e sanar as dúvidas. Essa tradução não reflete perfeitamente o mesmo sentido que foi dado na mensagem original, mas se assumiu que é uma boa aproximação. Para cada mensagem, existiam apenas duas alternativas para resposta, eram elas: positivo e negativo. O documento com o formulário aplicado pode ser visualizado no Apêndice B. A figura 11 apresenta o modelo das perguntas na avaliação:

The Federal Gov't shut down for President's Day, and the oil companies jacked the price of gas up 20 cents today, price went from \$3.79.9 to \$3.99.9 in the last couple of hours, WOW!
O GOVERNO FEDERAL fechada para Dia do Presidente, e as companhias petrolíferas levantado o preço do gás até 20 centavos hoje, o preço passou de \$ 3.79.9 para \$3.99.9 no último par de horas, WOW!

☒ Positivo
☐ Negativo

Figura 11 – Exemplo de pergunta da avaliação.

Na próxima seção serão mostrados os resultados e a análise feita a partir deles. Serão mostrados também, alguns resultados anteriores, que resultaram na melhora do sistema.

4.2 RESULTADOS E DISCUSSÃO

Como o principal objetivo do trabalho está na criação de um sistema de análise de sentimentos em mensagens do Facebook utilizando a abordagem do SenticNet e consequentemente a avaliação desse sistema, para que o mesmo possa mostrar resultados mais próximos da classificação humana, os resultados desse trabalho foram obtidos a partir da comparação dos resultados obtidos pelo sistema desenvolvido e os obtidos através do questionário.

As mensagens capturadas pelo sistema foram realizadas através de buscas por assuntos como “obama”, “meteor”, “samsung”, “iphone”, “life”, “windows”, etc. As mensagens capturadas e suas respectivas classificações atribuídas pelo sistema foram guardadas em um documento, que serviu posteriormente para elaboração da avaliação e do estudo comparativo. Esse documento encontra-se no Apêndice A deste trabalho. Essas mensagens passaram por um processo de análise manual uma a uma para que pudessem ser selecionadas apenas aquelas que traziam algum tipo de informação significativa sobre o tema, visto que o conteúdo retornado do Facebook é muito grande e

contem muitas mensagens que representam apenas fatos. Essa análise manual foi importante para eliminar classificações sobre mensagens irrelevantes e poluir o resultado final deste trabalho.

Foram selecionadas então 27 mensagens para fazerem parte do estudo. A partir delas foi criado o formulário para classificação dos sentimentos, no qual as pessoas possam dar sua opinião sobre o sentimento de cada mensagem. O formulário online com as perguntas da avaliação foi disponibilizado para respostas entre os dias 22 de fevereiro e 05 de março deste mesmo ano. Durante esse período foram recebidas 36 respostas.

Os resultados apresentados pelo sistema se mostraram bastantes compatíveis com os apresentados pela avaliação. Na classificação feita pelo sistema, 17 mensagens foram classificadas com sendo positivo e 10 como negativas. A Tab. 1 apresenta os resultados da classificação das 27 mensagens pelo sistema.

Tabela 1 – Resultado da classificação realizada através do sistema.

Conjunto	Quantidade	%
Positivos	17	62,96
Negativos	10	37,04

Já na classificação feita pelas pessoas através da pesquisa, 11 mensagens tiveram um percentual maior de resposta para a classificação positivo e 15 mensagens tiveram um percentual de resposta maior para classificação negativa, tendo ainda uma mensagem que obteve empate sobre a opinião, ficando com 50% das opiniões positivas e 50% negativas. As mensagens são consideradas positivas quando a maioria das respostas por positiva e negativa quando a maioria das respostas dada a ela for negativa. A Tab. 2 apresenta os resultados da classificação das mensagens feita através da avaliação. Os resultados da avaliação aplicada com as pessoas encontram no Apêndice C.

Tabela 2 – Resultado da classificação realizada através da pesquisa.

Conjunto	Quantidade	%
Positivos	11	40,74
Negativo	15	55,56

Ao compararmos os resultados, pode-se perceber que o sistema conseguiu uma taxa de acerto de 70,4%, que corresponde a 19 das 27 mensagens. A tabela 3 apresenta os resultados das classificações das mensagens feita através do sistema e da avaliação.

Tabela 3 – Tabela comparativa dos resultados

Mensagem	Classificação do sistema	Classificação da avaliação
1	Positivo	Negativo
2	Negativo	Negativo
3	Negativo	Negativo
4	Positivo	Positivo
5	Positivo	Negativo
6	Positivo	Positivo
7	Positivo	Positivo
8	Positivo	Incerteza
9	Positivo	Negativo
10	Positivo	Negativo
11	Positivo	Positivo
12	Positivo	Positivo
13	Positivo	Positivo
14	Positivo	Positivo
15	Positivo	Positivo
16	Negativo	Negativo
17	Positivo	Negativo
18	Positivo	Positivo
19	Positivo	Negativo
20	Positivo	Positivo
21	Negativo	Negativo
22	Negativo	Negativo
23	Negativo	Positivo
24	Negativo	Negativo
25	Negativo	Negativo
26	Negativo	Negativo
27	Negativo	Negativo

Essas 19 mensagens apresentaram valores bem próximos de classificação nos dois meios, sendo algumas com uma diferença um pouco para mais ou para menos, mas ainda se enquadrando como acerto para o sistema, o que mostra que o sistema desenvolvido é eficiente nas classificações. As figuras 12 e 13 mostram um resultado de uma classificação acertada pelo sistema. A figura 12 mostra a classificação dada pelas pessoas para certa mensagem. Já a fig. 13 é o resultado desta mesma mensagem, agora classificada pelo sistema. Nota-se que os resultados foram bem próximos sobre o sentimento em ambas as classificações, o que mostra um ponto positivo para o sistema.

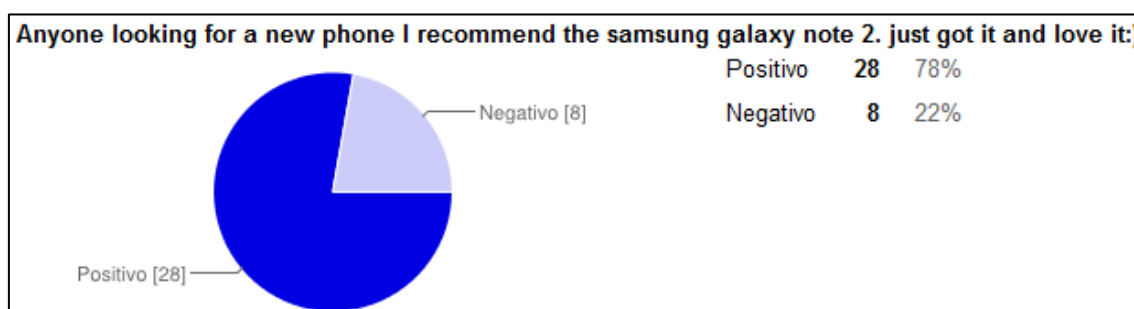


Figura 12 – Resultado da classificação através da avaliação para uma mensagem

Anyone looking for a new phone I recommend the samsung galaxy note 2. just got it and love it:)

