

Análise de sentimentos expressos no *Twitter* em relação aos candidatos da eleição presidencial de 2022

Sarah Maria Braga Silva¹, Elaine Ribeiro de Faria²

¹Faculdade de Gestão da Informação - Universidade Federal de Uberlândia (UFU)

²Faculdade de Computação - Universidade Federal de Uberlândia (UFU)

{sarah.braga,elaine}@ufu.br

Abstract. *This work aims to analyze the sentiments expressed by Twitter users concerning the candidates for the presidency in the 2022 election to verify if the candidates' performance is related to their popularity on social networks. For this, data from Twitter were collected, pre-processed, and classified with SVM. The results showed that the two most voted candidates in the elections are the ones with the most tweets on the social network. In addition, it was possible to perceive similarities between candidates' approval and performance in some regions of the country. Finally, it is possible to see an evolution in the approval of candidates on Twitter in the days before the election, which represent a challenging scenario for electoral polls.*

Resumo. *Este trabalho visa analisar os sentimentos expressos pelos usuários do Twitter em relação aos candidatos à presidência da eleição de 2022, com o objetivo de verificar se o desempenho dos candidatos está relacionado com a sua popularidade nas redes sociais. Para isso, dados provenientes do Twitter foram coletados, pré-processados e classificados com o SVM. Os resultados mostraram que os dois candidatos mais votados nas eleições são os que possuem mais tweets na rede social. Além disso, foi possível perceber similaridades entre a aprovação dos candidatos e seu desempenho em algumas regiões do país. Por último, é possível perceber uma evolução da aprovação dos candidatos no Twitter nos dias anteriores à eleição, o que representa um cenário desafiador para as pesquisas eleitorais.*

1. Introdução

Redes sociais são plataformas de comunicação que agrupa diversas informações entre elas opiniões e sentimentos expressos por seus usuários [Araújo et al. 2013]. Nos últimos anos, elas têm se moldado cada vez mais como um espaço para seus usuários debaterem e expressarem suas opiniões sobre assuntos em alta.

O *Twitter* é uma das redes sociais mais populares, com 1,3 bilhão de usuários e mais de 500 milhões de *tweets* publicados ao todo por dia [AHLGREN 2022]. Dentre os diversos assuntos debatidos na rede social, destaca-se a política, tema que gera interesse e comentários de várias pessoas. Nesse sentido, o *Twitter* se destaca como um local de divulgação de discussões políticas, sendo uma forma de obter informações políticas rápidas e sem filtro, de diversas fontes diferentes, além de auxiliar no anseio dos usuários que desejam participar ativamente do processo político [Rossetto et al. 2013].

Análise de sentimentos é o ramo da mineração de dados que tem como objetivo analisar opiniões e sentimentos expressos em textos, sendo uma ferramenta popular para análise de comentários realizados em redes sociais [Martins et al. 2015]. Ao ser aplicada no ramo da política, a análise de sentimentos possibilita mensurar a popularidade e aprovação de cada candidato, podendo ser utilizada como um possível meio para realização de pesquisas eleitorais [Belcastro et al. 2020].

Com a realização das eleições presidenciais no ano de 2022, e a grande movimentação de comentários que o evento gerou nas redes sociais, torna-se relevante analisar a opinião dos usuários da rede em relação aos candidatos. Dessa forma, é possível verificar se há uma correlação entre a aprovação dos candidatos nas redes e o resultado das eleições.

Este trabalho tem como objetivo utilizar técnicas de mineração de texto para analisar postagens oriundas da rede social *Twitter*, a fim de verificar se o desempenho de candidatos à presidência da República nas eleições brasileiras de 2022 está relacionado com a sua aprovação nas redes sociais. O sentimento expresso em cada *tweet* de cada candidato é identificado e contrastado com o seu desempenho das eleições. As principais contribuições do trabalho são: i) criação de uma base de dados pública e anonimizada com *tweets* coletados no período referente ao primeiro turno das eleições presidenciais de 2022, entre o período de 01/06/2022 e 01/10/2022; ii) análise da viabilidade do *Twitter* como ferramenta para predição de eleições, comparando dados extraídos do *Twitter* com o resultado final das eleições.

O artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 apresenta a proposta e metodologia utilizadas para realização do trabalho; a Seção 4 discute os resultados obtidos; por último, a Seção 5 conclui o trabalho.

2. Trabalhos relacionados

Existem vários trabalhos publicados com o intuito de analisar a relação entre a aprovação de candidatos em redes sociais e seu desempenho nas eleições presidenciais. Esta seção tem como objetivo apresentar os trabalhos utilizados como referência para realização deste estudo, que foram feitos para eleições presidenciais brasileiras de anos anteriores.

[Cristiani et al. 2020] apresentam um estudo referente às eleições de 2018, uma das áreas de aplicação da análise de sentimentos. Seu principal objetivo foi analisar uma possível relação entre opiniões expressas em redes sociais, no caso o *Twitter*, e o resultado das eleições. Os autores realizaram a coleta de dados em 8 eventos distintos durante a campanha eleitoral, resultando em uma base com 903.518 *tweets*. Após a coleta, os dados foram pré-processados e 600 *tweets* foram classificados manualmente, a fim de serem utilizados no treinamento dos classificadores *SVM* e *Naive Bayes*. Após a implementação e análise de desempenho, o *SVM* se destacou em todas as medidas utilizadas, obtendo 66,66% de acurácia. Após a análise dos resultados, os autores chegaram a conclusão de que o *Twitter* é uma ótima fonte de pesquisa sobre opiniões pois o candidato que obteve o maior número de votos, 55%, também foi aquele que obteve uma maior quantidade de *tweets* com sentimento positivo, 37,23%.

[Dutra and de Rezende Francisco 2018] buscaram compreender a influência do marketing durante o período eleitoral. No trabalho em questão, os autores coletaram

1.204.036 *tweets* utilizando os nomes dos candidatos, *hashtags* e palavras como *eleicoes* e *eleicoes2018*. Após a coleta, os dados foram pré-processados e classificados pelo algoritmo *Naive Bayes*, o qual obteve 85% de acurácia, 90% de precisão e 84% de revocação. Em sua conclusão, os autores apontaram uma maior classificação de *tweets* como positivos para Bolsonaro, com 34% do total de mensagens positivas, sendo ele também o vencedor das eleições em 2018.

O trabalho [Queiroz and Almeida 2020] propôs uma metodologia de análise de sentimentos que realiza a extração, tratamento e classificação dos *tweets* coletados, identificando seu grau de polaridade a partir da aplicação de dicionários para classificação das palavras em positivas ou negativas. Os dados coletados referem-se ao período entre 01 de agosto e 6 de outubro de 2018, totalizando 88 atributos (características associadas aos *tweets* como usuário e data da publicação) e 4.608 instâncias. Para o processo de mineração, a classificação dos textos como positivos, negativos ou neutros foi realizada por meio da utilização de pacotes de dicionários léxicos. Após a classificação, os autores conseguiram perceber relações diretas entre os comentários da rede social e pesquisas de intenção de voto realizadas pelas principais empresas do país. Posteriormente, foi aplicado o algoritmo K-means para identificação de grupos de palavras mais utilizadas. Dessa forma, os autores conseguiram perceber também os principais temas de discussão dos candidatos.

Vale destacar que as eleições presidenciais de 2022 representam um cenário bastante desafiador para análise de polaridade. A disputa entre os dois candidatos principais foi acirrada e até as pesquisas eleitorais não conseguiram prever com precisão o resultado das eleições.

3. Método experimental para contrastar o resultado das eleições com dados do *Twitter*

Esta seção tem como objetivo apresentar os procedimentos metodológicos utilizados neste estudo, que visa analisar se o desempenho dos candidatos nas eleições presidenciais está relacionado com sua popularidade no *Twitter*. A Figura 1 apresenta as etapas realizadas no trabalho, que serão descritas nas próximas seções.

