

Análise de sentimentos em comentários de vídeos do YouTube utilizando aprendizagem de máquinas supervisionada

Alexandre Ribeiro Afonso

Doutor em Ciência da Informação pela Universidade de Brasília (UnB) - Brasília, DF - Brasil.
Membro do grupo de pesquisa Research Expert Group for Intelligent Information in Multimodal Environment using Natural language Technologies and Ontologies (R.E.G.I.I.M.E.N.T.O.)

<http://lattes.cnpq.br/7017201402673628>

E-mail: rafonso.alex@gmail.com

Cláudio Gottschalg Duque

Doutor em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG) - Belo Horizonte, MG - Brasil. Líder do grupo de pesquisa Research Expert Group for Intelligent Information in Multimodal Environment using Natural language Technologies and Ontologies (R.E.G.I.I.M.E.N.T.O.)

<http://lattes.cnpq.br/8531105272766089>

E-mail: klauss@unb.br

Data de submissão: 22/07/2018. Data de aceite: 20/03/2019. Data de publicação:.

RESUMO

O artigo descreve um conjunto de três experimentos em análise de sentimentos, especificamente, para comentários textuais em português brasileiro e para um vídeo do YouTube. Utiliza-se o pacote de mineração de dados Weka para filtragem e um classificador baseado em aprendizagem de máquinas supervisionada (SMO). Os experimentos diferenciam-se pelo *corpus* a classificar: o primeiro utiliza três classes de polaridade (positiva, negativa e neutra), o segundo e o terceiro experimentos trabalham com duas classes (negativa e não negativa). No terceiro experimento são selecionadas somente postagens que comentam uma entidade (referente) específica. Os resultados de Acurácia e Medida-F Média são consideravelmente melhores para os experimentos contendo duas classes. O terceiro experimento atinge valores em volta de 81% para as duas medidas citadas, e sugere que quanto mais entidades são comentadas nos discursos do *corpus*, mais difícil seria a classificação de polaridades.

Palavras-chave: Análise de sentimentos. YouTube. Português brasileiro. Aprendizagem de máquinas.

Analysis of feelings in YouTube video comments using supervised machine learning

ABSTRACT

This paper describes a set of three experiments on sentiment analysis, specifically for texts in Brazilian Portuguese. The dataset is collected from comments about a video in YouTube. We use the data mining package Weka for filtering and a classifier based on supervised machine learning (SMO). Experiments are differentiated according to the corpus to be classified: the first experiment classified the texts into three categories of polarity (positive, negative and neutral), the second and third experiments classified in two classes (negative and non-negative). The input texts of the third experiment are about a unique entity (referent). The results of Accuracy and F-Measure Average are considerably better for experiments containing two classes. The third experiment reaches values around 81% for the two measures cited, and suggests that the more entities are commented on in the discourses of the corpus, the more difficult would be the classification of polarities.

keywords: *Sentiment analysis. YouTube. Brazilian Portuguese. Machine learning*

Análisis de sentimientos en comentarios de vídeos de YouTube utilizando el aprendizaje de máquinas supervisado

RESUMEN

En este documento se describe un conjunto de tres experimentos en el análisis de opiniones, específicamente para comentarios de texto en portugués brasileño, por un vídeo de YouTube. El paquete de minería de datos Weka se utiliza para el filtrado y un clasificador basado en el aprendizaje de máquinas supervisado (SMO). Los experimentos son diferenciados por el corpus a ser clasificado: el primero utiliza tres tipos de polaridad (positiva, negativa y neutra), el segundo y el tercer experimentos trabajan con dos clases (negativa y no negativa). En el tercer experimento, sólo se seleccionan los mensajes que comentan sobre una entidad específica (referente). Los resultados de Accuracy y F-Measure Average son considerablemente mejores para los experimentos que contienen dos clases. El tercer experimento alcanza valores alrededor del 81% para las dos medidas citadas, y sugiere que cuanto más entidades son comentadas en los discursos del corpus, más difícil sería la clasificación de polaridades.

Palabras clave: *Análisis de sentimientos. YouTube. Portugués brasileño. Aprendizaje de máquinas.*

INTRODUÇÃO

A descoberta de conhecimento a partir de grandes bases de opinião, advindas das mídias sociais digitais, tem sido denominada mineração de opiniões ou análise de sentimentos. Segundo Lima (2015), há diversas aplicações desta análise, em diversos setores da economia e política: identificação de fatores de risco em empresas, previsão de ganhos futuros, identificação da satisfação de clientes e mapeamento de intenções de voto.

Sobre o tema, Afonso (2017) observa que os estudos envolvendo as mídias sociais são os mais diversos, em áreas múltiplas, e mostram-se complementares. Estudos linguísticos, como o de Bezerra e Pimentel (2016), argumentam, por exemplo, sobre a língua em uso nos comentários das redes sociais, as expressões no ambiente digital e as características do internetês. Os estudos em informática, como os de Araújo, Gonçalves e Benevenuto (2013) ou de Duran, Avanço e Nunes (2015), propõem variedades de técnicas para classificar as postagens, seus sentimentos explícitos, e normalizar automaticamente o internetês para uma linguagem formal. Na ciência da informação e na comunicação é vislumbrada a possibilidade de levantar métricas e estatísticas de opinião sobre um fato ou produto, a partir de dados opinativos, utilizando-se um software de suporte. Cafezeiro, Da Costa e Da Silva Kubrusly (2016) descrevem um moderno panorama em que a ciência da informação e a ciência da computação convergem para além da simples “prestação de serviços” entre os dois campos, e enfatizam um possível modelo de coprodução, em que as ciências colaboram na conformação dos conceitos que as fundamentam.