A polaridade da mensagem é de: 0.3725

A mensagem expressa algo: positivo

Figura 13 – Resultado da classificação através do sistema para uma mensagem

Um total de oito mensagens apresentaram valores divergentes, sendo que uma obteve incerteza pela classificação das pessoas, recebendo 50% das classificações positivas e 50% de negativas, mas o sistema proposto a classificou como positiva. A mensagem mostrava que o seu ator queria morrer, mas quando ele se lembrava de uma pessoa muito especial, falava que não podia fazer isso, que não podia morrer, que a amava muito. Percebemos que a classificação vista de pontos de vistas diferentes geram classificações diferentes. Isso causou a incerteza nas pessoas. Mas o sistema considerou a visão de que ele não podia morrer e a classificou como positiva.

No geral, essas divergências foram classificações atribuídas pelo sistema como positiva, mas dadas como negativa pelas pessoas no questionário e vice-versa. Algumas

delas mostraram valores até próximos do acerto para o sistema, como mostra as figuras 14 e 15. As pessoas classificaram a mensagem como sentimento negativo, mas o valor da polaridade atribuído pelo sistema para a mesma mensagem foi um valor bem próximo a um valor negativo.

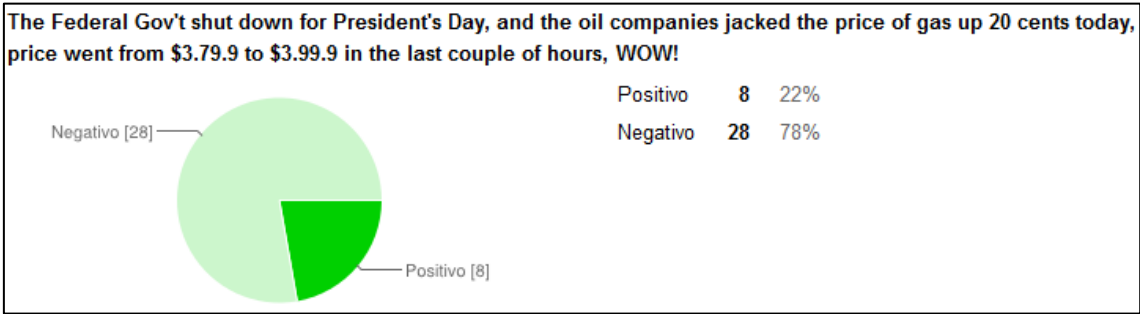


Figura 14 - Resultado da classificação através da avaliação para uma mensagem

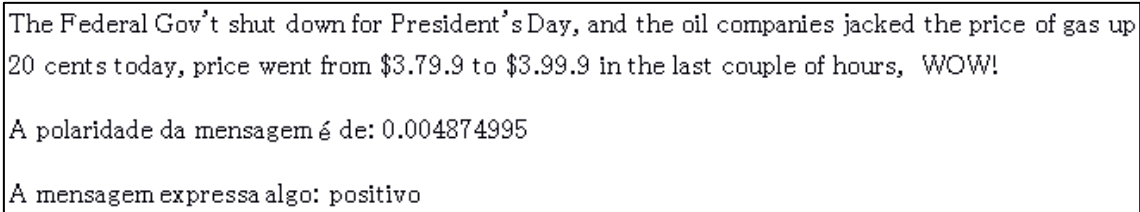


Figura 15 - Resultado da classificação através do sistema para uma mensagem

Diante dos resultados apresentados, podemos observar que o sistema se mostrou eficiente nas classificações das mensagens. Uma taxa de acerto de 70,4% é um valor bem proximo do esperado, que é se aproximar o maximo dos 100%. As mensagens que foram classificadas corretamente pelo sistema tiveram valores para a polaridade bem proximas do real. Algumas classificações foram acertadas por pouco pelo sistema. Tivemos tambem algumas mensagens em que a diferença de votos nas duas classificacoes foi pequena, mas em outros casos tiveram valores muito distantes. Algumas palavras em que valor para a polaridade seja muito forte, tanto positiva ou negativa, tiveram grande poder de decisão na classificação final da mensagem. Esses e outros fatores são importantes ser levados em consideração na hora de realizar ajustes no algoritmo e conseguir se aproximar o maximo do ideal.

5 CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho criou e avaliou um sistema que coleta mensagens do *Facebook* e, utilizando o SenticNet, classifica o sentimento delas de acordo com a sua polaridade. Através de um questionário, foi possível comparar os resultados e medir a eficiência do sistema, obtendo assim um resultado positivo para este trabalho.

Foi desenvolvido um sistema em Java para análise de sentimentos, que coleta mensagens do Facebook através da sua API disponibilizada para desenvolvedores e as classifica como positivo ou negativo, de acordo com o valor da polaridade, usando a técnica de seleção de palavras e o SenticNet como recurso para esse processo.

Após desenvolvimento da aplicação, foi necessário medir sua eficiência. E para isso foi realizado através do mesmo uma coleta de postagens no Facebook sobre temas como, por exemplo, “obama”, “meteor”, “samsung”, “iphone”, “windows” e realizado a classificação dessas mensagens. A partir dessas mensagens, foi montado e aplicado *online* um questionário para que as mensagens pudessem também ser classificadas por pessoas. Os resultados foram comparados e chegou a uma taxa de acerto favorável para o sistema. Dentre algumas limitações enfrentadas neste estudo e já mostradas no capítulo 2 deste trabalho, a taxa de acerto do sistema foi de 70,4%. Tivemos classificações que mostraram valores diferentes nos dois métodos, mas em alguns casos por diferenças pequenas ou por se tratarem de mensagens confusas, o que conclui que a taxa de erro do sistema poderá ser diminuída após uns ajustes no sistema.

Podemos concluir que este trabalho conseguiu atingir seus objetivos e que os resultados obtidos mostram que o mesmo é capaz de refletir os sentimentos expressos em mensagens postadas na rede social Facebook.

Como trabalhos futuros, destaca-se a busca por melhorias e ajustes no algoritmo desenvolvido que possam apresentar melhores resultados aos apresentados neste trabalho. Como o presente trabalho analisou apenas uma avaliação na rede social Facebook, sugere-se, também, a análise envolvendo mais redes sociais. Assim teríamos uma análise mais ampla das opiniões na Internet. Recomenda-se em trabalhos futuros também a uma avaliação com um numero maior de participantes. Assim teríamos uma precisão diferente do apresentado e talvez ate melhor do sistema.