Coleta de dados	<ul style="list-style-type: none">• Coleta de dados a partir de palavras chaves, utilizando a ferramenta <i>snsrape</i>
Criação da base rotulada	<ul style="list-style-type: none">• Realização da classificação manual de parte dos <i>tweets</i> coletados
Pré-processamento	<ul style="list-style-type: none">• Limpeza da base de dados coletada
Construção do modelo	<ul style="list-style-type: none">• Construção de 2 modelos de classificação diferentes
Avaliação dos classificadores	<ul style="list-style-type: none">• Avaliação e escolha final dos modelos através de indicadores de desempenho
Interpretação dos resultados	<ul style="list-style-type: none">• Apresentação dos resultados obtidos com a classificação e comparação com as eleições 2022

Figura 1. Procedimentos metodológicos

Os códigos utilizados na realização deste trabalho estão disponíveis no GitHub¹.

¹bit.ly/projetoAnaliseSentimentos

3.1. Coleta de dados

Para a coleta de dados da plataforma *Twitter* utilizou-se a ferramenta *snsrape*, disponível como uma biblioteca *Python*², que permite extrair *tweets* de um período específico a partir de palavras-chaves.

Os candidatos escolhidos para coleta e subsequente análise, foram aqueles que mais se destacaram em relação ao percentual de intenção de votos, nas pesquisas eleitorais realizadas ao longo do ano de 2022. Durante essa etapa, foram realizadas cinco consultas por candidato, sendo cada uma delas referente a cada região analisada. As palavras-chave utilizadas estão descritas na Tabela 1.

Tabela 1. Palavras-chave utilizadas na busca de *tweets* de cada candidato

Candidato	Palavras-chave
Jair Bolsonaro (PL)	bolsonaro
Lula (PT)	lula
Ciro Gomes (PDT)	ciro
Simone Tebet (MDB)	simone tebet, tebet
Felipe D'avila (Novo)	felipe davila, felipe d'avila

Para cada *tweet* foi armazenado a data da publicação, id, texto, e região. Ao final da coleta, cada candidato listado possuía um conjunto de *tweets* para cada região do país, totalizando portanto, cinco conjuntos.

3.2. Criação da base de dados rotulada

Neste trabalho, para classificar cada texto como positivo, negativo ou neutro, utilizou-se dos mesmos algoritmos do trabalho [Cristiani et al. 2020]: Naive Bayes e SVM³. Por serem modelos supervisionados, foi necessário a construção de uma base de *tweets* rotulados, a fim de obter uma base de treinamento para os classificadores.

Dessa forma, 200 instâncias de cada candidato analisado, escolhidas aleatoriamente, foram rotuladas manualmente. Os rótulos utilizados foram: (1) Positivo, para quando o comentário demonstra apoio ao candidato; (2) Negativo, para quando o comentário demonstra rejeição ao candidato; e (3) Neutro, para quando o comentário não demonstra apoio nem rejeição ao candidato.

É importante destacar que para o treinamento do modelo, os dados foram balanceados utilizando a técnica *undersampling*. Dessa forma, a quantidade total de instâncias contidas na base de treinamento passou de 1000 para 890, sendo 297 instâncias positivas, 297 negativas e 296 neutras.

3.3. Pré-processamento e Extração de Características

Após a coleta e rotularização manual da base de treinamento, os dados foram pré-processados. Esse processo é essencial para a tarefa de mineração de dados, tendo como finalidade limpar e padronizar o dado para as próximas etapas. Neste trabalho, essa etapa foi realizada através da linguagem *Python*, com as seguintes técnicas descritas a seguir.

²Documentação disponível no link <https://github.com/JustAnotherArchivist/snsrape>

³Os algoritmos utilizados neste trabalho foram construídos com bibliotecas *Python* que suportam classificações multinominais

