

EXTRAÇÃO DE SENTIMENTOS EM LÍNGUA PORTUGUESA COM SENTIWORDNET

Matheus Henrique Silva Salmi, Márcio Giovane Cunha Fernandes

mathsalmi@gmail.com, marcio.giovane@ueg.br

Universidade Estadual de Goiás – CCET – Sistemas de Informação
Anápolis – GO

RESUMO – Este trabalho visa identificar meios de utilizar o método SentiWordNet na extração de sentimentos expressos em trechos textuais escritos em língua portuguesa. Para tal, há uma breve revisão sobre o WordNet e como ele difere do SentiWordNet, explica-se problemas enfrentados para realizar a extração e, por fim, há a apresentação do Sentiment Analyzer, que é a solução posposta a fim de possibilitar análise de sentimentos em português.

Palavras-Chave – análise de sentimento, mineração de opinião, SentiWordNet.

SENTIMENT EXTRACTION IN PORTUGUESE WITH SENTIWORDNET

ABSTRACT – This paper aims to identify means to use SentiWordNet to extract sentiments expressed in texts written in Portuguese. To do so, there is a brief review about WordNet and how it differs from SentiWordNet, it explains problems faced during the sentiment extraction and finally there are considerations about Sentiment Analyzer, the proposed solution in order to enable the sentiment analysis in Portuguese.

KEYWORDS – sentiment analysis, opinion mining, SentiWordNet.

I. INTRODUÇÃO

A grande disponibilidade de opiniões, virtualmente, acerca de qualquer assunto vem despertando interesses e demonstrando a necessidade de se estudar mecanismos para automação de análises textuais (SALMI, 2015).

Existem grandes estudos sobre processamento de linguagens naturais (NLP), todavia, estes negligenciam os sentimentos expressos nas opiniões publicadas em meios como o microblogue Twitter (SALMI, 2015).

Áreas como ciências políticas e sociais, marketing, entre outras, são impactadas diretamente

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

por opiniões, sendo assim, o estudo e, eventual, desenvolvimento de técnicas para extração de sentimento se mostra muito relevante e impactante (SALMI, 2015).

Nesse âmbito, este trabalho visa demonstrar a aplicação do método SentiWordNet em língua portuguesa. Para isso, há uma breve revisão sobre sua concepção e sua relação com o WordNet.

Há, também, a evidenciação de problemas e desafios enfrentados a fim de colocar o método em prática e é desenvolvida e explorada uma solução para análises de textos em português: o Sentiment Analyzer.

Por fim, há a demonstração do comparativo dos resultados obtidos nas análises manual, portuguesa (com a solução proposta) e inglesa (método SentiWordNet), em que se é feita algumas reflexões sobre os métodos de construção do algoritmo utilizado e são feitas algumas sugestões para pesquisas futuras.

II. ANÁLISE DE SENTIMENTO

A exorbitante quantidade de dados disponível na Web a torna um lugar ideal para aprendizado. Hoje, 25 anos após sua criação, é possível perceber que o problema de compartilhamento de informação foi resolvido por meio de sua arquitetura aberta, agnóstica à tecnologia e cooperativa (BERNERS-LEE, [200-]).

Todavia, a existência de conteúdo por si só não garante a distribuição de informação, uma vez que a análise manual dos dados se torna impossível devido ao alto número de publicações diárias, assim, faz-se necessário encontrar mecanismos computadorizados capazes de recuperar tais informações de modo que destaquem as mais relevantes e em tempo útil para aplicação (KING et al., 2009).

Apesar dos esforços a fim de resolver o problema de busca de informações realizados nos últimos anos, os modelos que vêm sendo aplicados não são capazes de extrair todas as informações contidas na Web; um exemplo é a presença de opiniões, acerca de diversos assuntos, negligenciadas pelos algoritmos atuais (SALMI, 2015).

A aplicação de opiniões é importante em diversas áreas, como no marketing, em que se poderia elaborar ou reajustar uma campanha baseado em comentários extraídos de redes sociais em tempo real (BENEVENUTO, 2015).

Há, também, a possibilidade de aplicações bastante inusitadas e interessantes, como a utilização de técnicas de análise de sentimento para verificar os ânimos da população durante as enchentes de Gênova, na Itália, em 2014 (BUSCALDI et al., 2015).

Portanto, segundo Salmi (2015), a análise de sentimento tem impacto positivo no público-geral, pois busca identificar algoritmos capazes de realizar a extração computadorizada de sentimentos existentes em publicações online, como as presentes em mídias sociais.

III. SENTIWORDNET

Para Salmi (2015), análise de sentimento é um campo bastante abrangente, capaz de envolver áreas como processamento de linguagens naturais (NLP), economia, ciências políticas e, em tese, quaisquer outras áreas de estudo afetadas por opiniões.