REFERÊNCIAS BIBLIOGRÁFICAS

- [MOREIRA et al, 2009] MOREIRA, D.S.;DIAS,M.S.; **WEB 2.0 – A Web Social**. Revista CEPPG – CESUC, Nº 20 – 1/2009, p.196 à 208.
- [PANG et al, 2002] PANG, B.;LEE, L.;VAITHYANATHAN, Shivakumar. **Thumbs up? Sentiment Classification using Machine Learning Techniques**. Proceedings of EMNLP 2002, pp. 79–86.
- [SOUZA, 2011] SOUZA, L. V. **Análise de sentimentos no Twitter utilizando o SentiWordNet**. Pernambuco, 2011, 29p. Monografia (Bacharel em Ciências da Computação). Centro de Informática – Universidade Federal de Pernambuco.
- [SIQUEIRA, 2010] SIQUEIRA, H. **WhatMatter: Extração e visualização de características em opiniões sobre serviços**. Dissertação (Mestrado em Ciência da Computação) - Centro de Informática/UFPE. Recife. 2010.
- [LIU, 2010] LIU, B. **“Handbook of Natural Language Processing – Sentiment Analysis and Subjectivity”**. Department of Computer Science, University of Illinois at Chicago, 2010.
- [SANTOS, 2010] SANTOS, L. M. **Protótipo para mineração de opinião em redes sociais: estudo de casos selecionados usando o Twitter**. Lavras, 2010, 103p. Monografia (Bacharel em Ciencias da Computação) – Universidade Federal de Lavras.
- [RODRIGUES et al, 2012] RODRIGUES, C. A. S.; VIEIRA, L. L.; MALAGOLI, L. L.; TIMMERMAN, N. **Mineração de Opinião / Análise de Sentimentos**. Universidade Federal de Santa Catarina, Santa Catarina, 2012, 6p.
- [CARRILHO JUNIOR et al, 2007] CARRILHO JUNIOR, João Ribeiro; PASSOS, Emmanuel Pisces Lopes (Orientador). **Desenvolvimento de uma Metodologia para Mineração de Textos**. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2007, 98p.
- [RECUERO, 2009] RECUERO, Raquel. **Redes sociais na internet** / Raquel Recuero. – Porto Alegre: Sulina, 2009. (Coleção Cibercultura)191 p.
- [BOUÇAS, 2013] BOUÇAS, Cibelle. **Redes sociais atraem interesse de companhias**. Disponível em: <http://www.fenacon.org.br/>. Acesso em: 23 jan. 2013.
- [DALLACORTE, 2010] DALLACORTE, André Luiz. **Protótipo de rede social voltada para a conscientização ambiental baseada nos princípios da web 2.0**. Unochapecó, Chapecó, 2010.
- [OLÉQUES, 2010]OLÉQUES, Jussara Fernandes. **Transvendo o mundo em redes sociais e comunidades de aprendizagem**. Porto Alegre, 2010. 4f. Artigo – Escola Municipal Vila Monte Cristo.
- [FACEBOOK, 2012] FACEBOOK. **Site Oficial do Facebook**. Disponível em: <http://www.facebook.com/>. Acesso em: 21 jan. 2012.

[LOPES, 2009] LOPES, B. O.; **A linguagem dos blogs e as redes sociais virtuais.** Monografia de Tecnólogo – Faculdade de Tecnologia da Zona Leste, São Paulo, 2009, 47p.

[CAMBRIA et al, 2010] CAMBRIA, E., SPEER, R., HAVASI, C., HUSSAIN, A.: **SenticNet: A publicly available semantic resource for opinion mining.** In: AAAI CSK, pp. 14–18. Arlington (2010).

[PORIA et al, 2012] PORIA Soujanya, GELBUKH Alexander, CAMBRIA Erik, DAS Dipankar, BANDYOPADHYAY Sivaji. **Enriching SenticNet Polarity Scores Through Semi-Supervised Fuzzy Clustering.** Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference, (2012)

[ESULI et al, 2006] ESULI, A., SEBASTIANI, F.: **SentiWordNet: A publicly available lexical resource for opinion mining.** In: LREC (2006)

[SANTANA, 2013] SANTANA, O.G. **Porque Java?** DevMedia, Bahia, Agos. 2012. Canal Java. Disponível em < <http://www.devmedia.com.br/por-que-java/20384>> Acesso em: 24 jan. 2013.

[LIMA, 2005] LIMA, J.C.; CARVALHO, C.L.; **Resource Description Framework (RDF).** Instituto de Informática, Universidade Federal de Goiás, 2005.

[RICHARDSON et al, 2007] RICHARDSON, L.; RUBY, S. **RESTful Web Services.** 1.ed. United States of America: O'Reille Media, 2007.

[SALUSTIANO, 2013] SALUSTIANO, Sergio. **Monitoramento das redes sociais: Muito mais que uma análise de sentimentos.** Disponível em: <<http://www.slideshare.net/skrol/monitoramento-de-redes-sociais>>. Acesso em: 12 de fev. 2013

[LIU, 2010] LIU B. **Sentiment analysis: A multi-faceted problem.** IEEE Intell Syst. 2010; 25:1-5.

[PANG et al, 2008] PANG B, LEE L. **Opinion mining and sentiment analysis.** Found Trends Inf Retr. 2008;2(1-2):1-135.

[SEBASTIANI, 2002] SEBASTIANI, S. Machine learning in automated text categorization. **ACM Computing Surveys.** New York - USA, v. 34, Issue 1, p. 1-47, Março 2002.

[TORRES, 2009] TORRES, Cláudio. **A Bíblia do marketing digital:** tudo o que você queria saber sobre marketing e publicidade na internet e não tinha a quem perguntar. São Paulo: Novatec, 2009.

[KIRKPATRICK, 2011] KIRKPATRICK, David. **O efeito Facebook: os bastidores da história da empresa que conecta o mundo.** Rio de Janeiro: Intrínseca, 2011.

[ROSSI, 2012] ROSSI, Filipe. **As mídias sociais na empresa: um estudo comparativo de casos sobre a utilização do Facebook nas empresas.** São Paulo: FACECAP, 2012.

APÊNDICE

Apêndice A – Resultado das mensagens capturadas e classificadas pelo sistema desenvolvido.

meteor about to hit earth. say you've got 24 hours to think about it. what would be your last song request?

A polaridade da mensagem é de: 0.2405

A mensagem expressa algo: positivo

did anybody else hear about the meteor that hit russia yesterday?

A polaridade da mensagem é de: -0.103

A mensagem expressa algo: negativo

Pretty pathetic when they can't even understand the difference between an asteroid, a meteor and a meteorite.

A polaridade da mensagem é de: -0.055999998

A mensagem expressa algo: negativo

Pass this on. Put pressure on your Senators to remove all petroleum products from the stock market. This is how control is done. There is no other reason for these products on the "speculation" list. Nothing but greed. PASS THIS ON, WE CAN MAKE A CHANGE FOR OUR COUNTRY AND PEOPLE.