Sob tal ideia, de complementação entre áreas de conhecimento, este artigo descreve um conjunto de experimentos em análise de sentimentos, aplicando técnicas de descoberta de conhecimento em textos, em um *corpus* coletado de um canal do YouTube e de um vídeo específico, no qual três comentaristas discutem aspectos do filme Batman versus Superman: a origem da justiça, filme escolhido por ter sido alvo de polêmica nas diversas mídias.

O *corpus* contém comentários escritos em português, considerando-se na montagem, anotação e filtragem do *corpus* os critérios citados no trabalho de Afonso (2017).

Nesta pesquisa, procurou-se verificar o grau de acertos de um classificador automático de polaridades em comentários, utilizando o *corpus* compilado. Ou seja, deseja-se saber se a partir de um treinamento do classificador com comentários positivos, negativos e neutros, previamente anotados por um humano, a máquina conseguiria inferir corretamente a polaridade dos comentários restantes coletados. O ponto principal do trabalho encontra-se no experimento três, o qual sugere que a diversidade textual nos comentários do YouTube torna o processo de análise de sentimentos menos certa para o classificador automático.

Para tal tarefa, descrevem-se a seguir as etapas de três experimentos para essa mídia específica, com três corpora de teste, um para cada experimento, extraídos de uma coleta primária, e relata-se a comparação dos resultados obtidos.

CARACTERÍSTICAS DO DISCURSO NO YOUTUBE

O que motivou o desenvolvimento deste trabalho foram as percepções sobre as características do discurso nos comentários do YouTube. A principal observação, que gera o experimento três relatado neste trabalho, é a existência de múltiplas entidades (referentes) nos textos coletados que são comentadas.

Se o número de entidades comentadas é alto, os adjetivos e qualificações para tais entidades também são diversos, pois o usuário pode escrever, por exemplo, que “uma das pessoas do vídeo não deixa mais ninguém falar”, mas seria sem sentido dizer que “o filme não deixa ninguém falar” ou “o roteiro não deixa ninguém falar”. Isso significa que além das entidades comentadas serem múltiplas, a maneira de indicar no texto a positividade ou negatividade está particularmente ligada à entidade comentada, o que gera variadas formas de expressão da opinião.

A grande parte das pesquisas, quando se faz a montagem do *corpus* de treinamento e testes, em análise de sentimentos, ocorre através de uma API (Application Programming Interface), onde é possível selecionar os tweets, ou comentários do Facebook, por palavras-chave ou hashtags, treinar o sistema com uma parte dos comentários coletados e testar com a segunda parte. Para afirmarmos que os conjuntos de dados (os desta pesquisa e com a coleta por uma API) são equivalentes e produzem os mesmos resultados, as mesmas entidades comentadas teriam que estar presentes nos dois conjuntos, porém, isso não ocorre, pois os dados coletados para um único vídeo são sobre entidades variadas.

A multiplicidade das entidades pode ser notada, inclusive, numa única postagem do YouTube: pode-se comentar no mesmo post sobre o filme, os personagens e um dos críticos, ou ainda, encontram-se citações repetidas de uma mesma entidade de quem se comenta, com nomeações e categorizações distintas para essa entidade, fenômeno linguístico apontado como referenciação textual por Koch (2008) e relatado no contexto de mídias sociais por Afonso e Té (2017).

Se línguas diferentes levariam a resultados equivalentes, ou seja, se as taxas de acerto obtidas para o inglês, para o mesmo filme comentado no YouTube, nas mesmas condições, valem para o português, essa é uma questão que também exige investigações. Uma característica das mensagens coletadas no *corpus* de estudo obtido, por exemplo, é o uso de um palavrão para sintetizar um sentimento sobre algum aspecto do filme, podendo ser empregado tanto como intenção de elogio, como crítica negativa; algo recorrente no *corpus* de estudo. Ainda nas características gramaticais, seria necessário verificar, por exemplo, se a quantidade de adjetivos em postagens é equivalente nas diversas línguas, já que um dos importantes parâmetros para análise de sentimentos automática são os adjetivos, conforme descrito por Silva, Lima e Barros (2012).

Outro questionamento seria a respeito de expressões como ironias, sarcasmos e humor: elas seriam usuais em todos os ambientes de discussão?

Ou somente seriam frequentes para temas ou assuntos específicos? A partir da literatura, como apontado em Pedro (2018), sabe-se que a identificação automática de ironias é um problema clássico em análise de sentimentos. Não se poderia afirmar, sem investigações adequadas, que a presença de tais expressões ocorrem igualmente, por exemplo, em ambientes que discutem sobre medicina e programas de TV.

O quadro 1, a seguir, descreve os tipos de entidades comentadas nas postagens, segundo o *corpus* de estudos, contendo 918 comentários para um único vídeo.

Quadro 1 – Classes de entidades comentadas do YouTube, obtidas pela observação dos comentários de um vídeo escolhido para a pesquisa

ator de outro filme específico; ator específico; canal do vídeo ; cena específica; cena em sentido genérico; cenas específicas; cenas em sentido genérico; cinematografia do filme; crítica específica; crítica em sentido genérico; críticas específicas; críticas em sentido genérico; crítico específico; críticos específicos; críticos em sentido genérico; direção de outro filme específico; direção do filme; edição do filme; efeito em sentido genérico; efeitos especiais; elemento do personagem; elemento do vídeo da crítica; enredo do filme; fan service; filme BvS; grupo da platéia; HQ; montagem do filme; outro filme específico; outro roteirista específico; outros desenhos; outros filmes; outros filmes específicos; outro vídeo específico; parte do filme; parte do roteiro; parte dos críticos; partes do filme; partes do roteiro; personagem específico; personagem específico de outro filme; personagens específicos; personagens em sentido genérico; plateia em sentido genérico; produtora do vídeo da crítica; produtora específica; roteirista do filme; roteiristas do filme; roteiro em sentido genérico; tempo do filme; trailer do filme; trailers em sentido genérico; trilha sonora; universo dos super-heróis; vídeo da crítica; vídeos específicos.