Há uma gama de métodos existentes que podem ser utilizados a fim de inferir sentimentos nos textos, tais como: emoticons, LIWC, SentiStrength, entre outros (BENEVENUTO et al., [2015?]), porém, aqui, foca-se somente no estudo e aplicação do SentiWordNet.

O SentiWordNet é baseado no estudo de 2010, realizado pela Universidade de Princeton, nos Estados Unidos: o WordNet. Nele, foi realizada a reunião das palavras de língua inglesa (verbos, adjetivos, substantivos e advérbios), agrupando-as em conjuntos conceitualmente relacionados, denominados *synsets* (PRINCETON, 2010; SALMI, 2015).

É importante perceber que embora o *synset* seja único, ele pode ser formado por palavras já utilizadas em outros *synsets*, portanto, é comum perceber que as palavras agrupadas são repetidas, mas o conceito expressado é único (SALMI, 2015).

A estrutura do WordNet é, relativamente, simples, composta de vários *synsets*, dispostos, fisicamente, um por linha, que, por sua vez, são formados por um identificador único, o agrupamento de palavras que o formam e um *gloss*, que é a seção responsável por exemplificar o conceito ali formado e garantir a unicidade conceitual proposta pelo projeto (SALMI, 2015).

Para exemplificar, pode-se utilizar a palavra “clown” (palhaço, em português); o sistema online de Princeton (2010), que implementa o WordNet, identifica dois *synsets* relacionados à palavra, sendo eles:

1. {09950623} clown#1 buffoon#1 (a rude or vulgar fool); e
2. {09950334} clown#2 #buffoon#2 #goof#2 goofball#2 (a person who amuses others by ridiculous behavior).

A grande diferença entre o WordNet e o SentiWordNet é que o estudo de Princeton não foi pensado para análise de sentimento, assim, não possui informações suficientes para que, *in natura*, seja utilizado em algoritmos de extração de sentimento (SALMI, 2015).

O SentiWordNet, por sua vez, criado a partir do WordNet, possui o objetivo de reaproveitar todo o conhecimento já solidificado durante a construção de seu projeto-pai, aplicando-o ao contexto de computação social, mais especificamente, no de análise de opiniões (SALMI, 2015).

Para tal tarefa, Esuli e Sebastiani (2006) desenvolveram e aplicaram uma metodologia a fim de atribuir as características necessárias ausentes no WordNet, resultando, então, no processo de inclusão de um terceto de valores, cada um para identificar positividade, negatividade e objetividade (neutro) de cada *synset*; processo chamado de polarização PN (positivo-negativa).

No objetivo de tornar o método mais hábil a perceber nuances nos trechos textuais, Esuli e Sebastiani (2006) optaram por desenvolver um classificador não binário, de modo que todos os *synsets* possuam o terceto acima mencionado e que seus valores sejam complementares.

Caso a implementação fosse diferente, provavelmente, o classificador funcionaria apenas para os casos em que um dos valores fosse forte, identificando, incorretamente, como neutros os casos em que os três valores são aproximados, ou seja, não possuiria a capacidade de perceber diferentes graus de intensidade (SALMI, 2015).

É importante destacar que, ao final do processo de polarização PN, Esuli e Sebastiani (2006) realizaram um passo adicional de normalização dos valores obtidos a fim de garantir que o terceto estivesse no intervalo inclusivo de 0 a 1.

IV. SENTIMENT ANALYZER

Conforme explicado, o SentiWordNet, por meio da herança de conhecimentos do WordNet e da polarização dos *synsets*, é capaz de analisar trechos textuais e inferir os sentimentos ali expressos, classificando-os, então, como positivos, negativos ou neutros.

Todavia, sua base de conhecimento está restrita ao idioma inglês, assim, requer adaptações e modificações para que possa analisar textos em outros idiomas, como o português (SALMI, 2015).

Dessa forma, esse trabalho proporciona uma maneira de construir a base de dados necessária para funcionamento do método em língua portuguesa. Para tal, procurou-se seguir os passos explorados por Esuli e Sebastiani (2006), adaptando-os à realidade do idioma e tentou-se criar um ambiente dinâmico de treinamento por meio da utilização de uma plataforma virtual, escrita para tecnologia Web, chamada Sentiment Analyzer (SALMI, 2015).

Nessa plataforma, Salmi (2015) visou resolver os dois principais problemas para o projeto: criar a base léxica em português e atribuir a polarização para cada *synset* cadastrado nessa base. Além disso, garantiu o controle das palavras cadastradas no dicionário léxico por meio da existência de um ambiente administrativo e disponibilizou gratuitamente o uso da ferramenta, a fim de fomentar a pesquisa de análise de sentimento em língua portuguesa.

