

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-8 – Informação e Tecnologia

ANÁLISE DE SENTIMENTOS UTILIZANDO O ALGORITMO NAIVE BAYES MULTINOMIAL

SENTIMENT ANALYSIS USING THE NAIVE BAYES MULTINOMIAL ALGORITHM

Lúcia Helena de Magalhães - UFMG/ IF SUDESTE MG

Edimilson Batista dos Santos - UFSJ

Igor Campos Moraes - UFSJ

Rogers Ricardo de Avelar Carvalho - UFSJ

Modalidade: Trabalho Completo

Resumo: Os Serviços de Redes Sociais *Online* tornaram-se importantes para usuários expressarem seus sentimentos. Os comentários postados nesses ambientes são valiosos para as empresas saberem as opiniões dos clientes sobre seus produtos e serviços, para governantes tomarem conhecimento sobre o ponto de vista dos cidadãos a respeito das suas propostas políticas, para candidatos perceberem a reação dos eleitores a respeito da campanha eleitoral além de outras finalidades. Assim, este artigo tem como objetivo fazer a análise dos sentimentos expressos nos *tweets* usando o algoritmo de classificação *Naive Bayes* Multinomial. Nos experimentos, utilizaram-se os *tweets* referentes ao segundo turno das eleições presidenciais do Brasil em 2018 e comparou o resultado com outros contextos. Primeiramente, coletaram-se os *tweets* em tempo real utilizando a Interface de Programação de Aplicativos do *Twitter*. Em seguida, os comentários foram pré-processados e as bases de treino e de teste foram criadas. Por fim, os sentimentos foram classificados como negativos, positivos e neutros. Os resultados mostraram que a metodologia apresentou bom desempenho, mas para isso a amostra de treino precisa ter um tamanho razoável e apresentar características de todas as classes para que o algoritmo possa adquirir um bom conhecimento e assim, obter melhor performance na classificação dos sentimentos.

Palavras-Chave: Análise de sentimentos; redes sociais; Algoritmo *Naive Bayes* Multinomial.

Abstract: Online Social Networks' Services have become important for users sentiments expression. The comments posted in these environments are valuable for businesses to know customers' opinions about their products and services, for governors to know about citizens' views on their political proposals, for candidates understanding of the electors' reaction to their campaign, along with further purposes. Therefore, this paper aims to analyse sentiments expressed in tweets using the Naive Bayes Multinomial classification algorithm. During the experiments, tweets from the Brazilian presidential election in 2018 were used and the results were compared to other contexts. Firstly, real time tweets were collected through na Application Programming Interface from Twitter. Later, the comments were preprocessed and training and tests bases were created. Lastly, the sentiment were classified as negative, positive or neutral. The results showed a good methodology performance. However, the training sample must have a reasonable size and present characteristics of all classes, so that the algorithm would be able to acquire good knowledge and hence, obtain a better performance in sentiment classification.

Keywords: Sentiment analysis; social networks; Naive Bayes Multinomial Algorithm.

1 INTRODUÇÃO

A web é um ambiente democrático onde pessoas de diferentes culturas, localizações e pontos de vista podem trocar informações, opinar e criar conteúdo de qualquer tipo de assunto. É um dos maiores repositórios de informação do mundo contemporâneo (LEITÃO, 2004) e a tendência é crescer cada vez mais, pois a quantidade de dados no formato digital tem aumento de forma exponencial (MAÇADA; CANARY, 2014). Segundo Dobre e Xhafa (2014) são gerados no mundo aproximadamente um bilhão de gigabytes, sendo aproximadamente 90% desses dados oriundos de fontes não estruturadas. Existem estimativas que em 2020 o mundo já terá gerado 40 trilhões de gigabytes (SIVARAJAH et al., 2017). Conseguir conhecimento útil nesse ambiente é uma tarefa cada vez mais desafiadora, pois é impossível analisar manualmente esse volume de dados disponíveis na rede, mas o desenvolvimento de novas tecnologias tem possibilitado a análise e visualização de informações a partir de uma grande quantidade de dados.

Nesse contexto, o Aprendizado de Máquina é a área de maior destaque em análise de *Big Data*, pois permite a criação de modelos computacionais para processamento automático ou semiautomático. Uma dessas aplicações é a Análise de Sentimentos ou Mineração de Opinião, área que tem ganhado o interesse dos pesquisadores nos últimos anos. Dentre suas aplicações, têm-se a criação de sistemas de recomendações como o catálogo personalizado do Netflix¹, otimização de mecanismos de busca, como as sugestões de busca pelo Google², análise de mercado para *feedback* de produtos e serviços, predições do preço de ações do mercado financeiro e análise de comportamento em geral.

Assim, partindo da percepção do *Twitter*³ como um canal de difusão de informações, que proporciona a postagem de mensagens de até 140 caracteres (*tweets*) para uma rede de seguidores, bem como a troca de mensagens entre usuários de forma pública (*Replies*) e privada (*Messages*) (RECUERO; ZAGO, 2011, p.21), este artigo tem como objetivo analisar os sentimentos dos usuários expressos nessa rede social durante o segundo turno das eleições de 2018 para presidente da república do Brasil e comparar os resultados com outros fatos que também aconteceram nesse mesmo ano.