Fonte: Elaborado pelos autores.

Observando agora o quadro 2, a seguir, veja que há entidades retiradas de 215 comentários do Twitter (sem repetições de *tweets*, *retweets* e postagens sem sentido para o domínio), por uma API, cuja busca foi efetuada com as palavras-chave: *batman v superman*, *BvS*, *batmanvsuperman*, *a origem da justiça*. Dados coletados entre 1º de dezembro e 16 de dezembro de 2018.

Verifica-se que o quadro 2 possui classes de entidades associadas ao filme, não existem classes associadas a um vídeo do YouTube, com seus específicos comentaristas críticos e suas críticas sobre o filme, nem sobre o canal do vídeo, ou sobre a produtora do vídeo, nem sobre os diversos elementos multimodais presentes no vídeo, entre outras classes para esse vídeo. Ou seja, os comentários do vídeo são diversos, únicos e característicos para aquele vídeo, com suas entidades comentadas sobre o filme, mas muitas relacionadas também aos itens do vídeo.

Quadro 2 – Classes de entidades comentadas no Twitter, obtidas pela extração com palavras-chave

ator específico; ator de outro filme específico; filmeBvS; crítico específico; cena específica; cenas específicas; críticos em sentido genérico; direção do filme; direção de outro filme específico; grupo da plateia; outro filme específico; outros filmes; parte do roteiro; personagem específico; produtora específica; personagem específico de outro filme; pôster do filme; parte do filme; trilha sonora.

Fonte: Elaborado pelos autores.

Além disso, verifica-se no *corpus* de estudo do YouTube, que das 918 postagens, 507 têm a entidade «*Filme Batman versus Superman*» presente, mas 411 postagens não têm essa entidade presente, mesmo renomeada. Ou seja, 44,77% das mensagens coletadas do YouTube não contêm menções ao filme, ao contrário da base advinda do Twitter, que guarda 100% das postagens com menções ao filme, devido às palavras-chave de coleta.

Em tal processo, para evitar generalizações em resultados, coloca-se que campos de estudo como a ciência da informação, a comunicação e a linguística têm um papel de contribuição notável, através das análises textuais e do discurso, e da verificação do impacto de tais variedades nos algoritmos de classificação de sentimentos, como também observado no trabalho de Afonso (2017).

TRABALHOS RELACIONADOS

Considerando-se o português do Brasil, a mineração de opiniões (ou análise de sentimentos) em mídias sociais é um assunto que tem, recentemente, despertado interesse nos grupos de pesquisa nacionais, visto que tem surgido nas publicações, em eventos e periódicos de áreas diversas.

Os capítulos descritos na obra de Farzindar e Inkpen (2015) relatam as adaptações necessárias às técnicas de Processamento de Linguagem Natural para tratar os textos de mídias sociais, com suas características específicas. Os principais desafios, tais como a informalidade das comunicações, os ruídos e o conjunto limitado de características na representação dos textos em tarefas de classificação são problemas descritos pelos autores. Tais problemas também foram identificados na base textual de testes utilizada.

O trabalho de Araújo, Gonçalves e Benevenuto (2013) traz uma descrição geral e a comparação dos métodos de análise de sentimentos, com enfoque computacional. Relata os métodos baseados em léxico de sentimentos que contêm listas de termos, os quais definem sentimentos de raiva, alegria, tristeza, positividade, negatividade, entre outros sentimentos, como também descreve o enfoque baseado em aprendizagem de máquina. Os sistemas que adotam a aprendizagem de máquina são treinados utilizando-se textos anotados com seu respectivo sentimento, objetivando que o sistema possa inferir novas classificações de sentimentos em novos textos, partindo-se do modelo treinado e previamente construído.

Sobre a avaliação no acerto da polaridade de postagens em análise de sentimentos, o trabalho descrito por Balage Filho, Pardo e Aluísio (2013) é um exemplo que descreve os experimentos de análise que utilizam o *corpus* ReLi, descrito em Freitas *et al.* (2012). Esse *corpus* contém resenhas sobre livros em português brasileiro. Na análise de sentimentos, utilizam-se três léxicos, e a técnica de classificação baseia-se em um algoritmo denominado SO-CAL, já descrito em trabalho anterior, para a atividade de classificar as postagens em positivas ou negativas.

Na linha do tempo, em Balage Filho e Pardo (2013) é descrito o sistema NILC_USP, que utiliza não só a técnica baseada em léxicos de sentimentos citada, mas também o sistema classificador de sentimentos que baseia-se em regras e aprendizagem de máquina. O sistema atinge a Medida-F Média (uma das medidas de corretude utilizada em classificação automática) de 56,31% para postagens da rede social Twitter, considerando duas classes de classificação de sentimentos: positiva e negativa.