A polaridade da mensagem é de: 0.0976

A mensagem expressa algo: positivo

The Federal Gov't shut down for President's Day, and the oil companies jacked the price of gas up 20 cents today, price went from \$3.79.9 to \$3.99.9 in the last couple of hours, WOW!

A polaridade da mensagem é de: 0.004874995

A mensagem expressa algo: positivo

In America we should always look to promote opportunity for one to move up. Obama is that kind of guy. He became president from a apartment in Chicago and from a college loan. He gets it. He's a product of it.

A polaridade da mensagem é de: 0.49949995

A mensagem expressa algo: positivo

He did what i never knew would have been possible in my life so, i have every reason to praise Him and He's worthy to be praised

A polaridade da mensagem é de: 0.35842854

A mensagem expressa algo: positivo

i hate this life i wish i die but when i remember you i say how i can leave her i can't die i love you a lot m...ay

A polaridade da mensagem é de: 0.2526

A mensagem expressa algo: positivo

i laugh at people who put their life story's on here and beg for sympathy from it!

A polaridade da mensagem é de: 0.402
A mensagem expressa algo: positivo

My brain hurts now! Man almost crossed the center median on capital Circle and if he did he would've hit me now I can't think. It hurt my brain. My life flashed before my eyes. That's why my brain hurts. My brain hasn't moved that fast in years.
A polaridade da mensagem é de: 0.061166663
A mensagem expressa algo: positivo

What a beautiful day!!!! I love my life and kids.
A polaridade da mensagem é de: 0.3516667
A mensagem expressa algo: positivo

What a woman! Just reading about her life inspires me! I have read a little of her stuff and highly recommend it...I'll post the titles below, just in case one grabs you. Library probably has them all...
A polaridade da mensagem é de: 0.205
A mensagem expressa algo: positivo

Anyone looking for a new phone I recommend the samsung galaxy note 2.just got it and love it:)
A polaridade da mensagem é de: 0.3725
A mensagem expressa algo: positivo

I am sure by now...everyone has a tablet?! Well I don't but was wondering what everyones Favorite Tablet is so far? From Ipad to Samsung to Windiws 8? I am looking to get one for our business. Just looking for suggestions!?
A polaridade da mensagem é de: 0.315625
A mensagem expressa algo: positivo

my life leads only be happy with the samsung galaxy s 3 :D b
A polaridade da mensagem é de: 0.37433335
A mensagem expressa algo: positivo

Seriously considering throwing Roxanne Lacasse's new computer out the window!!!!
RRRRRRRRRAAAAAWWWWRRRRRRRR!!!!!!
Windows 8 is so bad that I'm begging for vista back! YUK!
A polaridade da mensagem é de: -0.003999993
A mensagem expressa algo: negativo

I thought these android tablets were flash player compatible? How come adobe is telling me its not? Ready to go back to best buy and throw my new tablet at them! Ugh, totally frustrated
A polaridade da mensagem é de: 0.30033335
A mensagem expressa algo: positivo

I like my new galaxy tablet. Still learning how to use it.
A polaridade da mensagem é de: 0.39725

A mensagem expressa algo: positivo

Since I uploaded my iPhone to iOS 6.1, I started to discover some strange WiFi usage has been used. That's odd, very odd. >:\

A polaridade da mensagem é de: 0.23866665

A mensagem expressa algo: positivo

The only promise in life ... Is death ... So live ...Before you die....

A polaridade da mensagem é de: 0.32033333

A mensagem expressa algo: positivo

I AM BORED TO DEATH. CAN'T DO ANYTHING CUZ I HURT SO BAD.

A polaridade da mensagem é de: -0.135

A mensagem expressa algo: negativo

Lol He Gets Mad When Were Talking On The Phone & I Randomly Play Justin Bieber Songs cx

A polaridade da mensagem é de: -0.796

A mensagem expressa algo: negativo

i will never abandon anybody I love...we might drift apart but ill never abandon you.

A polaridade da mensagem é de: -0.448

A mensagem expressa algo: negativo

There now you went and did it puppy, you brought out the tears in my eyes! I hate people who abandon their pets!

A polaridade da mensagem é de: -0.237

A mensagem expressa algo: negativo

So I'm watching freddy vs jason and my 4 yr old loves the crap out of scary movies should I be afraid!!!!

A polaridade da mensagem é de: -0.3635

A mensagem expressa algo: negativo

Damm everywhere you turn theres danger everywhere even here when i'm on Facebook someone could look me up on google. scary

A polaridade da mensagem é de: -0.4795

A mensagem expressa algo: negativo

51% percent of America voted for Obama the 49% percent of people who arent stupid voted against him...-Herman Cain on Hannity tonight

A polaridade da mensagem é de: -0.691

A mensagem expressa algo: negativo

Apêndice B – Avaliação aplicada com as pessoas para classificação das mensagens.

Avaliação sobre análise sentimentos em redes sociais

Olá, esse é o formulário de avaliação criado para experimento do TCC do aluno George Alcântara sobre análise de sentimentos em redes sociais.

Responda as questões abaixo com bastante atenção, pois a sua opinião é muito importante para a pesquisa. Abaixo encontram-se mensagens públicas retiradas do Facebook, você só precisa ler a mensagem e responder se considera ela uma mensagem com sentimento positivo ou negativo. As mensagens estão todas em inglês, mas abaixo de cada uma delas você pode encontrar a tradução feita pelo Google Translate. O resultado não é perfeito, mas pode ajudá-lo a compreender o sentido da mensagem.

meteor about to hit earth. say you've got 24 hours to think about it. what would be your last song request?

meteoro prestes a atingir a Terra. dizer que você tem 24 horas para pensar sobre isso. o que seria o seu pedido última música?

- ☐ Positivo
- ☐ Negativo

did anybody else hear about the meteor that hit russia yesterday?
se ninguém ouvir sobre o meteoro que atingiu a Rússia ontem?

- ☐ Positivo
- ☐ Negativo

Pretty pathetic when they can't even understand the difference between an asteroid, a meteor and a meteorite.

Patético quando não pode nem mesmo entender a diferença entre um asteróide, um meteoro e um meteorito.

- ☐ Positivo
- ☐ Negativo

Pass this on. Put pressure on your Senators to remove all petroleum products from the stock market. This is how control is done. There is no other reason for these products on the "speculation" list. Nothing but greed. PASS THIS ON, WE CAN MAKE A CHANGE FOR OUR COUNTRY AND PEOPLE.

Passa esta mensagem. Exercer pressão sobre seus senadores para remover todos os produtos petrolíferos do mercado de ações. Esta é a forma como o controle é feito. Não há outra razão para estes produtos na "especulação" lista. Nada mas a ganância. Passe este, podemos fazer uma mudança para o nosso país e as pessoas.