¹ <https://www.netflix.com>

² <https://www.google.com.br/>

³ <http://www.twitter.com>

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

A coleta de dados de conjuntura política foi feita no dia 28 de outubro, data que correspondia o dia de votação do segundo turno das eleições. Os demais contextos utilizados para a comparação foram: Competição entre Flamengo e São Paulo pelo campeonato brasileiro de futebol no dia 04 de novembro de 2018; *Reality Show Master Chef* Brasil durante a noite de eliminação, em 06 de novembro de 2018; Lançamento de novo videoclipe da cantora Anitta, dados coletados no dia 09 de novembro de 2018; *Reality Show A Fazenda* durante competição denominada “Prova de Fogo”, também em 09 de novembro de 2018.

O artigo foi estruturado da seguinte forma: a seção 2 apresenta alguns conceitos relevantes e descreve alguns trabalhos relacionados ao tema desta pesquisa. A seção 3 detalha a metodologia, desde a coleta de dados até a visualização dos resultados. A seção 4 faz uma análise das informações obtidas a fim de definir a porcentagem de assertividade do modelo. A seção 5 conclui a pesquisa além de listar melhorias para implementação em trabalhos futuros.

2 ANÁLISE DE SENTIMENTOS

Análise de Sentimento ou Mineração de Opinião é uma área da Mineração de Textos cujo objetivo não é categorizar os textos por tópicos e sim classificá-los com base na emoção contida em determinado documento. Normalmente, a classificação é baseada em sentimentos positivos, negativos ou neutros. Para Pang e Lee (2008), a análise, de uma forma mais abrangente, é usada para o tratamento computacional de opinião, sentimento e subjetividade em textos. Para esses autores, o objetivo é usar Processamento de Linguagem Natural (PNL), estatísticas ou métodos de Aprendizado de Máquina para extrair, identificar ou caracterizar esse tipo de comentários em uma unidade de texto.

Segundo Esuli (2006), Análise de Sentimento é um campo da linguística computacional, em que a preocupação é descobrir a opinião que um documento expressa sobre tal tópico ou produto ao qual se refere. De acordo com Liu (2010), essa área é útil tanto para aplicações voltadas para empresas e organizações quanto para indivíduos. Porém, no Brasil, esse ramo de atuação ainda é bem recente. Pang e Lee (2008) alegam que em outros países, principalmente nos Estados Unidos, já tem muitas empresas que se dedicam a oferecerem serviços de Análise de Sentimentos. Os autores também destacam que a aplicabilidade dessa área é grande e apresentam algumas, tais como, análise de resenhas; sistemas ligados ao

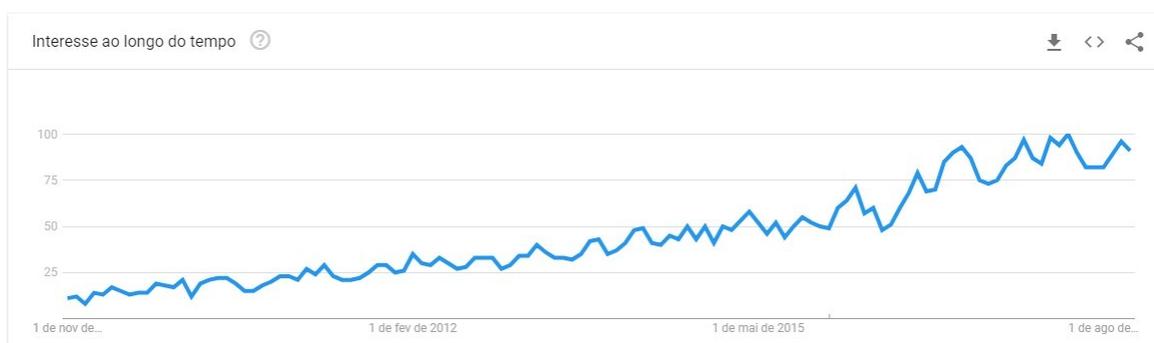
**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

mercado e ao governo e aplicações através de diferentes domínios como, por exemplo, Sociologia, Administração ou Direito.

Segundo Foschiera (2012, p. 14), “o tratamento computacional da opinião vem representando um desafio na atualidade, quando temos à disposição um grande número de informações relevantes que nos permite entender o ponto de vista dos outros”. Kim e Hovy (2006) alegam que, em decorrência disso, a identificação automática dos sentimentos de documentos eletrônicos tem se tornado o centro da atenção de diferentes campos de pesquisas.

A análise de sentimentos possui grande potencial devido à variedade de informações que ela pode revelar. Saber a opinião das pessoas sobre um determinado assunto influencia positivamente na tomada de decisões, além de permitir a detecção de tendências ou comportamentos sociais em geral. O *Google Trends*⁴ mostra o aumento do interesse por esse assunto nos últimos 10 anos ao fazer uma pesquisa pelo tema “Análise de sentimento”, conforme ilustrado na Figura 1.

Figura 1: Progressão de pesquisas por Análise de Sentimentos nos últimos 10 anos.



Fonte: Extraído do Google Trends (2018).

A fim de representar a linha de construção da metodologia, optou-se por abordar primeiramente os artigos que trazem um embasamento que justifique a escolha do *Twitter* como o Serviço de Redes Sociais Online (SRSO) provedor dos dados. Posteriormente, são apresentados trabalhos sobre análise de sentimentos que utilizam ou não Aprendizado de Máquina para finalmente justificar a escolha do método *Naive Bayes* Multinomial (NBM).

Um estudo feito por Recuero e Zago (2010) buscou categorizar o conteúdo das informações no *Twitter* e concluiu que mais da metade dos *tweets* são de conteúdo

⁴ <https://trends.google.com.br/trends/?geo=BR>

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

informativo, sendo que 25,3% são de opinião. Foi apontado ainda que o tipo de capital social que o usuário busca é do tipo conectivo, ou seja, prioriza a conexão com vários usuários e não o aprofundamento de comunicações pessoais.