No trabalho de Avanço e Nunes (2014) descreve-se também um sistema de classificação de polaridades utilizando-se léxicos de sentimentos. Dessa vez, o *corpus* de testes é retirado de um site com resenhas de produtos em português brasileiro. Os resultados mostram que a melhor combinação é a versão do algoritmo que lida com negação e intensificação e usa o léxico de sentimentos SentiLex. A Medida-F Média alcançada, considerando-se apenas as duas classes, positiva e negativa, é de 0,73.

Lima (2015) tem uma abordagem diferenciada para o tema, e ressalta que, em muitos casos, a avaliação desse tipo de dado não estruturado depende do contexto ou domínio específico de uma aplicação. Assim, o uso de ontologia surge como uma ferramenta para dar suporte à classificação de dados e inferência de informação a partir de dados não estruturados. O trabalho apresenta uma revisão sistemática sobre os temas: análise de sentimentos e ontologias. O objetivo específico é responder à questão sobre a contribuição do uso de ontologias para análise de sentimentos em dados não estruturados da Web.

No texto de Duran, Avanço e Nunes (2015) descreve-se um experimento no qual um normalizador baseado em léxico nomeado UGCNormal é proposto, também para corrigir ruídos em textos informais postados por internautas. Aplicando tal normalizador ao *corpus* de testes, obtém-se uma taxa de acertos maior por parte do classificador de polaridade. A Medida-F Média é melhorada do valor 0,736 para 0,758, considerando as classes positiva e negativa. O *corpus* de estudo foi elaborado a partir de resenhas de produtos em português do Brasil.

No ano seguinte, o trabalho de Avanço, Brum e Nunes (2016) traz o projeto de um sistema para classificação da polaridade de resenhas de produtos em português brasileiro. Considerando a classificação em positiva ou negativa, o artigo descreve experimentos utilizando tanto classificadores baseados em léxico de sentimentos, quanto por aprendizagem de máquina, híbridos ou reunidos. Aponta que o valor alcançado para o enfoque de classificação com léxico de sentimentos atinge o valor de Medida-F Média em 0,84 e para o enfoque com aprendizagem de máquina atinge o valor de 0,95.

Malini, Ciarelli e Medeiros (2017) vão além da análise de polaridade e constroem uma metodologia que combina classificação humana de textos com aplicação de algoritmos genéticos de análise de textos, no intuito de analisar sentimentos genéricos (baseado na polarização positivo/negativos) e sentimento específicos, baseados nas seguintes emoções: alegria, raiva, medo, antecipação, desgosto, tristeza, surpresa e confiança. A técnica é aplicada a conteúdos de opinião pública sobre cenários da política brasileira.

Recentemente, os resultados apresentados por Brum e Nunes (2018) foram alcançados utilizando o *corpus* TweetSentBR, constituído por postagens em português brasileiro advindas do Twitter, especificamente para o domínio de programas de TV, e manualmente anotados por diferentes julgadores. Os experimentos testam a classificação da polaridade utilizando seis classificadores

baseados em aprendizagem de máquina. O valor obtido na classificação binária é 80,38% de Medida-F Média e 64,87% quando inclusa a classe neutra.

MATERIAIS E MÉTODOS DE PESQUISA

Para os experimentos descritos neste artigo, utilizou-se um *corpus* textual extraído do YouTube. O *corpus* é composto por 918 comentários de um vídeo que faz críticas e análises ao filme *Batman versus Superman: a origem da justiça*¹. Os comentários textuais escritos coletados sobre o vídeo da crítica estão em português, assim como as falas dos críticos do vídeo. A amostra que gera o *corpus* correspondia, na época, a 15,2% do total de postagens para o vídeo da crítica. A coleta ocorreu entre os dias 15 de janeiro de 2018 e 15 de fevereiro de 2018.

Esse *corpus* foi compilado a partir do *corpus* descrito em Afonso (2017), o qual guardava a metade do número de postagens do *corpus* aqui descrito, e tinha outros propósitos de estudo. No *corpus* deste trabalho, complementou-se o número de postagens a partir do *corpus* descrito em Afonso (2017) até a marca de 1000 postagens, porém, retirando-se as postagens repetidas e fora do universo do filme, tal como propagandas e postagens sem sentido, o *corpus* final chega à marca citada de 918 comentários. Também, só foram consideradas postagens relacionadas diretamente ao vídeo, postagens em resposta à outra postagem foram desconsideradas. É importante frisar que a coleta das postagens é realizada de maneira aleatória, e somente os comentários do vídeo específico citado foram considerados, ou seja, o *corpus* não é originado a partir de outros vídeos.

Para cada postagem do *corpus* é anotada a positividade, negatividade ou neutralidade, seguindo o critério descrito por Afonso (2017), por apenas um anotador humano. No texto do autor citado, a razão, descrição e exemplificação das regras de anotação podem ser observadas.

Para cada postagem do *corpus*, identificam-se os referentes presentes nas 918 postagens. Os referentes são as entidades sobre quem se escreve: o filme, os personagens, a trilha sonora, o vídeo no YouTube, os críticos do vídeo, etc. Um referente é identificado quando uma opinião notável surge sobre ele no texto de comentário. A noção de referente e referenciação considerada, além de suas relações discursivas com a polaridade das opiniões e as regras de anotação de polaridade são também advindas do trabalho de Afonso (2017). A seguir, descrevem-se as regras de anotação de polaridade propostas por esse autor. Para exemplos de anotação de polaridade, veja o artigo citado.

Uma postagem é classificada como positiva:

- se existirem somente positivities para os referentes no texto postado;
- se existir a positividade do referente de hierarquia maior, ou dominante (no caso, foi considerado dominante “filme Batman vs Superman”) com aspectos negativos ou neutros para referentes de menor hierarquia em relação ao referente dominante;
- se só existem referentes de menor hierarquia que “filme Batman vs Superman”, com maior número de positivities para esses referentes.

