

Analyzing the Impact of Bots During the 2018 Brazilian Presidential Election Based on YouTube Comments

Cristian Amaral Silva
Universidade Federal de Viçosa
Florestal, Minas Gerais, Brasil
cristian.amaral@ufv.br

Daniel Mendes Barbosa
Universidade Federal de Viçosa
Florestal, Minas Gerais, Brasil
danielmendes@ufv.br

ABSTRACT

The tendency of discussing politics in Online Social Networks (OSNs) has become increasingly growing, especially in election periods. This growth may be justified by the little restrictive format of the information to be published and also by many politicians conducting their respective campaigns on these networks due to the limitation of the free political advertising time. According to this trend and to the large amount of data generated by users of these social networks, the objective of this work is to analyze the impact of bots during the candidacy of Fernando Haddad and Jair Messias Bolsonaro in the Brazilian presidential election of 2018 and, additionally, their possible voters, by using comments collected from YouTube. Our results demonstrate that, considering all the unique users of the collected data set, approximately 71.44% of them were considered undecided, 17.67% as potential Bolsonaro voters and 10.89% as potential Haddad voters, with no evidence of political manipulation by bots.

CCS CONCEPTS

• **Information systems** → **Data mining**; *Web searching and information discovery*; • **Networks** → **Online social networks**.

KEYWORDS

Bots, Election, Political Manipulation, Political Votes Estimation, YouTube

ACM Reference Format:

Cristian Amaral Silva and Daniel Mendes Barbosa. 2020. Analyzing the Impact of Bots During the 2018 Brazilian Presidential Election Based on YouTube Comments. In *Brazilian Symposium on Multimedia and the Web (WebMedia '20)*, October 29–November 1, 2020, Maranhão, Brazil. ACM, New York, NY, USA, 8 pages. <https://doi.org/xxx.xxxx/xxxxxxx.xxxxxxx>

1 INTRODUÇÃO

O volume de dados produzidos por usuários de redes sociais tem se tornado cada vez maior em uma escala global e vários estudos já mostraram que dentre esses dados existem muitas informações relevantes e valiosas [1, 3–5, 8, 9]. Com isso, a importância da

mineração dos mesmos quanto a um determinado contexto também aumenta.

Para esse fim, muitas estratégias e ferramentas vêm sendo estudadas e utilizadas, até mesmo em conjunto, como por exemplo a análise de sentimentos e a detecção de posicionamentos, com o objetivo de estabelecer uma relação da opinião das pessoas com a realidade de um assunto específico, como entretenimento, notícias, saúde e política [5].

Considerando o alcance e a taxa de utilização crescentes de redes sociais online e tendo em mente o pouco tempo concedido aos candidatos no horário eleitoral (propaganda eleitoral gratuita em emissoras de rádio e televisão), personalidades políticas passaram a utilizar esses ambientes como possibilidades alternativas de realizarem suas respectivas campanhas políticas e de compartilharem seus posicionamentos com a população.

Partindo dessas considerações e da popularização do YouTube¹, realizamos em [9] a análise da aceitação dos candidatos Jair Bolsonaro e Fernando Haddad durante a eleição presidencial brasileira de 2018 a partir de uma base de, aproximadamente, 7,5 milhões de comentários da rede social em questão, com o auxílio do método léxico de análise de sentimentos SentiStrength [12]. Após a realização desse trabalho, como uma forma de reforçarmos ainda mais os índices de aceitação e, conseqüentemente, de rejeição obtidos, surgiu a ideia de investigarmos também se houve ou não manipulação política a partir de *bots* durante o período eleitoral considerado. Como uma ideia adicional, também tomamos como objetivo a análise de usuários indecisos e possíveis eleitores de cada um dos candidatos.

De forma a alcançar todos os objetivos apontados, foi necessário realizar uma nova coleta dos comentários contidos nos vídeos previamente coletados (entre novembro e dezembro de 2018) para que fosse possível determinar quais os comentários foram excluídos desde o período da coleta original e quais ainda estão disponíveis. É importante mencionar que o atributo de identificação única dos usuários, o qual não tinha sido considerado inicialmente, também foi coletado.