Partindo do entendimento do *Twitter* como sendo o melhor SRSO para análise de sentimentos, Moraes e Magalhães (2017) desenvolveram uma metodologia para extrair a opinião dos usuários a partir das *hashtags*, que são etiquetas que apontam de forma específica um termo que não apenas constrói um contexto, mas também permite que o *tweet* seja buscado e recuperado e que, em geral, é representada pelo sinal “#” (RECUERO, 2014, p.61). O estudo utilizou técnicas de aprendizado e obteve bons resultados. No entanto, a metodologia é muito restrita a um assunto específico, o que evidencia a necessidade de modelos mais elaborados. Outros exemplos que demonstram a limitação de metodologias são abordados por Lima e Castro (2012) e Lugonov e Panchenko (2011), em que ambos utilizaram um dicionário de *emoticons* para classificação de sentimento em frases. Apesar de obterem altos índices de acurácia, a abrangência é extremamente baixa, pois nem todas as frases contêm esses símbolos.

Souza, Pereira e Dalip (2017) fizeram uma comparação entre métodos para análise de sentimentos em texto que usaram Aprendizado de Máquina supervisionado ou não supervisionado. Segundo os autores, a maior eficiência foi alcançada em modelos que utilizavam o aprendizado supervisionado.

Dentre vários algoritmos supervisionados, Gonçalves *et al.* (2013) compararam oito deles e o resultado mostrou que há uma grande discrepância entre as técnicas. Em uma das bases de testes, a média de diferenciação entre o modelo mais eficiente e os sete demais é de mais de 50%. Algumas tecnologias utilizam a mesma metodologia, porém bases de treino diferentes, mostrando a importância da fase de treinamento para uma boa análise de dados.

Para a escolha do método de aprendizagem deste trabalho, foi constatado que o trabalho de Filho, Carvalho e Pappa (2014) alcançou ótimos resultados utilizando o NBM. O estudo criou dois classificadores de identificação a partir de *tweets*, o primeiro para classificar o sexo do usuário e o último, a idade. Os *datasets* para treino foram rotulados manualmente e aplicados nos modelos *Naive Bayes*, *Random Forest*, *Support Vector Machine (SVM)* e *Multinomia Naive Bayes*. O NBM foi o mais bem-sucedido, atingindo acurácias próximas de

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

90% e 80%, respectivamente. Por isso ele foi o escolhido para ser usado nos experimentos desta pesquisa.

2.1 Naive Bayes Multinomial

O classificador *Naive Bayes* é considerado ingênuo (*naive*) por assumir que os termos de uma instância são condicionalmente independentes entre si, não exercendo influência um sobre o outro. Os termos apenas influenciam a classe a que se remetem (SCHMITT, 2013). É amplamente utilizado em classificação de textos e sua análise é baseada na ocorrência das palavras no documento a ser analisado. Segundo Baeza-Yates e Ribeiro-Neto (2013), os classificadores Bayesianos podem ser abordados de diferentes formas, algumas delas são os modelos binário e multinomial. De acordo com Schmitt (2013), no modelo binário, cada documento é representado por um vetor binário e a presença ou ausência de um termo são representadas pelos valores 1 ou 0, respectivamente. Nessa técnica, a proximidade entre uma palavra e outra é ignorada, pois ele assume que as palavras não possuem relações umas com as outras. Já no modelo multinomial, cada documento é representado por um vetor de inteiros que indica o número de vezes que cada termo acontece no documento (SCHMITT, 2013).

Nesta pesquisa, optou-se por utilizar o multinomial. Desse modo, não foi considerado apenas se uma palavra ocorre ou não em um texto, mas também o número de ocorrências de cada uma em toda a base de treino. Para esclarecer o funcionamento dessa técnica, um exemplo será ilustrado na Figura 2.

Figura 2: Exemplo do funcionamento do modelo NBM.

	Negativo	Neutro	Positivo	
eu	1	0	1	<p align="center">BASE DE TREINAMENTO</p> <p>Eu não gosto desse partido. - Negativo -1 -1 -1 -1 -1</p> <p>Eu gosto desse partido. - Positivo 1 1 1 1</p> <p>Bolsonaro é do PSL. - Neutro 0 0 0 0</p> <p>Haddad é do PT. - Neutro 0 0 0 0</p> <p>Esse é o pior partido do Brasil. - Negativo -1 -1-1 -1 -1 -1</p> <p>Esse é o pior candidato do Brasil. - Negativo -1 -1-1 -1 -1 -1 -1</p> <hr/> <p align="center">BASE DE TESTE</p> <p>Bolsonaro é o pior candidato do Brasil. - Negativo</p> <p>Haddad é o pior candidato do Brasil. - Negativo</p> <p>O pior país do mundo é o Brasil. - Neutro</p> <p>Daciolo é o melhor candidato do Brasil. - Neutro</p>
não	1	0	0	
gosto	1	0	1	
desse	1	0	1	
partido	2	0	1	
bolsonaro	0	1	0	
é	2	2	0	
do	2	2	0	
PSL	0	1	0	
haddad	0	1	0	
PT	0	1	0	
esse	2	0	0	
o	2	0	0	
pior	2	0	0	
candidato	1	0	0	
brasil	2	0	0	

Fonte: Elaborada pelos autores.

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

Ao rotular uma sentença, todas as palavras recebem uma ocorrência da classe correspondente ao rótulo. Ao final da leitura da base de treino, o algoritmo define a classe de cada palavra e “aprende” a classificar novos períodos não rotulados. Uma oração na base de teste recebe a classe que possui mais termos representantes, conforme observado nas duas primeiras sentenças que, por possuírem maior número de palavras negativas, foram classificadas como ‘Negativo’.