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

As soluções propostas por Salmi (2015) foram:

1. possibilidade de cadastro, edição e exclusão de *synsets* por administradores;
2. ferramenta pública para inferir sentimentos em textos em português;
3. feedback do usuário por meio de um sistema de votação de sentimentos expressos em *synsets*.

Salmi (2015) acreditou que com a criação da ferramenta gratuita de análise de sentimentos, fosse possível despertar interesse por parte do público-geral em estudar e testar o método no idioma português e aperfeiçoar a ferramenta por meio das votações, proposta no item três.

Ademais, é importante salientar que o terceiro item é bem importante, já que ele proporcionará a polarização dos *synsets*. Essa funcionalidade apesar de poderosa e essencial, é simples, consiste apenas em expor, de maneira aleatória, uma palavra de um *synset* previamente cadastrado (por administradores) e exibir um dispositivo para que o usuário selecione o percentual de positividade, negatividade e objetividade expressos em seu *gloss*.

No método de Esuli e Sebastiani (2006), o classificador foi o item mais importante, pois, foi a partir dele que a base léxica do WordNet passou a ter as características necessárias que possibilitaram sua aplicação no contexto de extração de sentimento; aqui não é diferente.

O classificador o Sentiment Analyzer também segue as concepções elaboradas por Esuli e Sebastiani (2006), assim, utiliza uma abordagem complementar em detrimento à binária, justamente para também ser capaz de analisar nuances textuais e garantir visibilidade de intensidade nas análises (SALMI, 2015).

A polarização PN é baseada na mesma escala de Esuli e Sebastiani (2006), sendo assim, todos os valores pertencem ao conjunto dos números reais e estão no intervalo inclusivo de 0 a 1.

Conforme anteriormente explicado, um *synset* é composto de várias palavras, *gloss* e o identificador único. Por meio do sistema de votação, o usuário dará *feedback* acerca de uma palavra aplicada no contexto de um *synset*, assim, no final, obter-se-á vários votos para uma palavra, pertencente a um *synset*.

Vale salientar, novamente, que o voto está diretamente vinculado à palavra e ao *synset*, ao mesmo tempo, portanto, mesmo que a palavra esteja vinculada a outro *synset*, os votos de uma palavra não impactarão os outros *synsets*.

Desse modo, é possível relatar que o classificador identifica a pontuação de valores positivos (1) realizando uma média aritmética, em que se soma todos os valores positivos pertencente àquela palavra e àquele *synset*, e divide-os pelo número total de votos do *synset*

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

(SALMI, 2015).

Similarmente, o cálculo para a negatividade do *synset* (2) se dá aplicando média aritmética, somando todos os valores negativos da palavra e do *synset* e divido-os pelo total de todos do *synset* (SALMI, 2015).

Conforme dito, os valores do terceto (positividade, negatividade e neutro) são complementares, assim, para se descobrir o total neutro (3) para uma palavra, pode-se reduzir do valor máximo permitido (um) a somatória do total positivo e negativo.

As fórmulas dos três cálculos expressos por Salmi (2015) são:

$$vp_s = \frac{\sum_{i=1}^n vp_i}{n} \quad (1)$$

$$vn_s = \frac{\sum_{i=1}^n vn_i}{n} \quad (2)$$

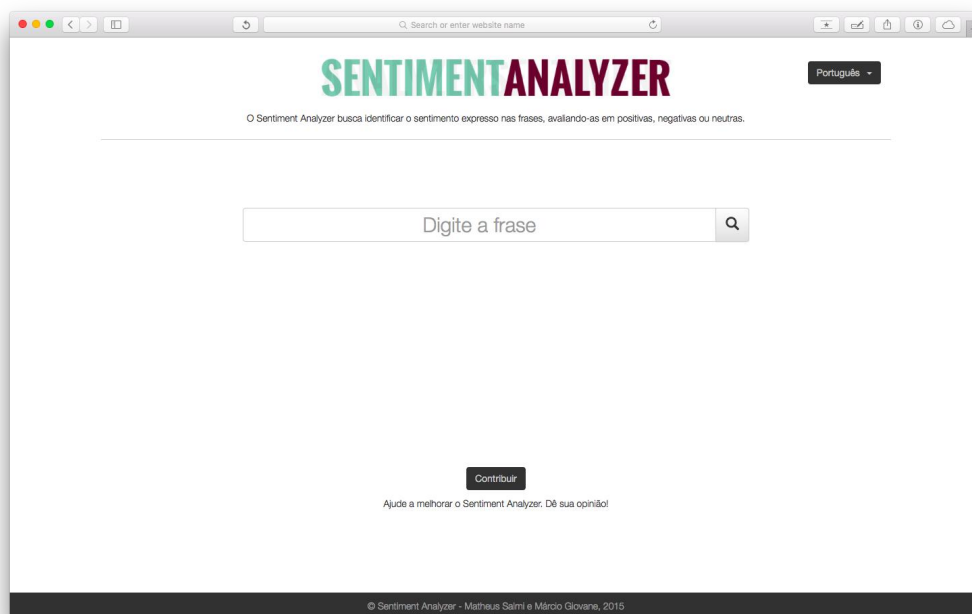
$$vo_s = 1 - (vp_s + vn_s) \quad (3)$$

Onde:

- vp_s - Valor positivo do *synset*.
- vp_i - Valor positivo de cada voto.
- vn_s - Valor negativo do *synset*.
- vn_i - Valor negativo de cada voto.
- vo_s - Valor neutro do *synset*.
- n - Total de votos.

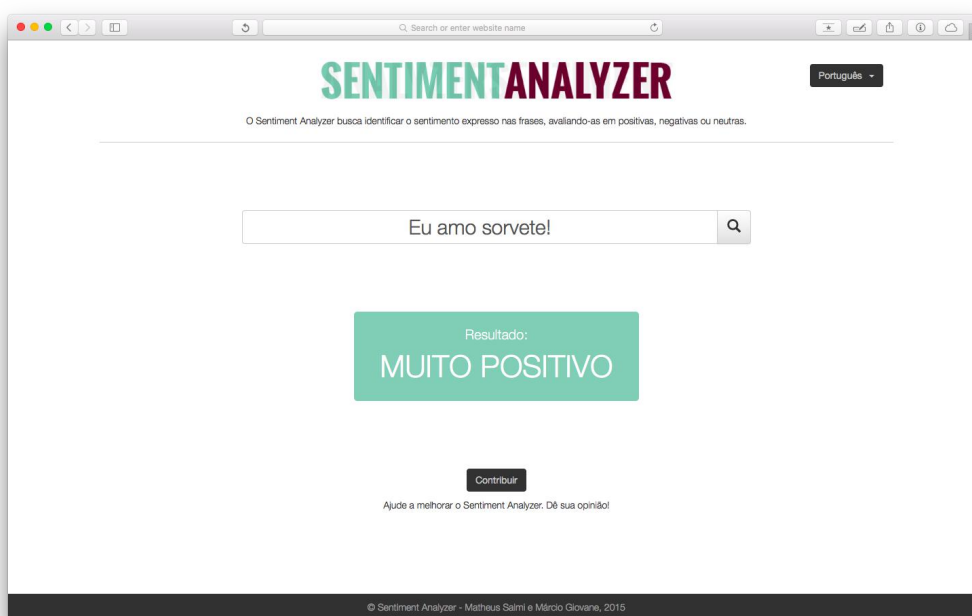
Explicado o funcionamento do principal mecanismo do sistema (o classificador) e antes de testá-lo e verificar seus resultados, é interessante ilustrar partes do *software* produzido, de modo a auxiliar na visualização da materialização da solução proposta.

Fig. 1: Tela inicial do sistema.



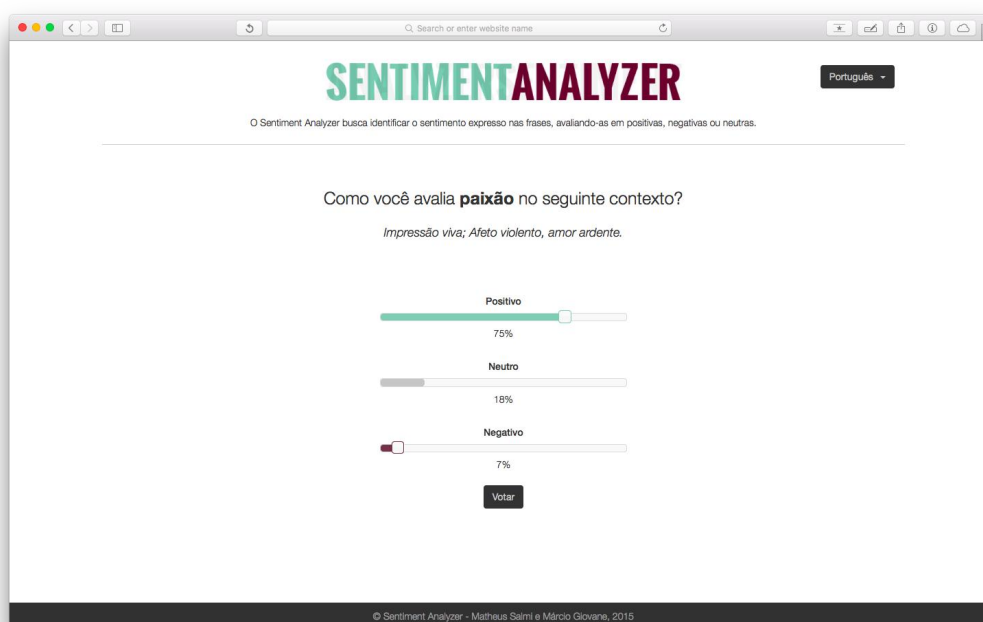
Fonte: Salmi (2015).