- **Preparação:** foram removidos caracteres especiais, hiperlinks, pontuações, marcações de usuários e acentos, pois não agregam na tarefa de classificação. A remoção foi feita através da biblioteca *Re*, de expressões regulares no *Python*;
- **Remoção da palavra utilizada na busca;**
- **Correção de palavras:** palavras escritas de forma incorreta e gírias foram corrigidas⁴ através da construção de um dicionário em *Python*, o qual continha a palavra incorreta como chave, e a forma correta de escrita como valor;
- **Padronização:** todas as instâncias foram padronizadas em letras minúsculas;
- **Remoção de stopwords:** utilização da biblioteca NLTK do *Python* para remover palavras que são consideradas irrelevantes para a classificação. Foi criada uma lista com palavras adicionais a fim de completar a biblioteca⁵;
- **Tokenização:** cada instância foi quebrada em um conjunto de palavras, utilizando como delimitador o espaço em branco entre elas. A biblioteca *Python* utilizada nessa etapa foi o NLTK, com o módulo *word tokenize*;
- **Lematização:** redução das palavras para sua raiz, retirando as inflexões presentes.

É importante destacar porém que, *hashtags* não foram removidas na etapa de pré-processamento. A técnica utilizada para extração de características dos documentos foi o TF-IDF, escolhido pois equilibra a importância dos termos frequentes e não-frequentes no texto.

3.4. Construção do modelo e Avaliação do Modelo

Seguindo a mesma estratégia do trabalho [Cristiani et al. 2020], para a classificação, foram testados 2 algoritmos: *Naive Bayes* e *SVM*. O método de avaliação escolhido foi o *10-fold cross-validation*. As medidas de avaliação usadas foram acurácia, precisão e revocação, sendo elas escolhidas por serem as mais utilizadas nos trabalhos relacionados a este. Como este estudo trata de um problema de classificação multiclasse (classes Positiva, Negativa e Neutra), utilizou-se o cálculo macro das medidas de avaliação. Assim, calcula-se cada medida para cada classe e a seguir a média entre todas as classes.

3.5. Interpretação dos resultados

A fim de verificar de avaliar os resultados, algumas estratégias foram traçadas:

- Análise das principais discussões que envolviam cada candidato, através das palavras de maior frequência em cada base de dados;
- Criação de gráficos contrastando a aprovação no *Twitter* (sentimento positivo) com o resultado final do primeiro turno por meio de uma análise global;
- Criação de uma tabela comparativa contrastando o percentual de votos que cada candidato recebeu e a correspondente aprovação na rede social das regiões brasileiras.

4. Resultados

Esta seção tem como objetivo descrever os resultados obtidos em cada etapa discutida anteriormente.

⁴Dicionário utilizado disponível em bit.ly/dicionarioPython

⁵Lista completa disponível em bit.ly/stopwordsPython

4.1. Coleta e Pré-processamento

A coleta foi realizada na rede social *Twitter* por meio das palavras chaves apresentadas anteriormente na Tabela 1. A data inicial da coleta foi 01 de junho de 2022, e a final 01 de outubro de 2022, 1 dia antes da eleição do 1º turno. No total, foram coletados 6.778.481 *tweets*.

Após o pré-processamento, a base de dados foi reduzida para 3.481.203 *tweets*. O principal motivo para a redução da quantidade de *tweets* coletados foi a remoção de dados duplicados, que chegaram a compor em torno de 50% da base da maioria dos candidatos. Isso aconteceu devido ao parâmetro *near* da API de coleta, que seleciona os *tweets* feitos próximos à região da cidade informada no parâmetro. Como os dados foram coletados utilizando todas as capitais brasileiras, alguns *tweets* foram coletados mais de uma vez, pois foram considerados "próximos" a mais de uma cidade.

É importante destacar que cada *tweet* foi representando por um conjunto de 3.171 palavras. A Tabela 2 apresenta a quantidade de *tweets* recuperados para cada candidato após o pré-processamento.

Tabela 2. Quantidade de *tweets* por candidato após o pré-processamento

Candidato	Quantidade de <i>tweets</i>
Jair Bolsonaro (PL)	1.618.392
Lula (PT)	1.428.482
Ciro Gomes (PDT)	379.008
Simone Tebet (MDB)	49.987
Felipe d'Avila (Novo)	5.334

Nota-se que os dois candidatos mais comentados nas redes, Jair Bolsonaro e Lula, também foram aqueles mais votados e que conseqüentemente chegaram ao 2º turno. Por outro lado, Ciro Gomes, 3º colocado no ranking de comentários extraídos na rede social, não obteve um bom resultado nas eleições, ficando atrás de Simone Tebet.