Observa-se na da terceira sentença um erro na classificação, pois ela deveria ser classificada como ‘Negativo’, mas foi rotulada como ‘Neutro’. Isso ocorre devido à ausência de palavras da sentença na base de treino. Então o algoritmo, por padrão, define as palavras ausentes como neutras, ocasionando maior ocorrência de palavras neutras do que negativas. Nota-se outro problema na quarta sentença, pois foi rotulada como ‘Neutro’, porém a classificação correta seria ‘Positivo’. Conforme aconteceu na terceira sentença, a falta de sentenças positivas na base de treino faz com que o algoritmo não aprendesse a rotular sentenças positivas.

Por conseguinte, a amostra de treino precisa ter um tamanho razoável, deve apresentar características de todas as classes de forma que o algoritmo possa adquirir conhecimento para fazer a predição de uma coleção não rotulada. Assim, mais uma vez, prova-se a importância da base de treino para o bom funcionamento de qualquer metodologia que utiliza aprendizado supervisionado.

3 METODOLOGIA

O algoritmo foi desenvolvido na linguagem *Python 3*, que oferece métodos intuitivos e documentações oficiais do *Twitter* sobre coleta e análise de dados, além de técnicas para criar, sem dificuldade, modelos de *Machine Learning*. Assim, utilizaram-se nesta pesquisa a biblioteca *tweepy* que tem a função de comunicar com a *Application Programming Interface* (API) do *Twitter* e buscar os *tweets*; a biblioteca *pandas* que fornece ferramentas para manipulação, leitura e visualização de dados; a *scikit-learn* que é a biblioteca de Aprendizado de Máquina e a biblioteca *matplotlib* para a geração dos gráficos. O código do experimento está disponibilizado no GitHub⁵. Usou-se também o software Anaconda que

⁵ https://github.com/igorcmoraes/Sentiment_Analysis_MultinomialNB

possui entre seus pacotes de aplicações o *Jupyter Notebook*, ferramenta que facilita a organização, visualização e compartilhamento de códigos em *Python*.

Para melhor entendimento da metodologia, dividiu-se a descrição da mesma em três tópicos: coletor, treino e teste.

3.1 Coletor

O SRSO escolhido para a coleta de dados foi o *Twitter*, pelas características já mencionadas e também pela facilidade que a plataforma tem de conectar com outras aplicações. Deste modo, utilizou-se a *Streaming* (API) para coletar as publicações feitas nessa rede social em tempo real. Assim, os *tweets* que apresentavam a palavra-chave informada pelo usuário foram coletados e armazenados em arquivos txt, separados por tema.

Como os *tweets* podem conter nenhuma ou várias quebras de linha, adotou-se um identificador de final de cada comentário para evitar a mistura deles. Por conseguinte, no final de cada *tweet*, foi acrescentado o caractere “|” que foi escolhido por ser pouco usual e, portanto, eficaz na função atribuída a ele.

3.2 Treino

Araújo *et al.* (2014) desenvolveram um *framework online* capaz de analisar sentimentos contidos em textos publicados em SRSO. A ferramenta, conhecida como iFeel⁶ tem uma interface que facilita a interação entre o usuário e os vários métodos de classificação nela implementados. Assim, o utilizador pode enviar seus arquivos para serem analisados e fazer um comparativo dos resultados obtidos em cada técnica. Deste modo, fez-se o teste usando uma parte dos *tweets* coletados nesta pesquisa e comparou os resultados dos 18 modelos de classificação de sentimentos do iFeel. Cinco deles obtiveram excelente resultados: OPINION LEXICON, SO-CAL, SANN, STANFORD e EMOLEX. Desta forma, a primeira ideia foi adicionar os *datasets* de cada modelo na base de treino, porém eles não são disponibilizados para o público. Assim, a solução foi usar o iFeel para rotular uma amostra dos *tweets* coletados e adicionar os resultados na base de treino. Por último, anexou-se ao *dataset* uma base rotulada para o estudo piloto deste trabalho contendo 1.250 *tweets*, totalizando uma base final contendo 45.556 sentenças rotuladas.

⁶ www.ifeel.dcc.ufmg.br

3.3 Teste

A base de teste foi formada por 342.518 *tweets* que foram divididos conforme ilustrado no Quadro 1.

Quadro 1: Divisões dos *tweets* por assunto

Assunto	Período	Quantidade de <i>tweets</i> coletados
Eleições	Durante a votação	132.966 <i>tweets</i>
Eleições	Durante a apuração	21.120 <i>tweets</i>
Eleições	Divulgação do Resultado	69.889 <i>tweets</i>
Futebol	Durante o jogo São Paulo x Flamengo	42.297 <i>tweets</i>
MasterChef	c	19.756 <i>tweets</i>
Videoclipe da cantora Anitta	No lançamento do novo videoclipe	26.601 <i>tweets</i>
A Fazenda	Durante a competição “Prova de Fogo”	29.889 <i>tweets</i>

Fonte: elaborada pelos autores.