Uma postagem é classificada como negativa:

- se existirem somente negatividades para os referentes no texto postado;
- se existir a negatividade do referente de hierarquia maior, ou dominante (no caso, foi considerado dominante “filme Batman vs Superman”) com aspectos positivos ou neutros para referentes de menor hierarquia em relação ao referente dominante;
- se só existem referentes de menor hierarquia que «filme Batman vs Superman», com maior número de negatividades para esses referentes.

¹ <https://www.youtube.com/watch?v=FrsKsV2aSyE>

Uma postagem é classificada como neutra:

- se há somente referentes de mesmo nível hierárquico e a subtração do número de opiniões positivas com o número de opiniões negativas sobre os referentes resultam zero;
- se existem na postagem somente referentes sem sentido positivo ou negativo, seja o referente dominante ou não. Isso quer dizer que há somente outro tipo de sentido para os referentes.

As postagens, após a coleta, passam por uma limpeza de dados, deixando-se somente caracteres alfanuméricos. Isso traz ganhos no acerto da polaridade, pois os sinais de pontuação são retirados das palavras deixando-as sem caracteres adicionais, o que contribui para o sistema de classificação identificar determinada palavra no *corpus*. As desvantagens existem, pois comentários contendo notas de avaliação do filme com números e vírgulas são alterados, e quando retirados vão gerar ruídos na sentença original. Expressões e junções de caracteres, que remetem à positividade ou negatividade, também poderão ser alteradas ao retirar sinais de pontuação e símbolos não alfanuméricos do texto do usuário, o que pode também gerar ruídos. O ideal seria o tratamento prévio do corpus (normalização), etapa ainda não realizada neste trabalho, mas que deve ser mais bem investigada futuramente.

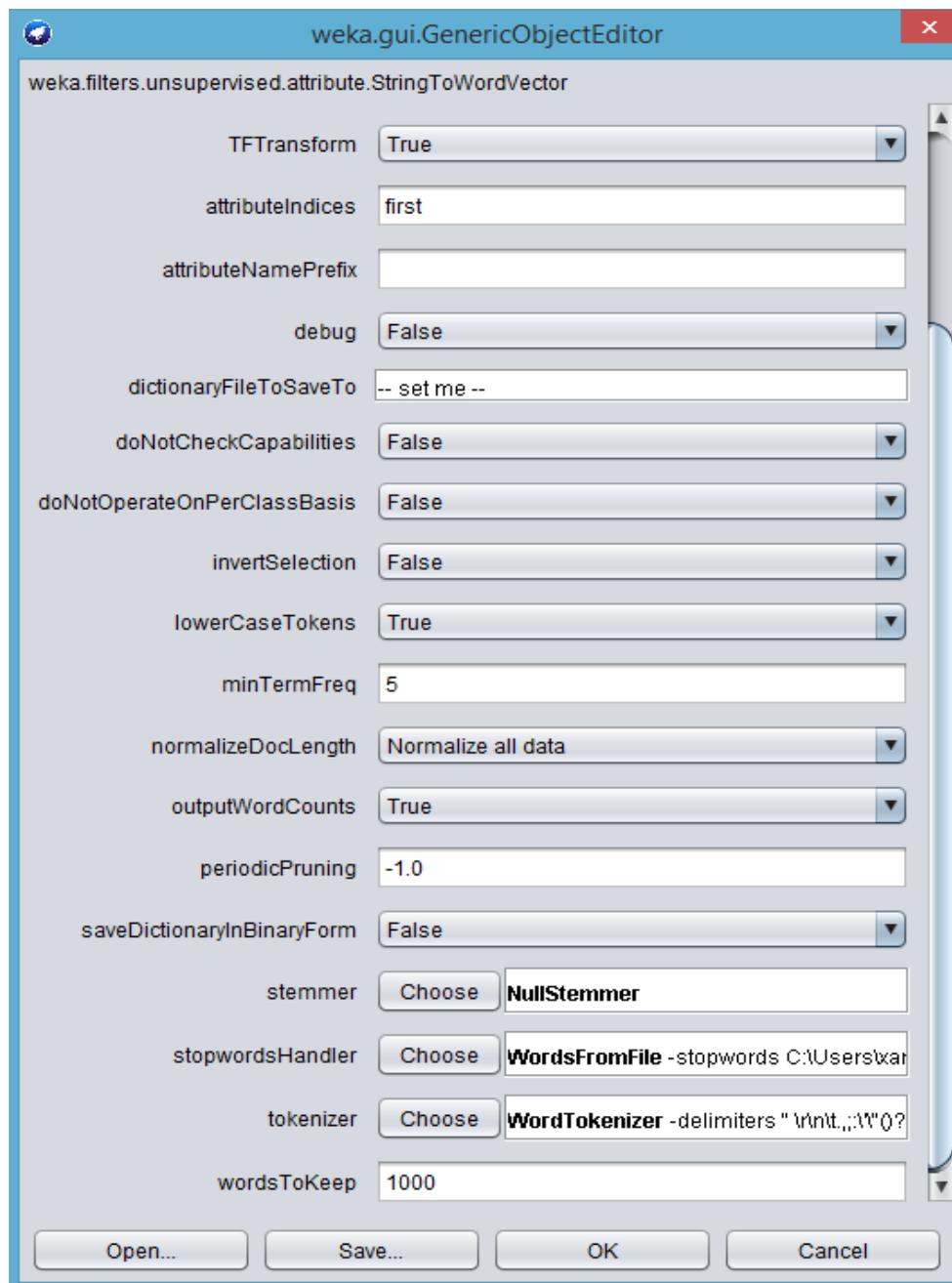
Para o *corpus* coletado, realizaram-se três experimentos. O primeiro trabalha com três classes: postagens positivas (p), negativas (n) e neutras (u). O segundo experimento trabalha com duas classes de polaridade: negativas (ne) e não negativas (nn), ou seja, as postagens do *corpus* com polaridade neutra foram juntadas com as postagens positivas formando a classe de postagens (nn). O experimento três também trabalha com duas classes: classe das postagens negativas (ne) e não negativas (nn), porém, são consideradas somente postagens em que o referente dominante “filme Batman vs Superman” está presente, ou seja, há pelo menos uma opinião sobre o filme na postagem coletada.