Fig. 2: Tela de análise de sentimento de frases.



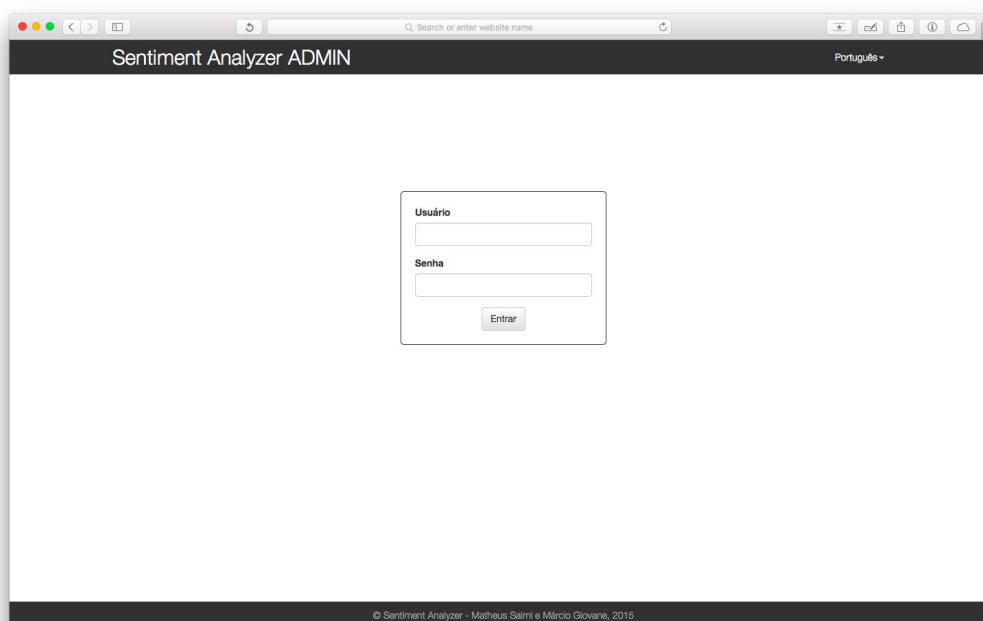
Fonte: Salmi (2015).

Fig. 3: Tela de votação do *synset*.



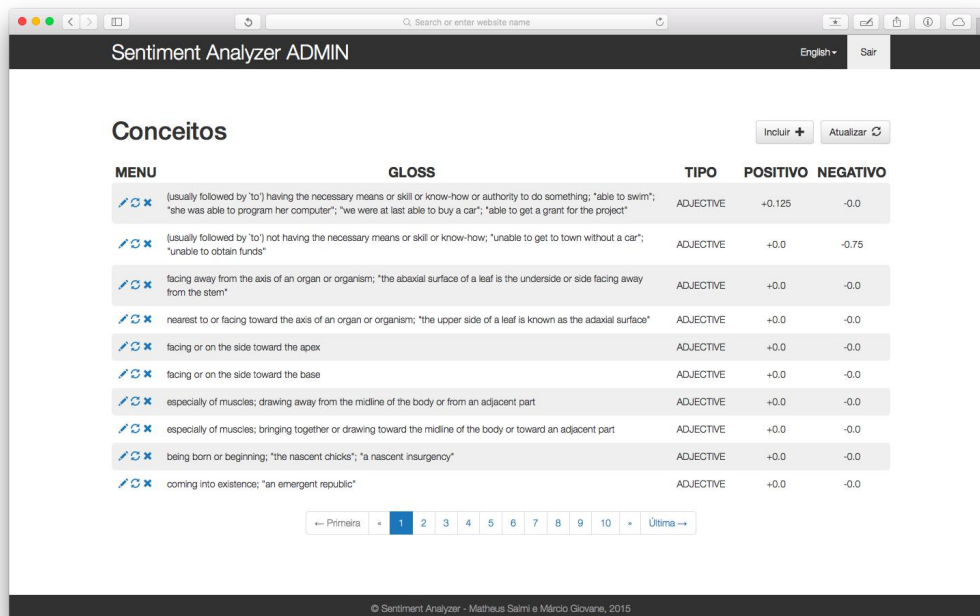
Fonte: Salmi (2015).