4.2. Avaliação dos classificadores

Esta seção apresenta um comparativo entre os classificadores SVM e Naive Bayes na classificação de sentimentos dos *tweets*. A Tabela 3 mostra a matriz de confusão obtida para o modelo construído com o classificador SVM. É possível perceber que uma das falhas do modelo foi o reconhecimento errôneo de classes neutras como negativas. Porém, no geral o classificador obteve 52% de acerto, sendo que para a classe negativa ele obteve o melhor desempenho.

Tabela 3. Matriz de confusão obtida para o classificador SVM

		Classificado como			Total
		Positivo	Neutro	Negativo	
Real	Positivo	131 (44%)	76 (26%)	90 (30%)	297
	Neutro	44 (14%)	126 (43%)	126 (43%)	296
	Negativo	29 (10%)	71 (24%)	197 (66%)	297
Total		204	273	413	890

Já a Tabela 4 mostra a matriz de confusão obtida para o modelo construído com o classificador *Naive Bayes*. Nesse caso, é possível perceber que o modelo construído possui dificuldade em aprender classes neutras, já que não conseguiu prever a maioria das instâncias dessa classe corretamente. Por esse motivo, o modelo preveu 50% das instâncias corretamente, sendo que ele obteve melhor resultado na classe positiva.

Tabela 4. Matriz de confusão obtida para o classificador *Naive Bayes*

		Classificado como			Total
		Positivo	Neutro	Negativo	
Real	Positivo	192 (65%)	55 (19%)	50 (16%)	297
	Neutro	112 (38%)	92 (31%)	92 (31%)	296
	Negativo	90 (30%)	46 (15%)	161 (55%)	297
Total		394	193	303	890

Portanto, com base nos dados obtidos na Tabela 5, o classificador escolhido para este projeto foi o *SVM* por possui as maiores métricas de desempenho.

Tabela 5. Avaliação dos indicadores

	Naive Bayes	SVM
Acurácia	50%	52%
Precisão	51%	54%
Revocação	50%	51%

4.3. Analisando os candidatos por meio de nuvens de palavras

Ao analisar a nuvem de palavras construída com a base do candidato Jair Bolsonaro, presente na Figura 2, é possível perceber algumas palavras em destaque relacionadas à discussões sobre o desempenho de seu primeiro mandato, como acabar, bobo corte, bom, reeleger e defender. A palavra Lula aparece em grande destaque também, sendo possível concluir que os usuários da rede social fazem comparações frequentes entre os dois candidatos.



Figura 2. Nuvem de palavras do candidato Jair Bolsonaro (PL)

Em relação ao candidato Lula, pode-se observar que a Figura 3 contém termos referentes ao tempo que passou na cadeia, como roubar e ladrão. Outro termo presente

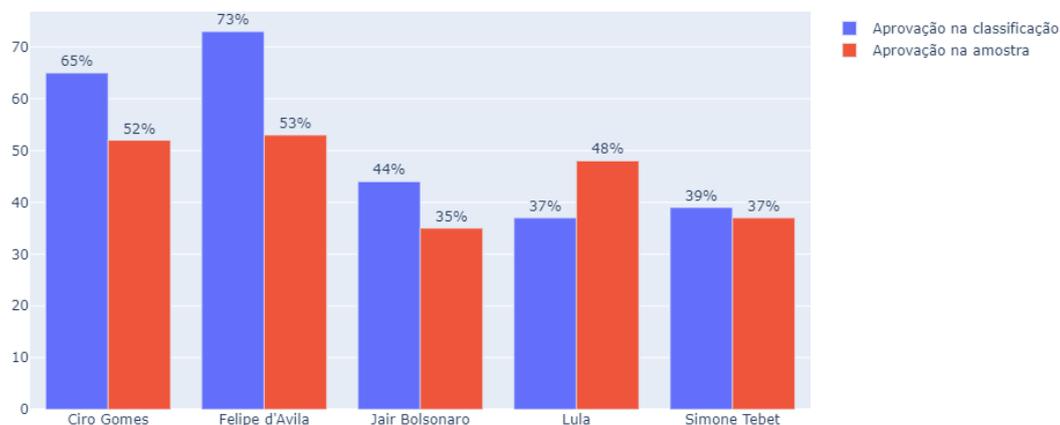


Figura 4. Comparação da porcentagem de aprovação na classificação e amostra por candidato