O *corpus* original foi então subdividido em três *corpora* diferentes para os três experimentos citados. Cada *corpus* de experimento foi colocado em um arquivo (.arff) que é a extensão para os arquivos de entrada do Weka. Utilizou-se o Weka na versão 3.8.2. Os arquivos (.arff) possuem dois atributos: o campo “registro” que contém a postagem tratada do tipo *string*, além do campo “class”, que é do tipo nominal e contém a polaridade da postagem, e pode ser de dois ou três tipos dependendo do experimento.

Ao ser aberto pelo Weka, o *corpus* de um experimento no formato (.arff) passa por apenas um filtro que converte cada postagem em um vetor de palavras, e para cada postagem, se uma palavra existe no *corpus*, um peso *tf.idf* é associado pelo Weka para tal palavra da postagem. A pesagem *tf.idf* tem sido amplamente utilizada na classificação e agrupamento automático de textos, como na análise de sentimentos. A medida permite, através de um valor, ter-se uma noção do peso semântico de uma palavra ou sintagma em relação a todo o *corpus* onde ela se insere, baseando-se na frequência dela nos registros. Trabalhos recentes sobre análise de sentimentos para o português, como De Souza, Pereira e Dalip (2017) e Dosciatti, Ferreira e Paraiso (2013) empregam a medida.

Nessa fase, também foi utilizada uma lista de *stopwords* para retiradas de palavras sem peso semântico do *corpus*, como alguns pronomes, preposições, artigos, conjugações do verbo haver, ser, estar e ter. Nessa filtragem, todas as palavras do *corpus* são colocadas em letras minúsculas; não se considerou a existência de palavras compostas e *n-grams*. Também, para que a palavra da postagem possa ser considerada na classificação, ela deve ter ocorrência mínima de 5 (cinco) vezes no *corpus* a classificar. A figura 1, a seguir, ilustra as configurações do filtro aplicado:

Figura 1 – Configurações do filtro Weka StringToWordVector aplicado aos três experimentos



Fonte: Elaborado pelos autores.

A figura 2, a seguir, exemplifica como fica a matriz de dados a ser classificada pelo Weka após a filtragem e a pesagem das palavras. Na primeira coluna temos o número da postagem, na segunda coluna a classe a que pertence, manualmente anotada por um único

observador segundo as regras citadas anteriormente, e nas colunas subsequentes as palavras do *corpus* com seus pesos *tf.idf*.

Figura 2 – Relação de postagens (cada linha numerada), sua anotação de polaridade manual (segunda coluna) e pesos *tf.idf* de cada palavra do *corpus* (valores numéricos)

No.	class	achar	achei	acho	adorei	agora	ai	ainda	algumas	alguns	algum	amei	ano	antes	apenas	apocalipse	aquele
Nominal	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
1	nn	0.0	2.61...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	nn	0.0	1.64...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	nn	0.0	0.0	2.20...	0.0	0.0	0.0	2.76...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	nn	0.0	1.64...	2.20...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.132...	0.0	0.0	0.0	0.0	0.0
13	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16	nn	0.0	0.0	0.0	0.0	0.0	0.0	2.76...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.89953173...	0.0
17	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
19	nn	0.0	0.0	0.0	2.851...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.37998...	0.0	0.0
21	nn	0.0	0.0	2.20...	2.851...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
22	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
23	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
26	nn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
27	nn	0.0	1.64...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.287...	0.0	2.89953173...	0.0	0.0
28	nn	0.0	1.64...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fonte: Elaborado pelos autores.

Para o processo de análise de sentimentos testou-se apenas um tipo de classificador, a Máquina de Vetores-Suporte. O teste foi realizado através do pacote de mineração de dados Weka, escolhendo o classificador denominado SMO (*Sequential Minimal Optimization algorithm for training a support vector classifier*) com valores *default* para esse classificador no Weka. O algoritmo foi escolhido porque na maioria dos trabalhos em análise de sentimentos ele tem sido utilizado devido ao alto índice de acertos. A separação de cada *corpus* nas partes de treinamento e testes foi realizada pela validação cruzada (*cross-validation*) em oito pastas.

RESULTADOS E ANÁLISE DE RESULTADOS

Os experimentos realizados objetivaram verificar como ocorreria a classificação dos comentários pelo algoritmo SMO, após a formatação dos dados e a utilização dos critérios e ferramentas descritas na sessão anterior.

O Experimento um utilizou três classes de polaridade: mensagens positivas (p), negativas (n) e neutras (u), sendo 407 positivas, 380 negativas e 131 neutras. Todas as postagens do *corpus* são consideradas, o que soma 918 comentários. A taxa de Acurácia alcançada foi de 64,4% e Medida-F Média em 0,629.