Nossos resultados indicam que não houve indícios de manipulação política por meio de *bots* devido à baixa quantidade de comentários postados por usuários (autores) únicos e pelas configurações de aceitação e rejeição não terem sofrido mudanças consideráveis ao analisarmos separadamente comentários indisponíveis e disponíveis. Também foi possível observar que, dos 1.527.075 usuários únicos da base, 269.875 são possíveis eleitores do Bolsonaro e 166.281 possíveis eleitores do Haddad.

Este texto está organizado da seguinte forma: alguns trabalhos relacionados são apresentados na Seção 2; na Seção 3 são discutidas

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebMedia '20, October 29–November 1, 2020, Maranhão, Brazil

© 2020 Association for Computing Machinery.

ACM ISBN xxx-x-xxxx-xxxx-x/xx/xx...\$15.00

<https://doi.org/xxx.xxxx/xxxxxxx.xxxxxxx>

¹<https://www.youtube.com>

todas as decisões tomadas e as estratégias adotadas para que fosse possível analisar o impacto de *bots* durante a eleição e os possíveis eleitores dos candidatos; todas as análises realizadas e seus respectivos resultados são discutidos na Seção 4; finalmente, na Seção 5, são apresentadas as conclusões que podem ser obtidas a partir das análises realizadas, as limitações do trabalho e o que pode ser feito futuramente para aprimorar os resultados.

2 TRABALHOS RELACIONADOS

Ao longo desta seção serão abordados alguns trabalhos presentes na literatura que possuem relação com o estudo do contexto político-eleitoral em redes sociais online e, mais especificamente, o impacto de *bots* nesse meio.

Tratando-se do idioma inglês, muitos trabalhos foram realizados com o intuito de relacionar o sentimento e a opinião de usuários de redes sociais com os resultados da eleição americana de 2016 [1, 4, 8], na qual Hillary Clinton e Donald Trump disputaram pela presidência dos Estados Unidos. Em [8], Paul *et al.* se empenharam em classificar a afiliação partidária de conteúdos compartilhados no Twitter² com métodos de medir similaridade de textos, avaliando, assim, os estados que mais apoiaram cada candidato. Por sua vez, Ardehaly *et al.* apresentaram, em [1], modelos de aprendizagem de máquina para inferir probabilidades conjuntas diárias de tweets quanto ao sentimento político e atributos democráticos. Já Caetano *et al.*, em [4], utilizaram *hashtags*³ provenientes do Twitter para inferirem e analisarem a variação do humor dos usuários quanto aos candidatos à presidência.

Quanto ao contexto político-eleitoral brasileiro, como já era esperado devido à complexidade sintática e semântica do idioma português, a quantidade de estudos realizados é muito menor. Em [5], Christie *et al.* consideram quatro possíveis candidatos à presidência do Brasil em 2018 para avaliar a popularidade de cada um de acordo com tweets coletados sobre cada um deles.

Considerando unicamente a ocorrência de *bots* em redes sociais em períodos eleitorais, é possível encontrar na literatura muitos estudos realizados [2, 3, 6, 11, 13]. Badawy *et al.* analisaram, em [2], o impacto de contas maliciosas criadas no Twitter por russos com o objetivo de manipular politicamente as pessoas através da ideologia política de sites compartilhados em tweets durante as eleições americanas de 2016. Para essa mesma eleição, Zannetou *et al.*, em [13], utilizaram como base tweets de 1.000 usuários identificados como *bots* russos com o objetivo de quantificar a influência política exercida por eles. Em contrapartida, Subrahmanian *et al.*, em [11], e Davis *et al.*, em [6], partiram do princípio de engenharia reversa para estabelecer técnicas de aprendizagem de máquina para identificar usuários que interagiram com *bots* russos no Twitter. Vale destacar que, por se tratar de um assunto muito complexo, a maioria dos estudos realizados quanto a esse contexto apresentam resultados subjetivos [7].

Apesar de já terem sido realizados diversos estudos nessa área, não foi encontrado nenhum trabalho que aborde o impacto de *bots* em períodos eleitorais ao se utilizar dados provenientes da plataforma de compartilhamento de vídeos YouTube, que nos dias

atuais já é considerada uma rede social. Por conta disso, pode-se considerar que a principal motivação para a realização deste trabalho foi de preencher essa lacuna da literatura.