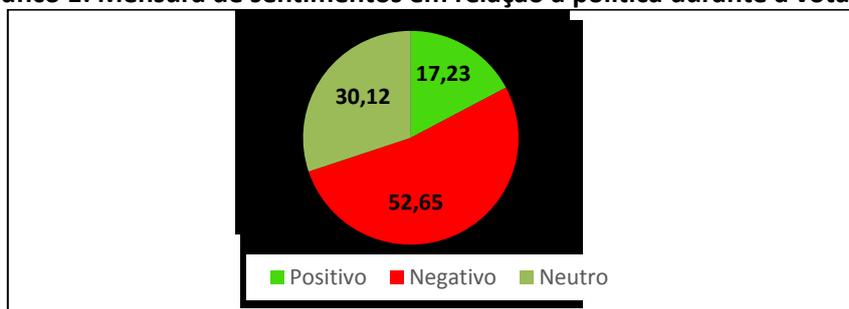
Para transferir os *tweets* para o modelo de predição usando o *Python* é necessário que os dados estejam em formato de lista. Transformar um arquivo de texto em lista não é uma tarefa tão simples, então, inicialmente o arquivo foi armazenado em uma *string* para, em seguida, a sequência de caracteres ser transformada em uma lista de *tweets*. Primeiramente, os dados da *string* passaram por um processo de planificação, onde foram retiradas todas as quebras de linha, pois elas são danosas para o modelo de predição. Com a *string* planificada, a lista foi criada e o caractere “|” foi usado para finalizar cada *tweet*.

4 RESULTADOS

A partir dos resultados das classificações de cada comentário, o número de ocorrências de cada classe foi passado como parâmetro para a criação dos gráficos, utilizando a biblioteca *matplotlib* do *Python*. O Gráfico 1 apresenta as porcentagens de sentimentos positivos, negativos e neutros dos *tweets* coletados durante a votação do segundo turno das eleições presidenciais do Brasil em 2018.

Observa-se no Gráfico 1 que durante a votação, mais de 50% dos comentários eram negativos, sendo 30,12% neutros e 17,23% positivos.

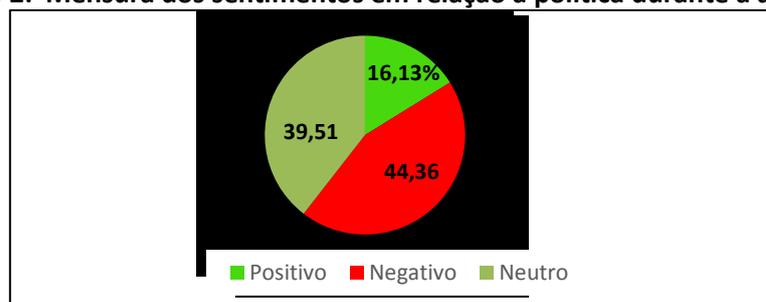
Gráfico 1: Mensura de sentimentos em relação à política durante a votação.



Fonte: Elaborada pelos autores

O Gráfico 2 apresenta as porcentagens de sentimentos positivos, negativos e neutros durante o período de apuração do resultado do segundo turno da eleição para presidente do Brasil no ano de 2018.

Gráfico 2: Mensura dos sentimentos em relação à política durante a apuração.

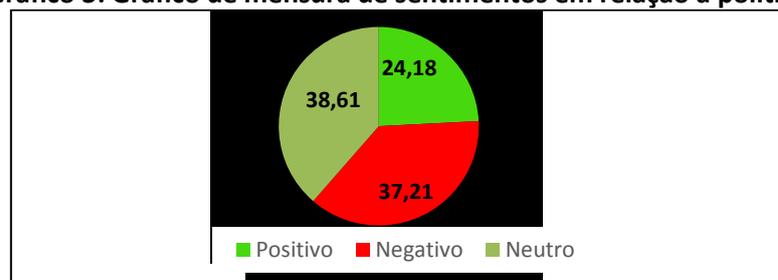


Fonte: Elaborada pelos autores

Houve uma redução dos comentários negativos durante a apuração, mas esse sentimento ainda predominou. Observa-se no Gráfico 2 que 44,36% dos sentimentos relacionados ao tema polícia eram negativos, 39,51% neutros e somente 16,13% positivos.

O gráfico 3 ilustra a porcentagem de sentimentos positivos, negativos e neutros durante a divulgação do resultado do segundo turno da eleição para presidente da república.

Gráfico 3: Gráfico de mensura de sentimentos em relação à política.



Fonte: Elaborada pelos autores.

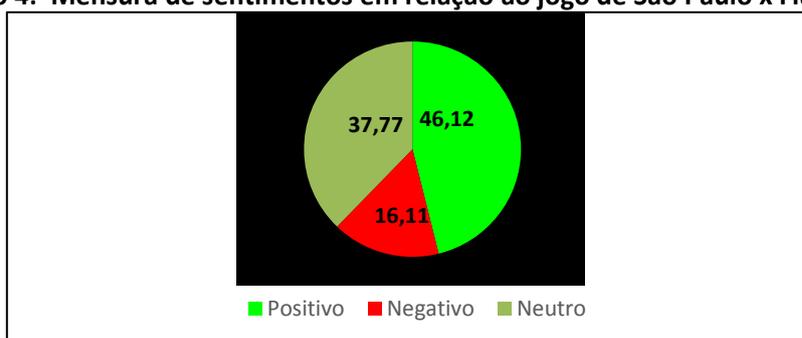
**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

Foi no período da divulgação dos resultados que apareceram mais publicações positivas, porém ainda uma porcentagem baixa, ou seja 24,18%, conforme ilustra o Gráfico 3.

Nota-se que, no contexto político, há uma diferença considerável entre o nível Negativo no momento da votação. Ao ler alguns *tweets* da base de dados, percebem-se que os eleitores apresentavam maior negatividade por expressarem mais o desejo do candidato adversário perder, ao invés de postar comentários positivos sobre o candidato favorito.