Em seguida, para o Experimento dois, duas classes foram utilizadas nos testes: mensagens negativas (ne) e não negativas (nn), sendo 538 não negativas e 380 negativas. Todas as postagens do *corpus* são consideradas, o que soma 918 comentários. A taxa de Acurácia alcançada foi de 73,9% e Medida-F Média em 0,736.

Finalmente, no Experimento três, duas classes também foram utilizadas, de negativas (ne) e não negativas (nn), sendo 335 não negativas e 172 negativas, o que soma 507 comentários. Somente as postagens com o referente ou entidade “filme Batman vs Superman” opinada foram selecionadas. Obteve-se a Acurácia em 81,04% e Medida-F Média em 0,806.

A tabela 1 resume a descrição dos resultados obtidos.

Tabela 1 – Resultados dos experimentos de análise de sentimentos com o algoritmo SMO

	Num. de Classes	Acurácia	Medida-F Média
Experimento 01	3	64,4%	0,629
Experimento 02	2	73,9%	0,736
Experimento 03	2	81,04%	0,806

Fonte: Elaborado pelos autores.

A melhor maneira de analisar os resultados é pela comparação entre os experimentos. Os experimentos um e dois possuem os mesmos dados de entrada, com as entidades possíveis descritas no quadro 1 deste artigo. Os dados são em mesmo número, e o que diferencia os experimentos um e dois são os números de classes de polaridade: são três classes para o experimento um (positivas, negativas e neutras) e duas para o experimento dois (negativas e não negativas). Como já visto nos trabalhos citados, o número de classes influencia nos resultados, sendo os resultados com duas classes com melhores valores obtidos.

O experimento dois mostra diferença próxima de sete pontos percentuais em relação ao experimento três, para as duas medidas. Eles têm o mesmo número de classes (negativas e não negativas), sendo que o experimento três tem maior Acurácia e Medida-F Média.

O experimento três tem como característica a presença somente de comentários que contenham, necessariamente, a entidade “*Filme Batman versus Superman*” citada, ou seja, são postagens que comentam sobre o filme. Os valores maiores nas medidas desse experimento sugerem que, quando se faz a restrição das postagens em relação a uma ou conjunto de palavras-chave, os valores de medidas aumentam, já que não há a diversidade de várias entidades comentadas, como mostra o quadro 1, exibido anteriormente neste artigo. Também deve-se considerar que quem guia a polaridade do comentário é a entidade dominante “*Filme Batman versus Superman*” presente em todas as postagens.

A maior contribuição desta pesquisa, considerando as pesquisas já realizadas em análise de sentimentos para o português brasileiro, é o apontamento sobre a influência das características da informação textual sobre os algoritmos de classificação e seus resultados, algo pouco trabalhado nas diversas pesquisas desse campo de estudo. Espera-se que a partir desse apontamento realizado, novas perspectivas (sociais, linguísticas e culturais) surjam em tal área, ou seja, a verificação de como tais fatores influenciam ou direcionam os algoritmos de classificação da informação.

CONSIDERAÇÕES FINAIS

Este artigo descreve um conjunto de três experimentos, em classificação de polaridades, para textos de comentários de um vídeo específico postado no YouTube.

O objetivo dos testes foi verificar o potencial da Máquina de Vetores-Suporte como classificador de polaridade, utilizando três *corpora* montados e filtrados a partir de uma única coleta. O que motiva tais experimentos sobre o YouTube é que ele possui características peculiares nos seus comentários de vídeo: existem referências a diversas entidades nos textos opinativos, conforme relata o quadro 1 do texto.

Ainda, a coleta dos textos ocorreu não por uma palavra-chave ou *hashtag* de vários usuários que postaram (como geralmente se faz para pesquisas com o Twitter), mas sim, são comentários de um único vídeo. A ideia, portanto, foi coletar comentários e verificar se a partir do treinamento do sistema com uma parte das postagens anotadas, a outra parte seria então classificada, nas suas polaridades, corretamente. As regras de anotação prévia das mensagens coletadas em polaridade positiva, negativa ou neutra são descritas em Afonso (2017).

O estudo permitiu iniciar a elaboração de uma metodologia para análise de polaridades de comentários para um único vídeo, quando as entidades comentadas são diversas.

Mesmo que os resultados não sejam próximos de 100%, considerando as medidas citadas, o método poderia ser adaptado e utilizado como suporte para uma análise realizada manualmente, assim como ocorre com os tradutores automáticos, que não possuem uma tradução totalmente certa para determinados formatos textuais, mas prestam auxílio para o tradutor humano.

O experimento três, que obteve taxas de acerto em volta de 81% considerando duas classes (negativo e não negativo), exigiria antes do treino e da classificação automática a localização de uma entidade específica (referente) nos textos de treinamento e de entrada. Se o analista quisesse treinar o sistema e classificar apenas mensagens sobre os personagens do filme, por exemplo, nesse caso, haveria a necessidade novamente de um filtro que selecionasse apenas tais postagens contendo essas entidades, e tal filtro ainda teria que ser planejado e construído.

Os resultados para duas classes (negativo e não negativo), nos experimentos dois e três, apresentam resultados de medição animadores e, talvez, seja possível aumentar esses valores ainda mais. Seria também necessário verificar se esses resultados ocorrem em outros *corpora* advindos de outros vídeos para uma confirmação de resultados. A neutralidade mostra-se uma classe com maior dificuldade de detecção por parte do classificador, fenômeno também descrito em outros trabalhos para outras mídias, como relatado nos trabalhos referenciados.