☐ Positivo

☐ Negativo

The Federal Gov't shut down for President's Day, and the oil companies jacked the price of gas up 20 cents today, price went from \$3.79.9 to \$3.99.9 in the last couple of hours, WOW!

O GOVERNO FEDERAL fechada para Dia do Presidente, e as companhias petrolíferas levantado o preço do gás até 20 centavos hoje, o preço passou de \$ 3.79.9 para \$3.99.9 no último par de horas, WOW!

☐ Positivo

☐ Negativo

In America we should always look to promote opportunity for one to move up. Obama is that kind of guy. He became president from a apartment in Chicago and from a college loan. He gets it. He's a product of it.

Nos Estados Unidos, deve sempre olhar para promover uma oportunidade para mover-se. Obama é esse tipo de cara. Ele tornou-se presidente de um apartamento em Chicago e de um empréstimo da faculdade. Ele recebe-lo. Ele é um produto dela.

☐ Positivo

☐ Negativo

He did what i never knew would have been possible in my life so, i have every reason to praise Him and He's worthy to be praised

Ele fez o que eu nunca soube que teria sido possível na minha vida assim, eu tenho todos os motivos para louvá-Lo e Ele é digno de ser louvado

☐ Positivo

☐ Negativo

i hate this life i wish i die but when i remember you i say how i can leave her i can't die i love you a lot m...ay

Eu odeio essa vida eu gostaria de morrer, mas quando eu lembro de você eu digo como eu posso deixá-la Eu não posso morrer Eu te amo muito ... m ay

☐ Positivo

☐ Negativo

i laugh at people who put their life story's on here and beg for sympathy from it!

Eu rio de pessoas que colocam suas histórias de vida aqui e imploram por simpatia dela!

- ☐ Positivo
- ☐ Negativo

My brain hurts now! Man almost crossed the center median on capital Circle and if he did he would've hit me now I can't think. It hurt my brain. My life flashed before my eyes. That's why my brain hurts. My brain hasn't moved that fast in years.

Meu cérebro dói agora! Homem quase atravessou a mediana centro no Círculo de capital e se ele fez ele teria me bateu agora eu não posso pensar. Doeu meu cérebro. Minha vida passou diante dos meus olhos. É por isso que meu cérebro dói. Meu cérebro não se moveu tão rápido em anos.

- ☐ Positivo
- ☐ Negativo

What a beautiful day!!!! I love my life and kids.

Que lindo dia!! Eu amo minha vida e as crianças.

- ☐ Positivo
- ☐ Negativo

What a woman! Just reading about her life inspires me! I have read a little of her stuff and highly recommend it...I'll post the titles below, just in case one grabs you. Library probably has them all...

Que mulher! Basta ler sobre sua vida me inspira! Eu li um pouco de suas coisas e recomendo ... eu vou postar os títulos abaixo, apenas no caso de um agarra-lo. Biblioteca provavelmente tem todos eles ...

- ☐ Positivo
- ☐ Negativo

Anyone looking for a new phone I recommend the samsung galaxy note 2. just got it and love it:)

Quem estiver procurando por um novo telefone eu recomendo o Samsung Galaxy Note 2. só tenho ele e adoro :)

- ☐ Positivo
- ☐ Negativo

I am sure by now...everyone has a tablet?! Well I don't but was wondering what everyones Favorite Tablet is so far? From Ipad to Samsung to Windiws 8? I am looking to get one for our business. Just looking for suggestions!?

Estou certo de que agora ... todo mundo tem um tablet?! Bem, eu não, mas queria saber o que everyones Tablet favorito é tão longe? De Ipad para a Samsung para Windiws 8? Eu estou olhando para obter um para o nosso negócio. Basta olhar para sugestões!?

- ☐ Positivo
- ☐ Negativo

my life leads only be happy with the samsung galaxy s 3 :D b
minha vida leva apenas ser feliz com o Samsung Galaxy S3: D b

- ☐ Positivo
- ☐ Negativo

Seriously considering throwing Roxanne Lacasse's new computer out the window!!!!
RRRRRRRRRAAAAAWWWWRRRRRRRR!!!!!! Windows 8 is so bad that I'm begging for vista back! YUK!

Considerando seriamente jogando computador novo Roxanne Lacasse de fora da janela!! RRRRRRRRAAAAAWWWWRRRRRRRR!!!!!! Windows 8 é tão ruim que eu estou implorando para o Vista de volta! YUK!

- ☐ Positivo
- ☐ Negativo

I thought these android tablets were flash player compatible? How come adobe is telling me its not? Ready to go back to best buy and throw my new tablet at them! Ugh, totally frustrated

Eu pensei que estes tablets Android foram flash player compatível? Como vêm adobe está me dizendo que não é? Pronto para voltar ao melhor comprar e jogar o meu novo tablet para eles! Ugh, totalmente frustrado

- ☐ Positivo
- ☐ Negativo

I like my new galaxy tablet. Still learning how to use it.
Eu gosto do meu novo galaxy tablet. Ainda está aprendendo a usá-lo.

- ☐ Positivo

☐ Negativo

Since I uploaded my iPhone to iOS 6.1, I started to discover some strange WiFi usage has been used. That's odd, very odd. >:\

Desde que eu carreguei meu iPhone para o iOS 6.1, comecei a descobrir algum uso estranho WiFi tem sido utilizado. Isso é estranho, muito estranho. >: \

☐ Positivo

☐ Negativo

The only promise in life ... Is death ... So live ...Before you die....

A única promessa na vida ... É a morte ... Então, viver ... Antes de morrer

☐ Positivo

☐ Negativo

I AM BORED TO DEATH. CAN'T DO ANYTHING CUZ I HURT SO BAD.

Estou entediado até a morte. Não pode fazer nada CUZ estou muito machucado.

☐ Positivo

☐ Negativo

Lol He Gets Mad When Were Talking On The Phone & I Randomly Play Justin Bieber Songs cx

Lol Ele fica louco quando estávamos conversando no telefone e eu tocando aleatoriamente canções de Justin Bieber cx

☐ Positivo

☐ Negativo

i will never abandon anybody I love...we might drift apart but ill never abandon you.

eu nunca vou abandonar ninguém que eu amo ... podemos se afastar, mas nunca abandonará.

☐ Positivo

☐ Negativo

There now you went and did it puppy, you brought out the tears in my eyes! I hate people who abandon their pets!

Há agora você foi e fez isso cachorro, você tirou as lágrimas nos meus olhos! Eu odeio as pessoas que abandonam seus animais de estimação!

☐ Positivo

☐ Negativo

So I'm watching freddy vs jason and my 4 yr old loves the crap out of scary movies should I be afraid!!!!

Então, eu estou assistindo Freddy vs Jason e meus 4 anos de idade ama o cagar de filmes de terror que eu deveria ter medo!!