Para fazer uma comparação dos sentimentos relacionados à política com diferentes contextos, elaboraram-se gráficos dos comentários sobre o jogo de São Paulo X Flamengo, o *reality show Master Chef Brasil*, o lançamento de novo videoclipe da cantora Anitta e o *reality show A Fazenda*. O Gráfico 4 mostra a porcentagem de sentimentos positivos, negativos e neutros expressos nos comentários dos torcedores durante o jogo de São Paulo e Flamengo.

Gráfico 4: Mensura de sentimentos em relação ao jogo de São Paulo x Flamengo.

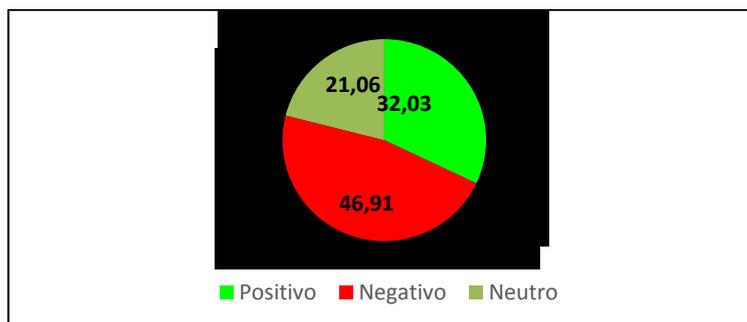


Fonte: Elaborada pelos autores.

Conforme apresentado no Gráfico 4, um total de 46,12% das opiniões foram classificadas como positivas, 37,77% como neutras e apenas 16,11% como negativas. Como são sentimentos expressos durante o jogo, os *tweets* classificados como positivos são mais comuns porque a maioria dos comentários ocorre nos primeiros momentos do jogo, onde os usuários publicam frases de incentivo ao time preferido, o que justifica a predominância desse sentimento.

O Gráfico 5 mostra as porcentagens de sentimentos positivos, negativos e neutros expressos na noite de eliminação do *MasterChef Brasil*.

Gráfico 5: Mensura de sentimentos em relação ao *MasterChef* Brasil.

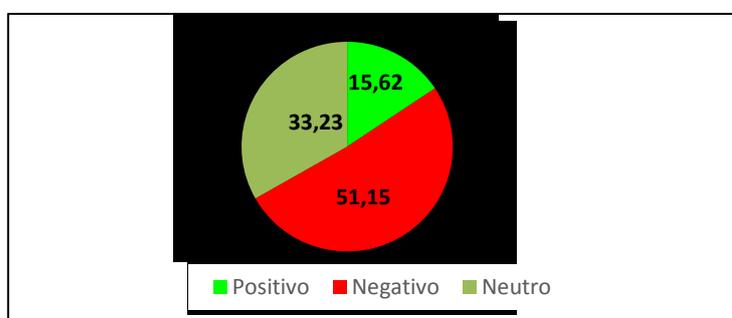


Fonte: Elaborada pelos autores.

Ao observar as porcentagens no Gráfico 5, nota-se que 46,91% dos sentimentos foram classificados como negativos, 32,03% como positivos e 21,06% como neutros. O motivo da negatividade é semelhante ao caso da eleição durante a votação, pois os *tweets* proclamam o desejo da eliminação de certos participantes e não a torcida pela vitória dos *chefs* preferidos.

O Gráfico 6 mostra as porcentagens de sentimentos positivos, negativos e neutros em relação ao lançamento do novo videoclipe da cantora Anitta.

Gráfico 6: Mensura de sentimentos em relação ao videoclipe da Anitta.



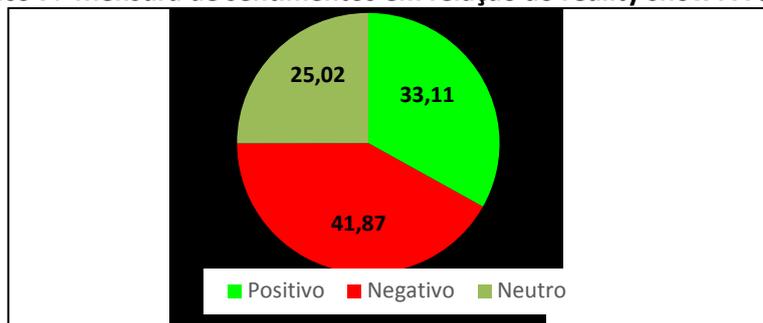
Fonte: Elaborada pelos autores.

Observa-se no Gráfico 6 que 51,15% dos sentimentos foram classificados como negativos, 33,23% como neutros e 15,62% como positivos.

Para entender a negatividade dos sentimentos expressos em relação ao novo videoclipe da Anitta, uma rápida pesquisa na internet revelou que realmente o videoclipe teve uma recepção negativa, como mostra a redação do *Catraca Livre* (2018).

O Gráfico 7 apresenta as porcentagens dos sentimentos positivos, negativos e neutros durante a competição Prova de Fogo do *reality show* “A Fazenda”.

Gráfico 7: Mensura de sentimentos em relação ao reality show A Fazenda



Fonte: Elaborada pelos autores.

Por fim, os comentários sobre o *reality show* “A Fazenda” também apresentaram uma porcentagem grande de negatividade, semelhante ao que aconteceu com o *MasterChef*, ou seja, os torcedores postam comentários negativos sobre quem eles não querem na competição. Com base nas análises, conclui-se que os usuários, ao expressarem suas opiniões em um contexto competitivo, preferem que os adversários tenham desempenho ruim, ao invés de desejarem que os competidores preferidos tenham sucesso.