3 METODOLOGIA

A metodologia proposta e utilizada neste estudo para que pudéssemos realizar as análises propostas foi dividida em cinco etapas subsequentes, as quais serão apresentadas e discutidas ao longo desta seção. Essa divisão foi definida da seguinte maneira: (i) exclusão de vídeos duplicados da base de vídeos original, (ii) verificação da disponibilidade atual dos vídeos, (iii) composição de uma nova base de comentários a partir do identificador dos vídeos disponíveis, (iv) junção externa da base de comentários original com a nova base de comentários, (v) separação de comentários disponíveis e indisponíveis. O fluxo sequencial das etapas estabelecidas está representado visualmente na Figura 1.

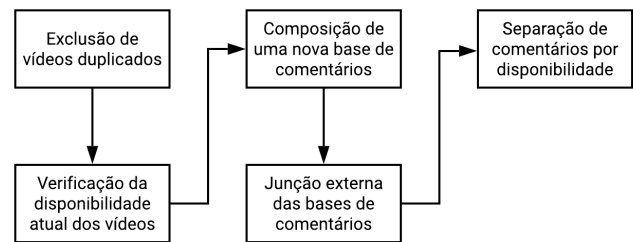


Figura 1: Etapas da metodologia

3.1 Exclusão de vídeos duplicados

Para que a composição da base de comentários referentes aos candidatos Bolsonaro e Haddad pudesse ser realizada, foi adotada uma estratégia de coletas diárias subsequentes, desde o início da campanha eleitoral até o final do segundo turno, por vídeos relacionados com os termos de busca "bolsonaro" e "haddad". Ao fim dessas coletas foi realizada uma união das informações de cada dia, o que possibilitou a composição de uma base de 16.537 vídeos.

Como especificado em [9], existe a possibilidade de vídeos estarem relacionados aos dois termos de busca que foram utilizados durante as coletas, o que acarretou em uma coleta duplicada dos vídeos que se encaixaram nessa condição. Sendo assim, para que as etapas posteriores da nossa metodologia pudessem ser realizadas com o mínimo de requisições possíveis à API de dados do YouTube, que é denominada de YouTube Data API⁴, todas as ocorrências de vídeos duplicados (mesmo identificador) foram eliminadas.

Após essa remoção, a base de vídeos que foi utilizada nas próximas etapas foi reduzida a 15.466 ocorrências únicas. É importante mencionar que apenas o atributo *video_id*, o qual se refere ao identificador único de um vídeo, será considerado neste trabalho.

3.2 Verificação da disponibilidade dos vídeos

Tendo em vista que desde a composição da base de vídeos original muitos vídeos podem ter sido indisponibilizados por diversos motivos, como disseminação de conteúdo de ódio, descumprimento dos

²<https://www.twitter.com>

³Palavras-chave utilizadas para fazer referência a um determinado assunto

⁴<https://developers.google.com/youtube/v3>

termos de utilização do YouTube ou até mesmo por vontade própria dos geradores de conteúdo, decidimos verificar quais deles foram indisponibilizados e quais deles ainda estão disponíveis a partir dos identificadores contidos na base de vídeos únicos.

Para que essa verificação fosse possível, utilizamos um script em Python⁵ para realizar uma requisição à API do YouTube para cada vídeo contido na base de vídeos, de forma que caso um conteúdo nulo fosse obtido, o vídeo foi removido e, caso fosse possível recuperar suas informações, o mesmo ainda está disponível. Ao fim das requisições conseguimos verificar que, do total de vídeos únicos, 10.357 ainda estão disponíveis e 5.109 foram removidos ou deixaram de ter visibilidade pública.

3.3 Composição de uma nova base de comentários

Todos os comentários que estavam associados aos vídeos indisponíveis também foram indisponibilizados, entretanto, também há a possibilidade de comentários de vídeos disponíveis terem sido removidos desde a data da composição da base original. Como esses comentários podem ter sido removidos por uma detecção de *bots* por parte do YouTube, também é muito importante identificar quais são esses comentários.