Fig. 4: Tela de *login* da área restrita (administrativo).



Fonte: Salmi (2015).

Fig. 5: Tela de listagem de *synsets* cadastrados (administrativo).

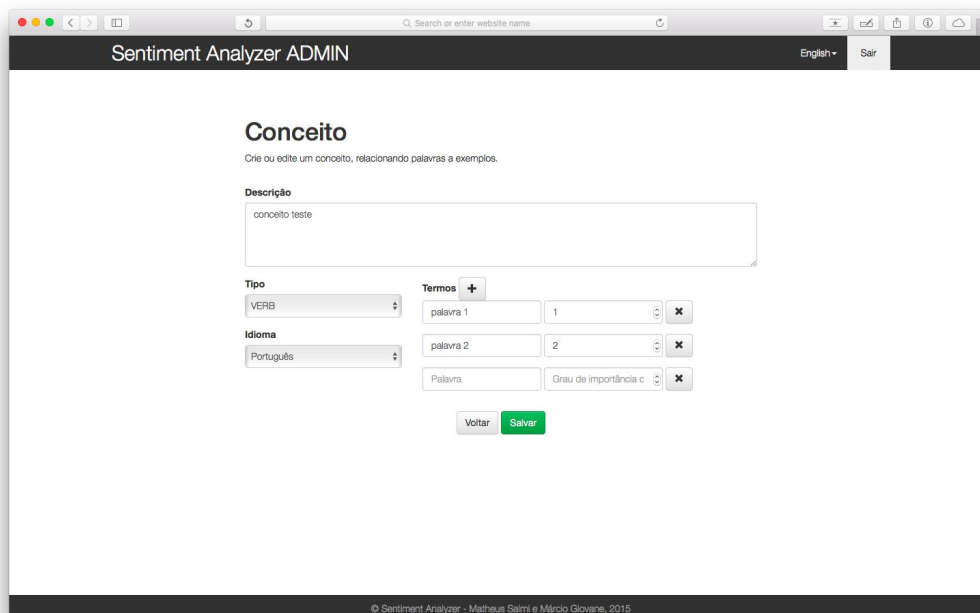


The screenshot shows the 'Conceitos' (Concepts) page in the Sentiment Analyzer ADMIN. It features a table with columns: MENU, GLOSS, TIPO, POSITIVO, and NEGATIVO. The table lists 10 synsets, all of type 'ADJECTIVE'. Each row has a 'MENU' icon (a blue square with a white 'x') and a 'Gloss' description. The 'POSITIVO' and 'NEGATIVO' columns show numerical values. At the bottom, there is a pagination bar with links for 'Primeira', '1', '2', '3', '4', '5', '6', '7', '8', '9', '10', and 'Última'.

MENU	GLOSS	TIPO	POSITIVO	NEGATIVO
	(usually followed by 'to') having the necessary means or skill or know-how or authority to do something; "able to swim"; "she was able to program her computer"; "we were at last able to buy a car"; "able to get a grant for the project"	ADJECTIVE	+0.125	-0.0
	(usually followed by 'to') not having the necessary means or skill or know-how; "unable to get to town without a car"; "unable to obtain funds"	ADJECTIVE	+0.0	-0.75
	facing away from the axis of an organ or organism; "the abaxial surface of a leaf is the underside or side facing away from the stem"	ADJECTIVE	+0.0	-0.0
	nearest to or facing toward the axis of an organ or organism; "the upper side of a leaf is known as the adaxial surface"	ADJECTIVE	+0.0	-0.0
	facing or on the side toward the apex	ADJECTIVE	+0.0	-0.0
	facing or on the side toward the base	ADJECTIVE	+0.0	-0.0
	especially of muscles; drawing away from the midline of the body or from an adjacent part	ADJECTIVE	+0.0	-0.0
	especially of muscles; bringing together or drawing toward the midline of the body or toward an adjacent part	ADJECTIVE	+0.0	-0.0
	being born or beginning; "the nascent chicks"; "a nascent insurgency"	ADJECTIVE	+0.0	-0.0
	coming into existence; "an emergent republic"	ADJECTIVE	+0.0	-0.0

Fonte: Salmi (2015).

Fig. 6: Tela de cadastro de *synset* (administrativo).



The screenshot shows the 'Conceito' (Concept) form in the Sentiment Analyzer ADMIN. It includes a 'Descrição' (Description) text area with the placeholder 'conceito teste'. Below this, there are dropdown menus for 'Tipo' (set to 'VERB') and 'Idioma' (set to 'Português'). To the right, there is a 'Termos' section with a '+' button and three rows of input fields for 'palavra 1', 'palavra 2', and 'Palavra', each with a corresponding 'Grau de importância' (Importance degree) field and a delete button 'x'. At the bottom, there are 'Voltar' (Back) and 'Salvar' (Save) buttons.