3 CONSIDERAÇÕES FINAIS

A análise de sentimentos é bastante desafiadora, pois os comentários dos usuários postados nos SRSO são apenas textos em formato livre que, tradicionalmente, os computadores têm dificuldade em interpretar, por isso a necessidade da etapa de pré-processamento para preparar os textos em formatos que os algoritmos de Aprendizagem de Máquina possam decifrar. Desta forma, realizou-se primeiramente nesse estudo uma preparação dos *tweets* e, posteriormente, os comentários dos usuários foram classificados em positivos, negativos ou neutros.

É importante chamar a atenção para as inovações que os Serviços de Redes Sociais *Online* trazem para a ação política na sociedade da informação, fato que promove novos temas de pesquisa, pois ainda há poucos estudos que examinam a complexa relação entre política e SRSO. É necessário realizar mais pesquisas para compreender de forma mais abrangente as opiniões dos cidadãos em relação a assuntos políticos bem como outros temas governamentais.

Além disso, ainda há necessidades de novas pesquisas no campo de Análise de Sentimentos, pois com o advento dos SRSO, a forma como as pessoas expressam seus pontos de vistas mudou drasticamente. Ambientes como fórum e *blogs* representam

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

enormes fontes de informações, no entanto, encontrar as opiniões desejadas e monitorá-las pode ser uma tarefa complexa porque há diversas fontes e cada uma delas pode conter um volume grande de textos opinativos. Devido ao seu grande valor para aplicações práticas, tem havido um crescimento explosivo tanto em pesquisas acadêmicas como em aplicações diversas. No entanto, a maioria das abordagens existentes baseia na estrutura sintática dos textos, que está longe de ser a forma como o ser humano processa a linguagem natural. Deste modo, novas pesquisas que levem em consideração a semântica expressa nesses comentários ainda se fazem necessárias para lidar com as tarefas de análise de sentimentos expressos em diferentes domínios e formatos de textos.

Por conseguinte, este trabalho procurou avaliar o desempenho do algoritmo de classificação *Naive Bayes* Multinomial na classificação de *tweets* publicados durante o período da eleição para presidente do Brasil em 2018 e comparou o resultado com outros contextos fazendo experiências com os *tweets* coletados em mais quatro situações: no decorrer da competição entre Flamengo e São Paulo pelo campeonato brasileiro, na noite de eliminação do *Reality Show Master Chef* Brasil, no lançamento de novo videoclipe da cantora Anitta e durante a competição Prova de Fogo do *Reality Show “A Fazenda”*. Os resultados mostraram que a metodologia apresentou bom desempenho, mas para isso a amostra de treino precisa ter um tamanho razoável e apresentar características de todas as classes para que o algoritmo possa adquirir um bom conhecimento e assim, obter melhor performance na classificação dos sentimentos.

Em trabalhos futuros são almejadas algumas melhorias que não foram aplicadas nesta pesquisa devido ao tempo necessário para implementação:

- a) Inclusão de mais *tweets* rotulados pelos métodos do iFeel:

Não foi possível incluir na base de treino todos os *tweets* que seriam rotulados pelos melhores métodos da plataforma iFeel. Esses *tweets* são de grande importância para o bom funcionamento do algoritmo, pois neles estão palavras específicas dos contextos aplicados que a base de treino até então utilizada não identificou de forma correta. Assim, a proposta é unir frases rotuladas pelos cinco melhores métodos e criar uma superbase.

- b) Implementar Validação Cruzada:

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

A validação cruzada, ou *Cross Validation*, é amplamente empregada em modelos de predição. É um método de reamostragem usado na avaliação de um modelo com o intuito de evitar o teste da técnica no mesmo conjunto de dados no qual ela foi treinada (SOUZA, 2019).

c) Utilização de *Stemming*:

A técnica consiste em reduzir o termo ao seu radical. Na visão de Monteiro e Gomes et al. (2006), *stemming* consiste na remoção de variações de palavras, tais como plural, gerúndio, afixos, gênero e número, de modo que a palavra fique somente com a *stem*, ou seja, com o radical.

d) Filtro de SPAM:

No *Twitter*, muitos usuários utilizam palavras-chave que estão em alta para aumentarem o alcance de um *tweet* e, assim, propagar mensagens que não se relacionam com os temas indexados ou propagar sites não confiáveis. Esses comentários atrapalham a análise de sentimentos e devem ser descartados.

e) Tratamento de idioma:

A *Streaming API* coleta *tweets* do mundo todo que podem apresentar frases em diversos idiomas. A base de treino funciona bem se utilizada apenas em uma linguagem, pois a classificação pode ficar confusa com palavras iguais, mas que em outros idiomas podem apresentar significado diferentes. Portanto, esse problema precisa ser resolvido, seja coletando apenas *tweets* em português, seja traduzindo as sentenças em outros idiomas ou simplesmente excluindo-as.

Além disso, em um próximo trabalho, seria interessante fazer uma comparação entre os resultados alcançados pela técnica utilizada e os resultados obtidos com as bases da plataforma iFeel para verificar a eficiência da metodologia proposta.

REFERÊNCIAS

ARAÚJO, M.; GONÇALVES, P.; CHA, M.; BENEVENUTO, F. iFeel: a system that compares and combines sentiment analysis methods. **Proceedings of the companion publication of the 23rd international conference on World wide web companion**. 2014. Disponível em: https://www.researchgate.net/publication/261959618_iFeel_a_system_that_compares_and_combines_sentiment_analysis_methods. Acesso em: 10 ago. 2019.