Uma forma de realizar essa identificação seria efetuar uma requisição para cada comentário da base de comentários original, assim como foi feito para a base de vídeos únicos, entretanto, devido à grande quantidade de comentários e ao limite de 10.000 requisições diárias à API do YouTube, essa estratégia se mostrou inviável. Como uma alternativa à essa limitação, decidimos utilizar outro script em Python para efetuarmos uma nova coleta de comentários a partir dos identificadores dos vídeos únicos disponíveis, pois, dessa forma, em uma única requisição é possível obter informações de 100 comentários.

Vale ressaltar que nenhum dos atributos dos comentários coletados em 2018 podem identificar unicamente seus respectivos autores, fato que comprometeria todo o trabalho. Entretanto, ao pesquisarmos a fundo sobre as possibilidades de atributos retornados pela API, descobrimos que cada usuário do YouTube possui um identificador único de canal, que é atribuído a ele independentemente se o mesmo criou ou não um canal na plataforma. Considerando a importância dessa informação para a análise do impacto de *bots* durante a eleição, o atributo adicional *author_channel_id* também foi considerado na coleta atual de comentários.

Após a nova coleta de todos os comentários referentes aos vídeos que ainda estão disponíveis, foi possível obter, ao todo, 5.258.199 comentários, os quais possuem todos os atributos contidos no Json de exemplo representado na Listing 1.

3.4 Junção externa das bases de comentários

Apesar da limitação de requisições ter sido amenizada pela estratégia de coleta apresentada na etapa anterior, outro problema acabou surgindo: a API do YouTube não permite a coleta de comentários filtrados por um determinado período de tempo, sendo assim, comentários publicados e/ou alterados após o segundo turno eleitoral de 2018 também foram coletados. Em outras palavras, também foram coletados comentários que não estão presentes na base de

```
1 {
2   "author_channel_id": "UCLCjLXct_YqMWTfZNFdFDKA",
3   "comment_id": "Ugxq7SMoTfm9cULG_g14AaABAg",
4   "like_count": 2058,
5   "published_at": "2018-09-01T18:49:04.000Z",
6   "text_original": "Aqui em Cariri no Ceara os bandidos estao
7   ↳ coagindo a populacao a nao fazer comite ou ajudar na
8   ↳ campanha do Bolsonaro...",
9   "updated_at": "2018-09-01T18:49:04.000Z",
  "video_id": "ungFeSiRNhw"
}
```

Listing 1: Exemplo de comentário coletado

comentários original e, conseqüentemente, não foram considerados na análise da aceitação dos candidatos. Dessa forma, não seria condizente com a proposta do trabalho considerá-los nem na análise de manipulação política e nem na de possíveis eleitores dos candidatos.

Para que somente os comentários coletados previamente fossem considerados na análise, mas com a adição do atributo referente ao identificador único dos seus respectivos autores, surgiu a ideia de realizarmos uma junção externa com a base de comentários original e a base de comentários coletada na etapa anterior da metodologia. Seguindo essa estratégia, foi possível reaproveitar por completo o pré-processamento, a detecção de posicionamentos e a análise de sentimentos já realizados em [9], tendo como única adição o atributo *author_channel_id*, que possui um valor válido caso o comentário esteja disponível e possui um valor nulo caso o comentário esteja indisponível.

Ao final da coleta percebemos que mesmo se um comentário estiver disponível, só é possível obter o identificador do canal do seu autor caso o usuário que efetuou a publicação não tiver tornado privado as informações de seu próprio canal. Dito isso, é importante mencionar que como essa informação dos comentários é indispensável para as análises propostas, todos os comentários que se encaixaram nessa condição foram desconsiderados.

Diante de todos os comentários coletados, apenas 216 se encaixaram nesse caso, sendo que desses somente 174 estavam contidos na base de comentários original. Sendo assim, por esses comentários representarem uma parcela muito pequena da base de comentários, os mesmos não são suficientes para comprometer este estudo.

3.5 Separação de comentários por disponibilidade

Após a realização da junção externa detalhada na seção anterior, tornou-se possível dividir a base de comentários resultante em comentários disponíveis e indisponíveis, divisão que foi contemplada como a última etapa da metodologia proposta. Para isso, comentários que possuem um valor válido para o atributo *author_channel_id* são classificados como disponíveis, enquanto comentários que possuem um valor nulo para esse atributo são classificados como indisponíveis.