Em relação ao Jair Bolsonaro, nota-se que a diferença entre a taxa de aprovação e reprovação diminuiu consideravelmente em setembro, o mês anterior à eleição, indicando uma possível melhora no seu desempenho na eleição presidencial. Além disso, o dia anterior à eleição foi a primeira vez que o candidato apresentou uma taxa de aprovação maior que a de rejeição.

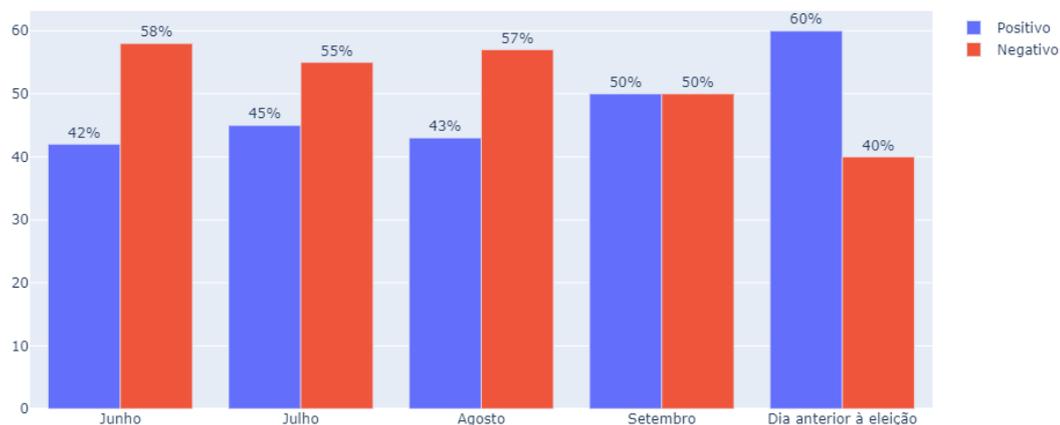


Figura 5. Taxa de aprovação e rejeição por mês - Jair Bolsonaro (PL)

Pode-se perceber um comportamento semelhante em relação ao candidato Lula. O candidato apresentou uma constante evolução em sua popularidade na rede social analisada, chegando no pico no dia anterior à eleição, indicando que talvez se destaque na

votação. Além disso, assim como o candidato Jair Bolsonaro, o dia anterior a eleição foi a primeira vez que o candidato Lula apresentou uma taxa de aprovação maior que a de rejeição.

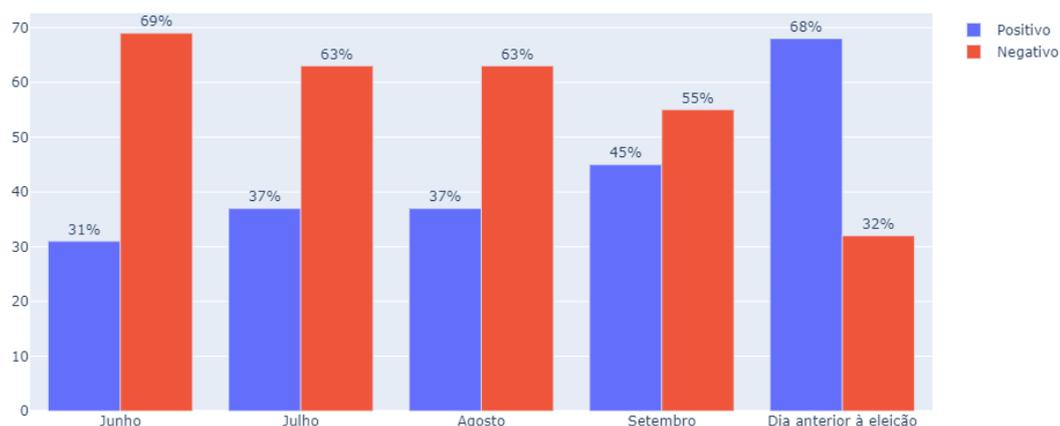


Figura 6. Taxa de aprovação e rejeição por mês - Lula (PT)