BAEZA-YATES, R. RIBEIRO-NETO, B. **Recuperação de Informação**: Conceitos e tecnologia das máquinas de busca. Tradução de Leandro Krug Wives, Viviane Pereira Moreira. 2. ed. Porto Alegre: Bookman, 2013.

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

CARRILHO JUNIOR, J. R. **Desenvolvimento de uma Metodologia para Mineração de Textos.**, 2007. 96p. Dissertação (Mestrado em Engenharia Elétrica) - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

CATRACA LIVRE. **Anitta é acusada de oportunismo após beijar mulheres em clipe.** 2018. Disponível em: <https://catracalivre.com.br/entretenimento/anitta-e-acusada-de-oportunismo-apos-beijar-mulheres-em-clipe/>. Acesso em: 11 nov. 2018.

DOBRE, C.; XHAFA, F. Intelligent services for Big data science. **Future Generation Computer Systems**, [S.l.], v.37, p. 267–281, 2014. <https://doi.org/10.1016/j.future.2013.07.014>

ESULI, A. **Opinion Mining.** Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa, Italy, 2006.

FILHO, R. M.; CARVALHO, A. I.; PAPPA, G. L. Inferência de Sexo e Idade de Usuários no Twitter. In: BRAZILIAN WORKSHOP ON SOCIAL NETWORKS ANALYSIS AND MINING, 2014. **Anais do XXXIV Congresso da Sociedade Brasileira de Computação – CSBC 2014.**

FOSCHIERA, S. M. P. A **semântica da Emoção:** Um Estudo Contrastivo a partir da FrameNet e da Roda das Emoções. 2013. Tese (Doutorado) - Programa de Pós-Graduação em Linguística Aplicada, Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, 2013.

GONÇALVES, P. et al. Comparing and combining sentiment analysis methods. **Proceedings Of The First Acm Conference On Online Social Networks - Cosn '13**, [s.l.], 2013. ACM Press. Disponível em: <<http://blackbird.dcc.ufmg.br:1210/pdfs/cosn127-goncalves.pdf>>. Acesso em: 03 nov. 2018.

LEITÃO, P. J. de O. **Organização da Informação em Subject Gateways.** 2004. Dissertação (Mestrado) - Programa de Pós-Graduação em Estudos de Informação e Bibliotecas Digitais, Departamento de Ciências e Tecnologias da Informação, Instituto Universitário De Lisboa, Portugal, 2004.

LIMA, A. C. E. S.; CASTRO, L. N. de. Uso de emoticons para análise de sentimento de tweets. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 2012, Campina Grande. **Anais...** Campina Grande: CBA, 2012. p. 2555 - 2561.

LIU, B. **Web Data Mining.** Exploring Hiperlinks, Contents, and Usage Data. Springer: Chigago, 2007.

LUGONOV, A.; PANCHENKO, V. Characteristics and predictability of Twitter sentiment series. In: 19TH INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION, 19., 2011, Perth. **Proceedings...** Perth: The Modelling And Simulation Society Of Australia And New Zealand Inc., 2011. p. 1617 - 1623.

MAÇADA, A. C. G.; CANARY, V. P. A Tomada de Decisão no Contexto do Big Data: estudo de caso único. In: **Encontro da ANPAD - EnANPAD**, XXXVIII, 2014.

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

MONTEIRO, L. d. O., I. R. GOMES, et al. Etapas do Processo de Mineração de Textos – uma abordagem aplicada a textos em Português do Brasil. **Anais do XXVI Congresso da SBC - I Workshop de Computação e Aplicações**, Campo Grande, Centro Universitário do Pará, 2006.

MORAES, I. C.; MAGALHÃES, L. H. de. Análise de sentimentos de usuários a partir de hashtags no Twitter: Um estudo de caso do Oscar de 2017. *In: WORKSHOP DE INICIAÇÃO CIENTÍFICA EM SISTEMAS DE INFORMAÇÃO, 4.*, 2017, Lavras. **Anais do WICSI**. Lavras, 2017.

PANG, B.; LEE, L. **Opinion Mining and sentiment analysis**. Computer Science Department, Cornell University, Ithaca, NY 14853, U.S.A., 2008.

RECUERO, R. Contribuições da Análise de Redes Sociais para o estudo das redes sociais na Internet: o caso da hashtag #Tamojuntodilma e #CalaabocaDilma. **Revista Fronteiras – estudos midiáticos**, São Leopoldo, v. 16, n. 2, p.60-77, maio/ago. 2014.

RECUERO, R.; ZAGO, G. “RT, por favor”: considerações sobre a difusão de informações no Twitter. **Fronteiras – Estudos Midiáticos**, [s.l.], v. 12, n. 2, p.69-81, 3 set. 2010.

RECUERO, R.; ZAGO, G. A Economia do Retweet: Redes, Difusão de Informações e Capital Social no Twitter. **Contracampo**, Niterói (RJ), v. 24, n. 1, p.20-43, jun. 2011.

SCHMITT, V. F. **Uma análise comparativa de técnicas de aprendizagem de máquina para prever a popularidade de postagens no facebook**. 2013. 57f. TCC – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

SIVARAJAH, U.; KAMAL, M. M.; IRANI, Z.; WEERAKKODY, V. Critical analysis of Big Data challenges and analytical methods. **Journal of Business Research**, [S.l.], v.70, p.263–286, 2017

SOUZA, A. **Validação Cruzada: Conceito e Exemplo em R**. 2019. Disponível em: <https://pessoalex.wordpress.com/2019/04/16/validacao-cruzada-conceito-e-exemplo-em-r/>. Acesso em: 10 ago. 2019.

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