Considerando todos os 6.173.805 comentários da base resultante da junção externa, exceto aqueles que possuem identificadores privados do canal do autor, 1.594.227 estão indisponíveis e 4.579.404 continuam disponíveis desde o período eleitoral de 2018. É de suma importância evidenciar que os comentários que possuem um identificador inválido de autor só poderão ser utilizados para avaliar

⁵<https://www.python.org>

a diferença de aceitação dos candidatos ao se considerar comentários disponíveis e indisponíveis, já que para as demais análises é necessário se conhecer os usuários únicos da base.

4 EXPERIMENTOS E RESULTADOS

Todas as análises realizadas após a definição e execução das etapas da metodologia com o objetivo de analisar o impacto de *bots* durante o período eleitoral serão explicadas e apresentadas ao longo desta seção.

4.1 Nova análise da aceitação dos candidatos

Como já mencionado ao final da Seção 3.5, os comentários indisponíveis não possuem um identificador válido de autor e, consequentemente, só poderiam ser utilizados em uma análise comparativa da aceitação dos candidatos considerando individualmente os comentários disponíveis e indisponíveis e os resultados obtidos em nosso trabalho anterior [9].

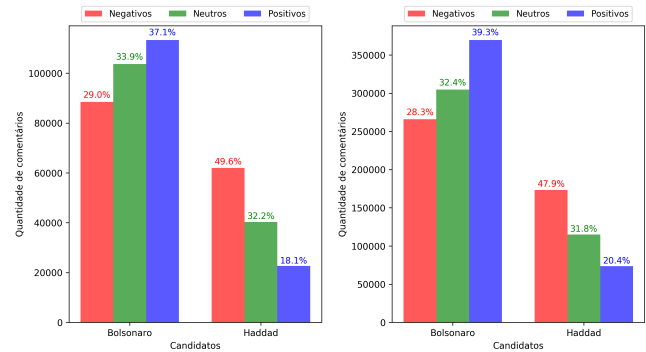
Visto que a base de comentários composta em 2018 não persistiu o atributo *author_channel_id*, decidimos tomar como primeira análise essa comparação de níveis de aceitação, pois só faria sentido verificar se houve manipulação política por meio de *bots* e analisar possíveis eleitores dos candidatos caso a aceitação estabelecida em ambas as classes de disponibilidades não fossem muito diferentes. Em outras palavras, caso os níveis de aceitação e rejeição se diferissem muito, os comentários que foram indisponibilizados possuiriam uma forte tendência de terem manipulado os resultados da eleição, entretanto, como não é possível identificar os usuários por traz desses comentários, o restante das análises ficaria comprometido.

De forma análoga à análise da aceitação realizada em 2018, também tomamos a decisão de realizar essa nova análise considerando separadamente cada turno eleitoral de acordo com duas métricas distintas, sendo elas: desconsiderando os *likes* dos comentários e considerando os *likes* dos comentários. Na primeira métrica, somente a quantidade de comentários é considerada, enquanto a quantidade de *likes* dos comentários é considerada na segunda métrica. Visto que para a análise da aceitação do Bolsonaro e do Haddad os comentários que tiveram a mesma similaridade para ambos os candidatos foram desconsiderados por não causarem impacto nos resultados, os mesmos continuaram sendo desconsiderados nesta nova análise.

4.1.1 Primeiro turno. Quando os *likes* dos comentários foram desconsiderados na análise do primeiro turno realizada em 2018, o candidato Jair Messias Bolsonaro obteve, respectivamente, 38,8%, 32,8% e 28,5% de aceitação, neutralidade e rejeição, enquanto o outro candidato, Fernando Haddad, obteve 19,8%, 31,9% e 48,3%, respectivamente.

Para a nova análise, também desconsiderando os *likes* dos comentários, temos na Figura 2a os resultados obtidos para os comentários indisponíveis e na Figura 2b os resultados obtidos para os comentários disponíveis. Após a comparação dos resultados obtidos quanto aos comentários indisponíveis e disponíveis é possível perceber que a menor diferença entre as classes aconteceu com comentários neutros referentes ao Haddad, que foi de 0,4%, e a maior diferença aconteceu com comentários positivos também referentes ao Haddad, que foi de 2,4%.