☐ Positivo

☐ Negativo

Damm everywhere you turn theres danger everywhere even here when i'm on Facebook someone could look me up on google. scary

Damm por todos os lados há perigo em toda parte, mesmo aqui quando estou no Facebook alguém poderia me procure no google. assustador

☐ Positivo

☐ Negativo

51% percent of America voted for Obama the 49% percent of people who are not stupid voted against him... -Herman Cain on Hannity tonight

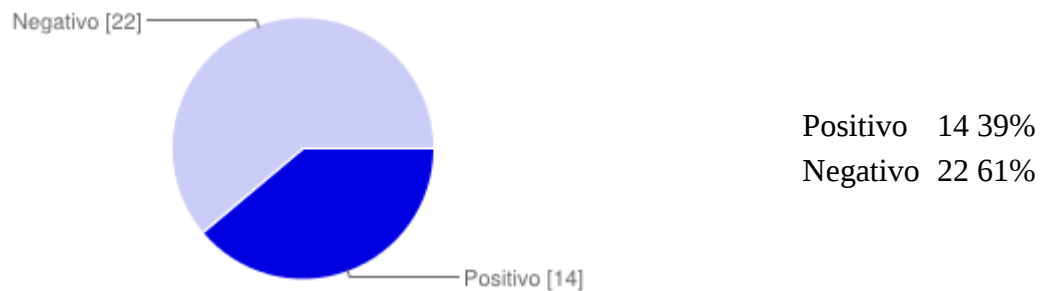
51% por cento da América votaram em Obama e 49% por cento de pessoas que não são estúpidos votou contra ele ...-Herman Cain esta noite Hannity

☐ Positivo

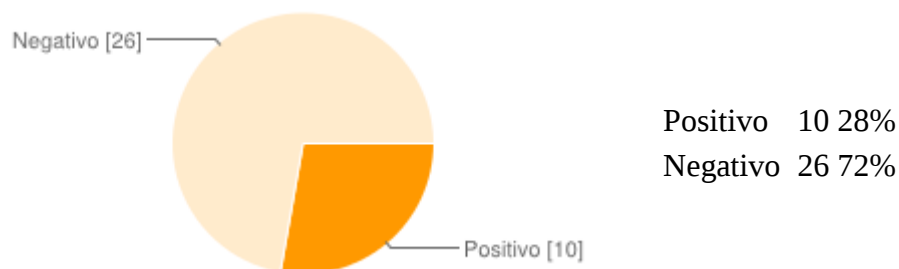
☐ Negativo

Apêndice C – Resultado da avaliação aplicada com as pessoas.

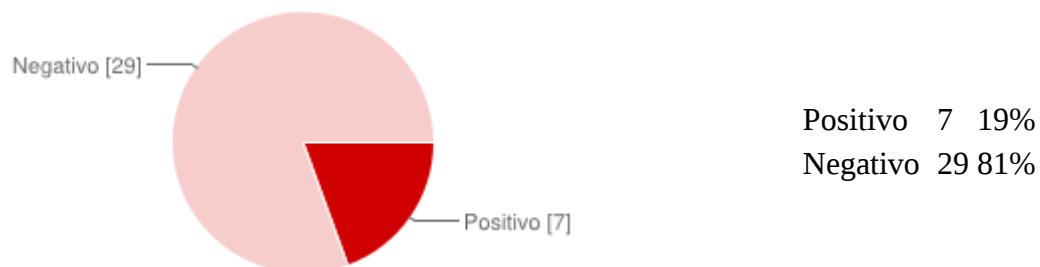
meteor about to hit earth. say you've got 24 hours to think about it. what would be your last song request?



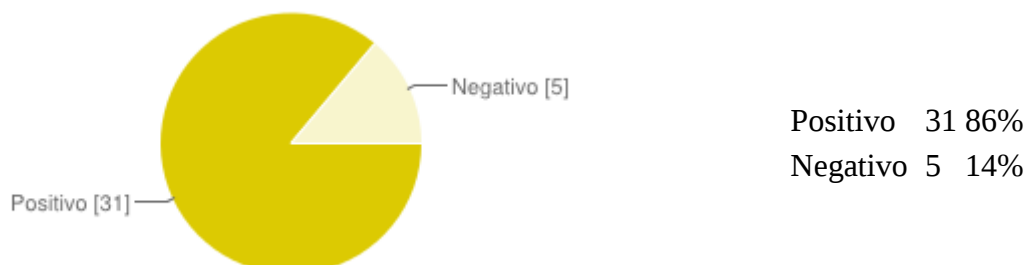
did anybody else hear about the meteor that hit russia yesterday?



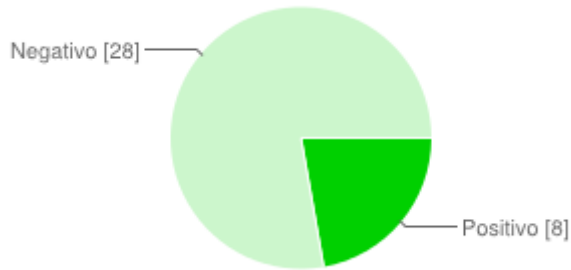
Pretty pathetic when they can't even understand the difference between an asteroid, a meteor and a meteorite.



Pass this on. Put pressure on your Senators to remove all petroleum products from the stock market. This is how control is done. There is no other reason for these products on the "speculation" list. Nothing but greed. PASS THIS ON, WE CAN MAKE A CHANGE FOR OUR COUNTRY AND PEOPLE.

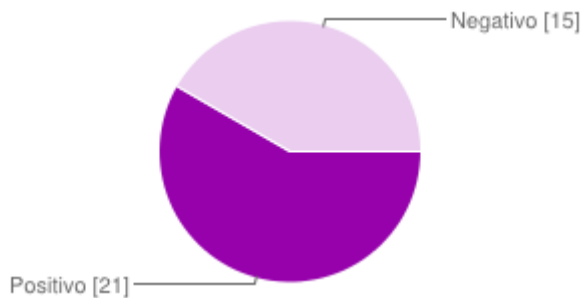


The Federal Gov't shut down for President's Day, and the oil companies jacked the price of gas up 20 cents today, price went from \$3.79.9 to \$3.99.9 in the last couple of hours, WOW!



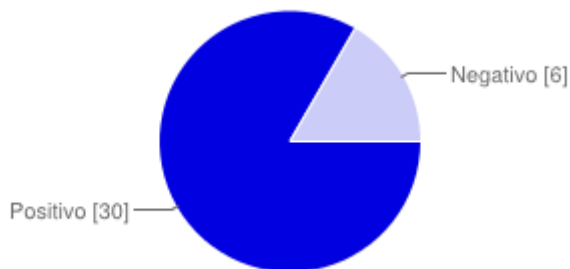
Positivo 8 22%
Negativo 28 78%

In America we should always look to promote opportunity for one to move up. Obama is that kind of guy. He became president from a apartment in Chicago and from a college loan. He gets it. He's a product of it.