4.5. Contraste dos resultados da classificação e resultados oficiais da eleição presidencial de 2022

A Figura 7 mostra um contraste entre a aprovação do candidato na base completa e o percentual de votos recebidos no primeiro turno. Pode-se perceber que não basta avaliar a taxa de aprovação, sem considerar o volume de *tweets* recebidos pelos candidatos.

Percebe-se que os candidatos Jair Bolsonaro e Lula foram aqueles que a taxa de aprovação na rede social mais se aproximou do percentual de votos recebidos. Esse fato pode ser consequência do tamanho da base de dados dos candidatos, já que os dois receberam a maior quantidade de comentários na rede social.

Baseado no experimento realizado, foi possível concluir que não houve semelhanças entre a taxa de popularidade do candidato na rede social *Twitter* e seu desempenho na eleição presidencial, considerando toda a base de dados coleta. Olhando as taxas de aprovação no dia anterior das eleições onde Lula obteve 68% e Jair Bolsonaro 60% pode-se perceber uma semelhança com o resultado do primeiro turno.

Por outro lado, quando analisamos a taxa de aprovação por região, presente na Tabela 6, é possível perceber algumas semelhanças com o resultado oficial da eleição. Em relação ao candidato Jair Bolsonaro, nota-se que o Nordeste é a região com a maior taxa de reprovação, sendo uma das regiões que o candidato menos recebeu votos. Por outro lado, a região Norte aparece com a maior taxa de aprovação, sendo também uma das regiões que o candidato mais recebeu votos.

Já em relação ao candidato Lula, percebe-se que a região Nordeste foi a que o candidato mais recebeu comentários positivos na rede social, e também foi uma das regiões que o candidato mais recebeu votos.

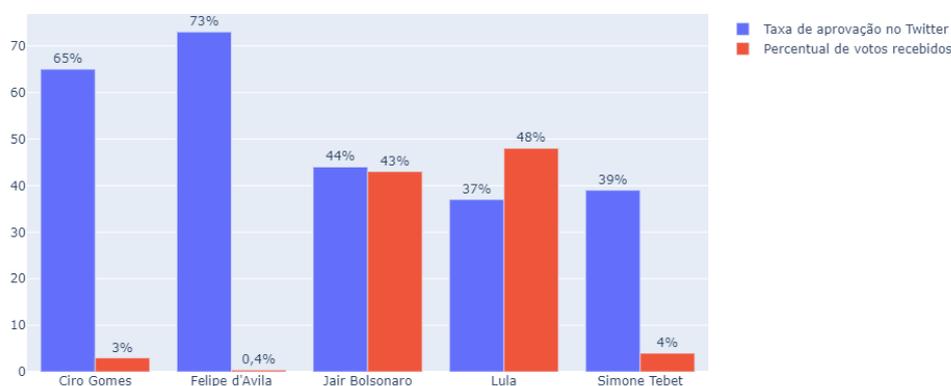


Figura 7. Comparação entre a aprovação no *Twitter* e o percentual de votos recebidos no 1º turno

Tabela 6. Taxa de aprovação por região e por candidato

	Ciro Gomes	Felipe d'Avila	Jair Bolsonaro	Lula	Simone Tebet
Centro-Oeste	65%	53%	44%	38%	39%
Nordeste	65%	70%	42%	43%	43%
Norte	65%	59%	45%	38%	43%
Sudeste	65%	51%	44%	36%	39%
Sul	64%	64%	43%	37%	37%

5. Conclusão

Este trabalho teve como objetivo verificar se o desempenho de candidatos à presidência nas eleições brasileiras de 2022 está relacionado com a sua popularidade nas redes sociais. Para isso, o *Twitter* foi escolhido como fonte de coleta dos dados, devido à sua facilidade e popularidade. Além disso, foram utilizadas as seguintes técnicas para extração de conhecimento textuais a fim de identificar o sentimento presente em cada *tweet*: coleta de dados, criação da base de dados rotulada, pré-processamento de textos, construção de nuvens de palavras, classificação dos textos usando algoritmo SVM e validação dos resultados. Ao fim, o sentimento dos *tweets* foi contrastado com o resultado das eleições.