REFERÊNCIAS

AFONSO, A.R. A referenciação em textos do YouTube: um estudo com vistas à análise de sentimentos Referenciation in YouTube Texts: a study for Sentiment Analysis. *Liinc em Revista*, v. 13, n. 2, 2017.

AFONSO, A.R.; TÉ, J. Um estudo sobre referenciação e a construção da opinião a partir de um corpus textual extraído do YouTube. *Domínios de Linguagem*, v. 11, n. 2, p. 339-350, 2017.

ARAÚJO, M.; GONÇALVES, P.; BENEVENUTO, F. Métodos para análise de sentimentos no Twitter. In: BRAZILIAN SYMPOSIUM ON MULTIMEDIA AND THE WEB, WEBMEDIA, 19., 2013, Salvador. *Proceedings...* Salvador: Brazilian Computer Society, 2013.

AVANÇO, L.V.; BRUM, H.B.; NUNES, M.G.V. Improving opinion classifiers by combining different methods and resources. In: Encontro Nacional de Inteligência Artificial e Computacional (ENIAC), 13., 2016, Recife. *Anais[...]*. Recife: [s.n.], 2016. p. 25-36, 2016.

AVANÇO, L.V.; NUNES, M.G.V. Lexicon-based sentiment analysis for reviews of products in Brazilian Portuguese. In: BRAZILIAN CONFERENCE ON INTELLIGENT SYSTEMS, 3., 2014, São Carlos. *Proceedings[...]*. Canada: IEEE, 2014. p. 277-281.

BALAGE FILHO, P.P.; PARDO, T.A.S.; ALUÍSIO, S.M. An evaluation of the Brazilian Portuguese LIWC dictionary for sentiment analysis. In: BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY, 9., 2013, Fortaleza. *Proceedings [...]*. Fortaleza: Sociedade Brasileira de Computação, 2013.

BALAGE FILHO, P.; PARDO, T. NILC_USP: A hybrid system for sentiment analysis in twitter messages. In: CONFERENCE ON LEXICAL AND COMPUTATIONAL SEMANTICS, 2013, Georgia. *Proceedings[...]*. Georgia: SemEval@NAACL-HLT, 2013. p. 568-572.

BEZERRA, B.G.; PIMENTEL, R.L. Normativismo linguístico em redes sociais digitais: uma análise da fanpage língua portuguesa no facebook. *Trabalhos em Linguística Aplicada*, v. 55, n. 3, p. 731-755, 2016.

BRUM, H.B.; NUNES, M.G.V. Building a Sentiment *Corpus* of Tweets in Brazilian Portuguese. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 11., 2018, Miyazaki. *Proceedings [...]*. Miyazaki, 2018. p. 4167-4172. v.1.

CAFEZEIRO, I.; DA COSTA, L. C.; DA SILVA KUBRUSLY, R. Ciência da Computação, Ciência da Informação, Sistemas de Informação: Uma reflexão sobre o papel da informação e da interdisciplinaridade na configuração das tecnologias e das ciências. *Perspectivas em Ciência da Informação*, v. 21, n. 3, p. 111-133, 2016.

DE SOUZA, K.F.; PEREIRA, M.H.R.; DALIP, D.H. UniLex: Método Léxico para Análise de Sentimentos Textuais sobre Conteúdo de Tweets em Português Brasileiro. *Abakós*, v. 5, n. 2, p. 79-96, 2017.

DOSCIATTI, M.M.; FERREIRA, L.P.C.; PARAISO, E.C. Identificando emoções em textos em português do Brasil usando máquina de vetores de suporte em solução multiclasse. *In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL- ENIAC*, 2013, Fortaleza. *Anais [...]*. Fortaleza: [s.n.], 2013.

DURAN, M.S.; AVANÇO, L.; NUNES, M.G.V. A normalizer for ugc in brazilian portuguese. *In: WORKSHOP ON NOISY USER-GENERATED TEXT*. China: Association for Computational Linguistics-ACL, 2015.

FARZINDAR, A.; INKPEN, D. (org.). *Natural language processing for social media*. [S.l.]: Morgan & Claypool Publishers, 2015.

FREITAS, C. *et al.* Vampiro que brilha... rá! Desafios na anotação de opinião em um *corpus* de resenhas de livros. *In: ENCONTRO DE LINGUÍSTICA DE CORPUS*, 11., 2012, São Carlos. *Anais [...]*. São Carlos: Instituto de Ciências Matemáticas e de Computação da USP, 2012.

KOCH, I.G.V. Como se constroem e reconstroem os objetos-de-discurso. *Revista Investigações*, v. 21, n. 2, 2008.

LIMA, P.O. Revisão sistemática sobre uso de ontologia para análise de sentimento em conteúdo da Web. *Ciência da Informação*, v. 44, n. 3, 2015.

MALINI, F.; CIARELLI, P.; MEDEIROS, J. O sentimento político em redes sociais: big data, algoritmos e as emoções nos tweets sobre o impeachment de Dilma Rouseff. *Liinc em Revista*, Rio de Janeiro, v. 13, n. 2, 2017.

PEDRO, G.W. *COMENTCORPUS: Identificação e pistas linguísticas para detecção de ironia no português do Brasil*. 2018. 95 f. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal de São Carlos. São Carlos, 2018.

SILVA, N.R.; LIMA, D.; BARROS, F. Sapair: Um processo de análise de sentimento no nível de característica. *In: INTERNATIONAL WORKSHOP ON WEB AND TEXT INTELLIGENCE (WTI)*, 4., 2012, [S.l.]. *Anais[...]*. [S.l.: s.n.], 2012.