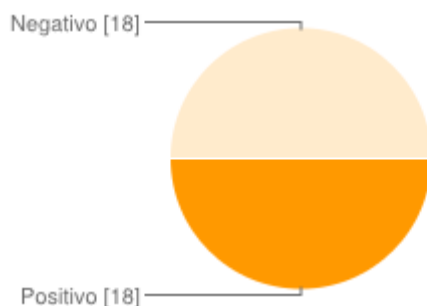
Positivo 21 58%
Negativo 15 42%

He did what i never knew would have been possible in my life so, i have every reason to praise Him and He's worthy to be praised



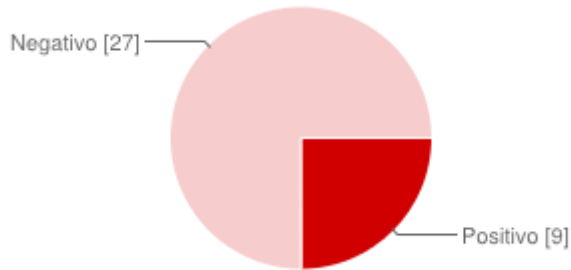
Positivo 30 83%
Negativo 6 17%

i hate this life i wish i die but when i remember you i say how i can leave her i can't die i love you a lot m...ay



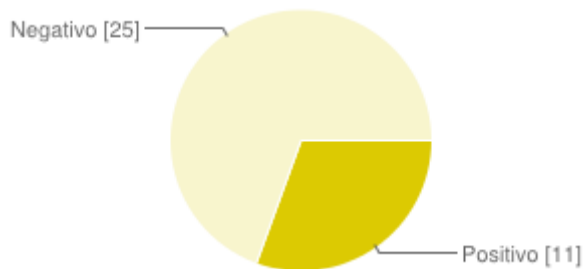
Positivo 18 50%
Negativo 18 50%

i laugh at people who put their life story's on here and beg for sympathy from it!



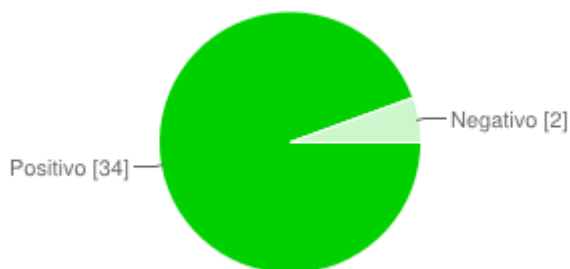
Positivo 9 25%
Negativo 27 75%

My brain hurts now! Man almost crossed the center median on capital Circle and if he did he would've hit me now I can't think. It hurt my brain. My life flashed before my eyes. That's why my brain hurts. My brain hasn't moved that fast in years.



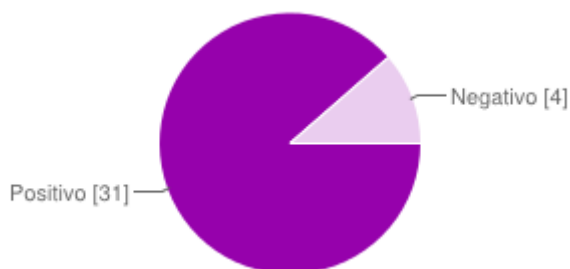
Positivo 11 31%
Negativo 25 69%

What a beautiful day!!!! I love my life and kids.



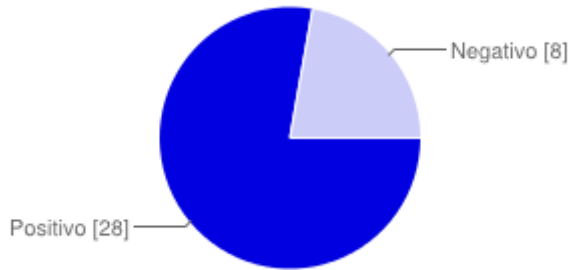
Positivo 34 94%
Negativo 2 6%

What a woman! Just reading about her life inspires me! I have read a little of her stuff and highly recommend it...I'll post the titles below, just in case one grabs you. Library probably has them all...



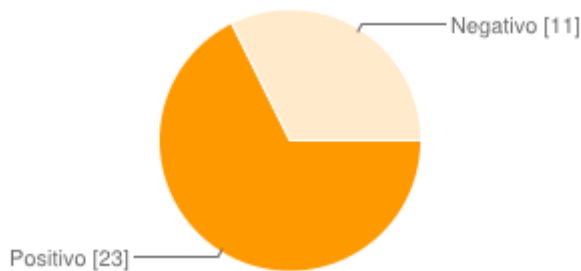
Positivo 31 89%
Negativo 4 11%

Anyone looking for a new phone I recommend the samsung galaxy note 2. just got it and love it:)



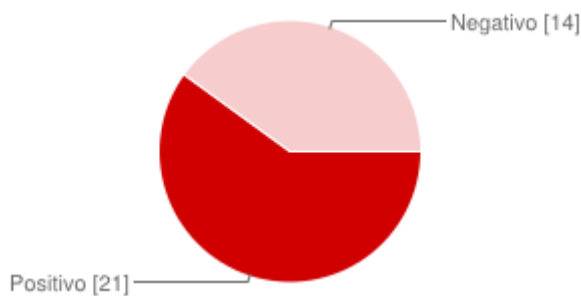
Positivo 28 78%
Negativo 8 22%

I am sure by now...everyone has a tablet?! Well I don't but was wondering what everyone's Favorite Tablet is so far? From Ipad to Samsung to Windiws 8? I am looking to get one for our business. Just looking for suggestions!?



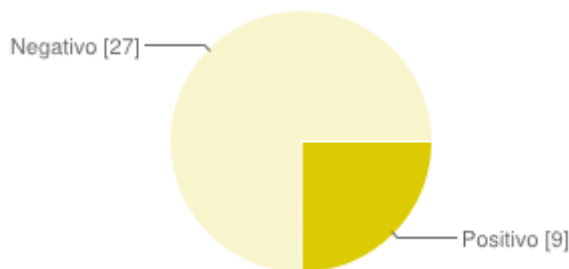
Positivo 23 68%
Negativo 11 32%

my life leads only be happy with the samsung galaxy s 3 :D b



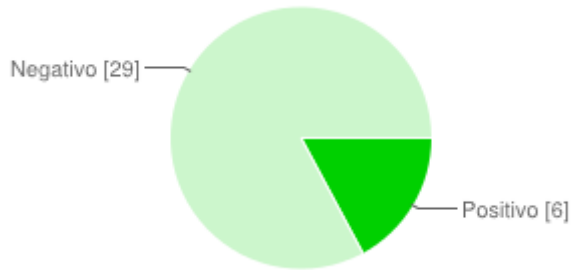
Positivo 21 60%
Negativo 14 40%

Seriously considering throwing Roxanne Lacasse's new computer out the window!!!!
RRRRRRRRRAAAAAWWWWRRRRRRRR!!!!!! Windows 8 is so bad that I'm
begging for vista back! YUK!