A partir das análises foi constatado que os candidatos mais comentados nas redes sociais são também os que receberam mais votos. Analisando a popularidade dos candidatos na véspera das eleições, nota-se que os dois que receberam mais *tweets* positivos foram os mais votados. Em especial, nota-se um aumento da porcentagem de *tweets* positivos em relação aos dois principais candidatos ao longo do mês. Por último, em uma análise considerando as regiões do Brasil, é possível traçar paralelos entre a popularidade no *Twitter* e o os votos recebidos na região.

Em relação ao classificador, o resultado obtido poderia ter sido melhor se houvesse um avanço no pré-processamento, como identificação de *tweets* realizados por *bots*, aprimoramento da lista de *stopwords* e correção de palavras/gírias. Além disso, outra técnica que poderia ter sido implementada é a identificação e remoção de textos jornalísticos que

não apresentam opiniões sobre candidatos. Métodos mais robustos para classificação de textos como o BERT (*Bidirectional Encoder Representations from Transformers*) poderiam trazer melhores resultados.

Outro ponto interessante a ser comentado em relação à classificação, é em relação aos comentários coletados. Foi identificado que em um mesmo *tweet*, o usuário pode ter sentimentos conflitantes para candidatos diferentes, como por exemplo, falar mal de um candidato para enaltecer outro que também é mencionado. Nesse caso, como o comentário possui palavras de cunho negativo, o classificador entende que o sentimento expresso é negativo. Outro caso bastante comum foi, elogiar um candidato e no mesmo comentário falar mal de outro, encontrando portanto palavras em tons negativos e positivos em um mesmo *tweet*. Visando esse problema, novas pesquisas mais avançadas têm surgido com o propósito de analisar o sentimento expresso no comentário mais detalhadamente. A Análise de Sentimentos Baseada em Aspectos busca obter o máximo de detalhes sobre a entidade que está sendo classificada, separando para cada comentário o alvo da opinião, sentimento atribuído e categoria correspondente. Dessa forma, é possível realizar a separação de opiniões sobre candidatos em um mesmo comentário. Assim, técnicas mais avançadas como essa poderiam trazer melhores resultados.

Referências

- AHLGREN, M. (2022). 50 + Twitter Estatísticas e Fatos para 2022. <https://www.websiterating.com/pt/research>.
- Araújo, M., Gonçalves, P., and Benevenuto, F. (2013). Measuring sentiments in online social networks. In *Proceedings of the 19th Brazilian Symposium on Multimedia and the Web, WebMedia '13*, page 97–104, New York, NY, USA.
- Belcastro, L., Cantini, R., Marozzo, F., Talia, D., and Trunfio, P. (2020). Learning political polarization on social media using neural networks. *IEEE Access*, 8:47177–47187.
- Cristiani, A., Lieira, D., and Camargo, H. (2020). A sentiment analysis of brazilian elections tweets. In *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*, pages 153–160, Porto Alegre, RS, Brasil. SBC.
- Dutra, D. A. M. and de Rezende Francisco, E. (2018). Text mining: Análise de sentimentos nas eleições 2018. In *Congresso Transformação Digital 2018*.
- Martins, R., Pereira, A., and Benevenuto, F. (2015). An approach to sentiment analysis of web applications in portuguese. In *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web, WebMedia '15*, pages 105–112, New York, NY, USA. Association for Computing Machinery.
- Queiroz, G. and Almeida, L. (2020). Uma metodologia de análise de sentimentos dos candidatos as eleições presidenciais de 2018 no twitter. *Revista de Engenharia e Pesquisa Aplicada*, 5:21–30.
- Rossetto, G. P. N., Carreiro, R., and Almada, M. P. (2013). Twitter e comunicação política: limites e possibilidades. *Compólitica*, 3(2):189–216.