Fonte: Salmi (2015).

Para aplicação do sistema, Salmi (2015) selecionou alguns tuítes extraídos no período da eleição presidencial do Brasil em 2014. A fim de testar, foi necessário, primeiramente, realizar um cadastro inicial de palavras de modo a formar a base léxica em português.

Após, houve a categorização dos *synsets* por meio do sistema de votação; essa parte foi difícil, uma vez que todo o funcionamento do sistema depende dela e foi necessário despertar interesse em usuários para que eles contribuíssem dando seu *feedback* em forma de votos (SALMI, 2015).

Então, houve a tradução manual dos tuítes selecionados para inglês, assim, como métrica, se utilizaria a versão original do SentiWordNet, possibilitando, então, visualizar como o algoritmo se saiu ao analisar textos em português e confrontar com os resultados obtidos com traduções, considerando que há perda semântica em virtualmente quase todas as traduções.

Tabela 1 – tuítes selecionados para análise.

Nº	Original	Tradução
1	A Dilma prometeu 800 aeroportos e não fez nenhum o Aécio construiu pelo menos um como governador.	Dilma promised 800 airports and build none. Aécio built at least one as governor
2	O caos está instalado. Dilma vem aí!!! http://t.co/gpTb0tpqaO	Chaos is here. Dilma is coming!!! http://t.co/gpTb0tpqaO
3	Eu aprendi uma nova conta: 13 - 4 = 7. A Dilma é uma comédia sem fim. Pobre, Brasil.	I learned a new operation: 13 - 4 = 7. Dilma is an endless comedy. Poor Brazil.
4	Dilma tá fazendo aliança até com o PRÓPRIO DEMÔNIO nessas eleições	Dilma is establishing alliance with the devil himself in this elections
5	A @dilmabr é uma covarde sem caráter. Te desafio a vir para o Maranhão sacudir a mão de @lobaofilho15 sua aproveitadora sem escrúpulos!	@dilmabr is coward. I challenge you to come to Maranhão hand shake @lobaofilho15 you unscrupulous exploiter!
6	O mentiroso e desocupado João Pedro Stidele ameaça Marina Silva http://t.co/IWkEgXIBC6	the liar and tramp João Pedro Stidele threatens Marina Silva http://t.co/IWkEgXIBC6
7	felicidade se resumindo a ver Dilma no topo do censo	happiness on seeing Dilma at the top of the census
8	Professor Tarcísio, Eduardo Jorge, Marina Silva, melhores pessoas.	Professor Tarcísio, Eduardo Jorge, and Marina Silva. The best people.
9	É, já vi que Dilma tá sem moral mesmo.	Yeah... I see nobody respects Dilma
10	Esses eventos pedindo o "impeachment" da Dilma beiram o ridículo...	These events asking for Dilma's impeachment are ridiculous...

Fonte: Salmi (2015)

Salmi (2015) ressaltou que durante a criação da base léxica, houve foco nas palavras extraídas dos tuítes, isso porque devido ao tempo que dispunha, não seria possível treinar o sistema o suficiente para que fosse capaz de realizar uma análise útil; caso uma abordagem aleatória de cadastro de palavras fosse adotada, o sistema provavelmente classificaria, erroneamente, todos os tuítes como neutros.

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

Tabela 2 – comparação entre avaliações manual, português e inglês.

Nº	Manual	Português	Inglês
1	Negativo	Neutro	Negativo
2	Negativo	Negativo	Negativo
3	Negativo	Neutro	Negativo
4	Negativo	Negativo	Negativo
5	Negativo	Negativo	Neutro
6	Negativo	Neutro	Negativo
7	Positivo	Positivo	Positivo
8	Positivo	Neutro	Positivo
9	Negativo	Neutro	Negativo
10	Negativo	Neutro	Neutro

Fonte: Salmi (2015)

Ao considerar que a coluna “manual” possua o valor mais correto, em relação ao sentimento expressos pelos autores, pode-se perceber que o sistema inglês, utilizando o SentiWordNet original, se saiu melhor, acertando mais casos.

Salmi (2015) explicita que esse comportamento ocorreu devido à pequena base léxica e de votos utilizada para os testes, com o aumento dos votos e de palavras registradas, o método português provavelmente se sairia melhor, produzindo resultados capazes de refletir melhor a realidade.

V. CONCLUSÃO

Análise de sentimento é um campo de estudo um tanto novo, porém muito interessante e que, certamente, impacta as demais áreas que contam com opiniões.