Levando em conta os resultados originais, a menor diferença dentre os comentários indisponíveis foi de 0,3% para comentários neutros referentes ao Haddad e a maior foi de 1,7% para comentários positivos relacionados aos dois candidatos. Agora, quanto aos comentários disponíveis, a menor diferença foi de 0,1% para comentários neutros referentes ao Haddad e a maior diferença foi de 0,6% para comentários negativos referentes ao mesmo candidato. Em suma, para essa análise, pode-se observar que a característica dos gráficos não se alterou e as diferenças encontradas foram muito pequenas.



(a) Comentários indisponíveis (b) Comentários disponíveis

Figura 2: Aceitação no primeiro turno desconsiderando *likes*

Para o mesmo turno eleitoral, quando os *likes* dos comentários foram considerados na análise original, o Haddad obteve 22,5%, 33,4% e 44,1% de rejeição, neutralidade e aceitação, respectivamente, enquanto seu adversário obteve 40,4%, 31,8% e 27,8%, respectivamente.

Considerando os *likes* dos comentários, temos na Figura 3a os resultados para os comentários indisponíveis e na Figura 3b os resultados para os comentários disponíveis. Comparando apenas esses resultados, pode-se visualizar que a menor diferença entre as classes foi de 0,1% para comentários positivos referentes ao Bolsonaro e a maior diferença, que foi de 4,5%, aconteceu com comentários neutros relacionados ao Haddad.

Agora, comparando os resultados obtidos com os resultados originais, a menor diferença perante os comentários indisponíveis foi de 0,1% para comentários positivos relacionados ao Bolsonaro e a maior diferença foi de 3,5% para comentários neutros referentes ao outro candidato. Tratando-se do outro conjunto de comentários, a menor diferença obtida foi de 0% dos comentários positivos referentes ao Bolsonaro e a maior foi de 1%, que diz respeito aos comentários neutros que falam sobre o Fernando Haddad. Para essa análise, apesar de se ter observado um leve aumento das diferenças entre as classes ao se comparar com a análise anterior, que desconsidera os *likes*, esse aumento ocorreu, majoritariamente, nas classes de comentários neutros, o que não causa impacto direto na aceitação e nem na rejeição dos candidatos.

4.1.2 Segundo turno. Quando os *likes* dos comentários não foram levados em consideração na análise da aceitação dos candidatos realizada em 2018, o Haddad acabou obtendo 55,3%, 27,9% e 16,9%

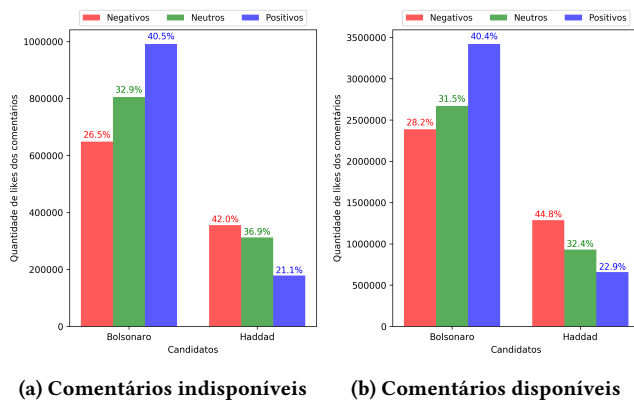


Figura 3: Aceitação no primeiro turno considerando likes

de rejeição, neutralidade e aceitação, respectivamente, enquanto seu concorrente obteve 33,9%, 33,8% e 32,3%, respectivamente.

Quanto à nova análise, a Figura 4a apresenta os resultados para o conjunto de comentários indisponíveis e a Figura 4b apresenta os resultados para o conjunto de comentários disponíveis. Facilmente pode ser identificada uma leve alteração da característica dos gráficos: enquanto no conjunto de comentários indisponíveis o Bolsonaro possui uma rejeição um pouco menor que neutralidade, no conjunto de comentários disponíveis essa configuração é invertida. Faz-se necessário evidenciar que essa alteração não condiz com uma manipulação política pois os níveis de aceitação, neutralidade e rejeição do Bolsonaro na análise original estavam muito próximos, ou seja, uma suave diferença nos resultados já foi suficiente para alterar a característica dos gráficos.