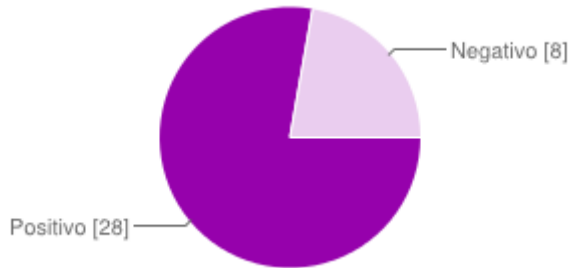
Positivo 9 25%
Negativo 27 75%

I thought these android tablets were flash player compatible? How come adobe is telling me its not? Ready to go back to best buy and throw my new tablet at them! Ugh, totally frustrated



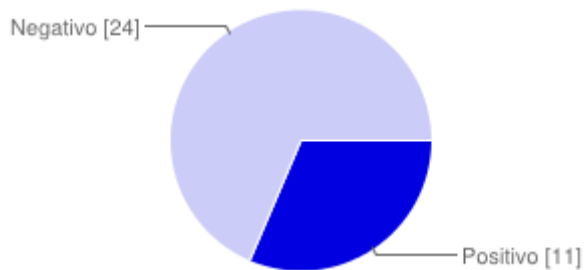
Positivo 6 17%
Negativo 29 83%

I like my new galaxy tablet. Still learning how to use it.



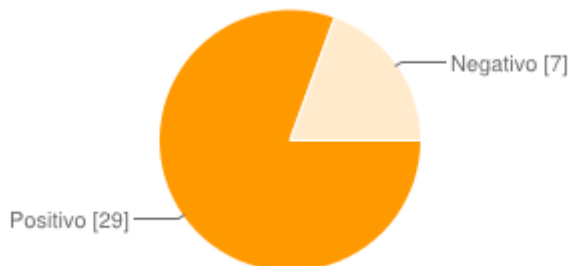
Positivo 28 78%
Negativo 8 22%

Since I uploaded my iPhone to iOS 6.1, I started to discover some strange WiFi usage has been used. That's odd, very odd. >:\



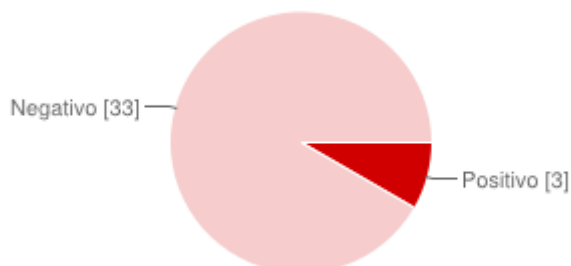
Positivo 11 31%
Negativo 24 69%

The only promise in life ... Is death ... So live ...Before you die....



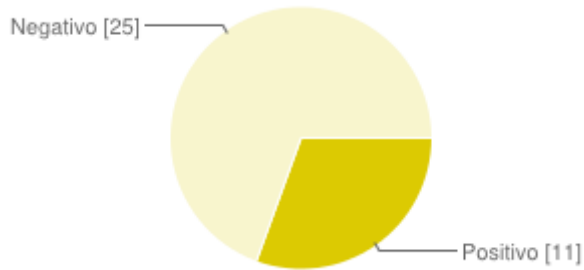
Positivo 29 81%
Negativo 7 19%

I AM BORED TO DEATH. CAN'T DO ANYTHING CUZ I HURT SO BAD.



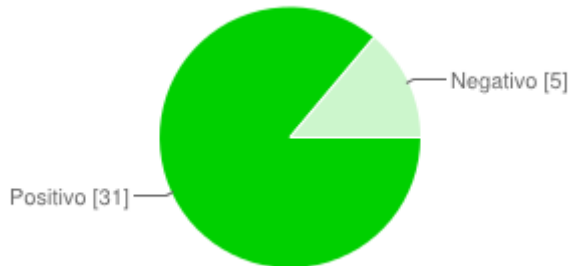
Positivo 3 8%
Negativo 33 92%

Lol He Gets Mad When Were Talking On The Phone & I Randomly Play Justin Bieber Songs cx



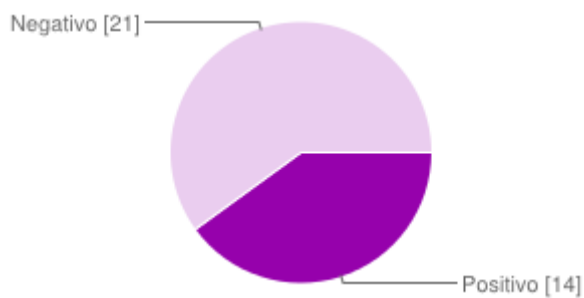
Positivo 11 31%
Negativo 25 69%

i will never abandon anybody I love...we might drift apart but ill never abandon you.



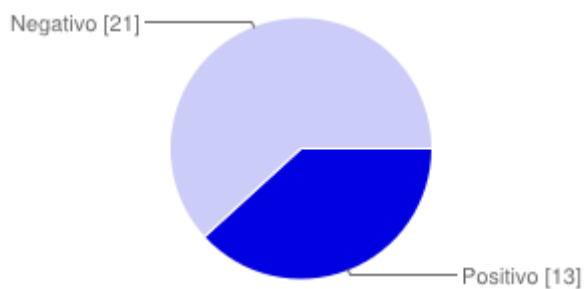
Positivo 31 86%
Negativo 5 14%

There now you went and did it puppy, you brought out the tears in my eyes! I hate people who abandon their pets!



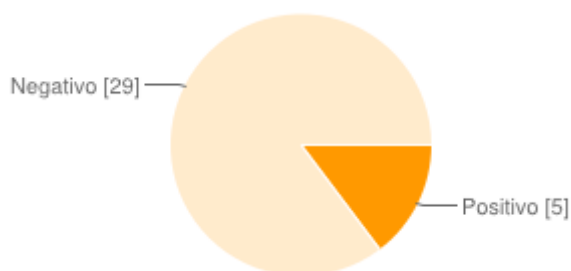
Positivo 14 40%
Negativo 21 60%

So I'm watching freddy vs jason and my 4 yr old loves the crap out of scary movies should I be afraid!!!!



Positivo 13 38%
Negativo 21 62%

Damm everywhere you turn theres danger everywhere even here when i'm on Facebook someone could look me up on google. scary



Positivo 5 15%
Negativo 29 85%

51% percent of America voted for Obama the 49% percent of people who are not stupid voted against him... -Herman Cain on Hannity tonight