Apesar de ter ganhado foco nas últimas décadas, ainda há muito o que se avançar, isso porque os problemas enfrentados são imensos. Como explicado no decorrer do trabalho, a computação social, bem como a análise de sentimento, influencia uma porção de campos de estudos, tais como o processamento de linguagens naturais (NLP), ciências políticas e econômicas etc.

Os métodos aqui expostos visam proporcionar maneiras de extrair sentimentos em textos; houve revisão acerca dos métodos WordNet e SentiWordNet, que, originalmente, focam no idioma inglês, mesmo que o objetivo do trabalho seja a análise de trechos textuais em português.

Tais revisões foram necessárias porque a solução proposta (Sentiment Analyzer) é um estudo

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

oriundo dos esforços acerca do SentiWordNet, com o objetivo de aplicar todo o conhecimento obtido durante sua concepção para elaborar um método funcional para a língua portuguesa.

Os resultados obtidos na execução da solução, o Sentiment Analyzer, demonstram que produzir um sistema funcional, capaz de demonstrar resultados mais similares à realidade de sentimentos expressos é bem complexo.

Parte da complexidade se dá porque com o método-base escolhido, o SentiWordNet, é necessário construir uma base léxica, o que dispense de oneroso trabalho humano. Além dessa construção, se faz necessário polarizar cada um dos itens (*synsets*) cadastrados.

No caso do Sentiment Analyzer, a polarização se deu por meio do sistema de votação, no qual os usuários caracterizam cada uma palavra no contexto de um *synset* em positivo, negativo e neutro (objetivo), o que, também, requer um grande trabalho manual, pois sem ele o sistema não é capaz de analisar corretamente as frases; tal comportamento pode ser visto nos resultados, onde se compara as análises manual, em português e do projeto original SentiWordNet, em inglês.

Como recomendação, Salmi (2015) sugeriu que o método de construção da base léxica sofresse leves alterações, exemplificando que a metodologia seguida na língua inglesa não é eficaz para o português porque as palavras do idioma lusitano se alteram conforme gênero e número, portanto, diferentemente do inglês, que é de gênero neutro, seria necessário cadastrar palavras no masculino, feminino, no singular e no plural.

Esse cadastro quaduplicado causaria problemas no sistema de votação, já que o mesmo conceito teria de ser votado quatro vezes para que fosse corretamente pontuado como positivo, negativo ou neutro. Assim, Salmi (2015) sugeriu a consideração do uso da raiz da palavra ao invés de utilizá-la por completo nos algoritmos de análise e de votação.

Além dessas considerações, é possível que haja outras, aqui desapercibidas, a serem realizadas em futuros estudos e espera-se que a atual pesquisa tenha contribuído positivamente na demonstração do fascinante estudo acerca da extração de sentimentos em português.

REFERÊNCIAS

BENEVENUTO, Fabrício. **Técnicas de Análise de Sentimento**. [2015/1]. Disponível em: <<http://homepages.dcc.ufmg.br/~fabricio/20151-sentiment-analysis-course.htm>>. Acesso em: 7 de setembro de 2016.

BENEVENUTO, Fabrício; ARAÚJO, Matheus; GONÇALVES, Pollyanna. **Métodos para Análise de Sentimentos no Twitter**. [2015?]. Disponível em:

SIUNI-UEG - Anápolis – Goiás – Brasil

07 a 09 de outubro de 2016

<homepages.dcc.ufmg.br/~fabricio/download/webmedia13.pdf>. Acesso em: 8 de setembro de 2016.

BERNERS-LEE, Timothy. **Answers for Young People**. [200-]. Disponível em:
<<https://www.w3.org/People/Berners-Lee/Kids.html#What>>. Acesso em: 7 de setembro de 2016.

BUSCALDI, Davide; HERNÁNDEZ-FARIAS, Irazú. **Sentiment Analysis on Microblogs for Natural Disasters Management: A Study on the 2014 Genoa Floodings**. Florença, Itália: International World Wide Web Conference Committee (IW3C2), 2015.

KING, Irwin; LI, Jeixing; CHAN, Kam Tong. **A brief survey of computational approaches in social computing**. Atlanta, Estados Unidos: Proceedings of International Joint Conference on Neural Networks, 2009.

PRINCETON, University. **About WordNet**. [2010]. Disponível em:
<<http://wordnet.princeton.edu/>>. Acesso em: 10 de setembro de 2016.

SALMI, Matheus. **Investigação de Algoritmos de Análise de Sentimento para a Língua Portuguesa**. Anápolis, 2015. Monografia – Curso de Sistemas de Informação, UnUCET, Universidade Estadual de Goiás.