Comparando os resultados originais com os resultados obtidos, pode-se identificar que a menor diferença obtida dentre as classes dos comentários indisponíveis foi de 0,3% para comentários positivos referentes ao Jair Bolsonaro e a maior diferença foi de 2% para comentários negativos sobre o seu adversário. De maneira parecida, dentre os comentários disponíveis, a menor diferença foi de 0,1% para os comentários positivos sobre o Bolsonaro e a maior foi de 1% para a classe de comentários negativos referentes ao Fernando Haddad.

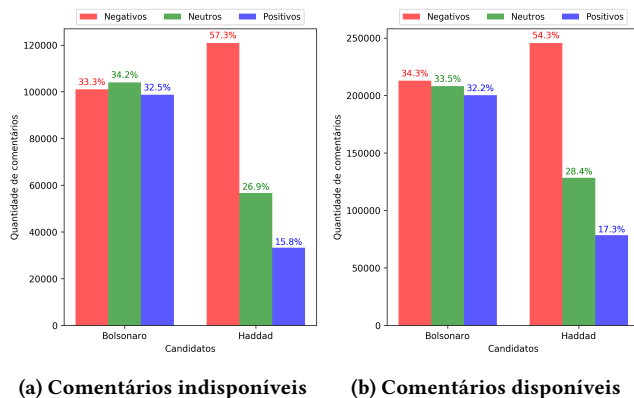


Figura 4: Aceitação no segundo turno desconsiderando likes

Ainda sobre o segundo turno eleitoral, na análise realizada em 2018, quando os *likes* dos comentários não foram desconsiderados, Jair Bolsonaro alcançou 32,2%, 33,3% e 34,5% de rejeição, neutralidade e aceitação, respectivamente, enquanto o Haddad obteve, respectivamente, 51,7%, 29,9% e 18,4%.

Podem ser visualizados graficamente na Figura 5a os resultados para os comentários indisponíveis e na Figura 5b os resultados para os comentários disponíveis. Tratando-se apenas desses resultados em específico, é possível observar que a menor diferença entre os comentários indisponíveis e disponíveis foi de 0,6% para a classe de comentários neutros e relacionados ao Bolsonaro e a maior diferença foi de 2,4%, que ocorreu ao se comparar os comentários negativos referentes ao Fernando Haddad. Adicionalmente, pode-se verificar que, assim como na análise original, o conjunto de comentários disponíveis possui mais *likes* de comentários negativos sobre o Haddad do que *likes* de comentários positivos sobre o Bolsonaro, ao contrário do conjunto de comentários indisponíveis.

Efetuada uma comparação dos resultados originais com os resultados obtidos na presente análise, é possível visualizar que a menor diferença dentre os comentários disponíveis foi de 0,2% para comentários neutros e positivos referentes ao Bolsonaro e a maior diferença foi de 0,7% quando foram comparados os comentários negativos sobre o Haddad. De forma similar, dentre os comentários indisponíveis, a menor diferença entre as classes foi de 0,4% para comentários neutros relacionados ao Jair Bolsonaro e a maior foi de 1,7% também para comentários negativos referentes ao Fernando Haddad. Assim como na primeira análise apresentada na Seção 4.1.1, pode-se perceber que o aspecto do gráfico foi mantido em ambos os conjuntos de comentários e, em geral, a diferença entre as classes de comentários se mostrou consideravelmente atenuada.

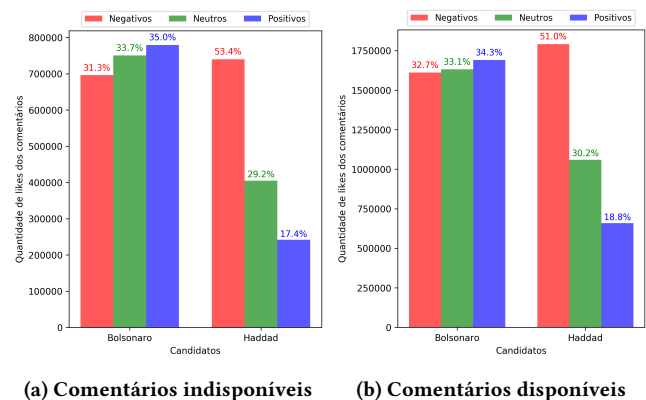


Figura 5: Aceitação no segundo turno considerando likes

4.2 Estatísticas de usuários únicos

Estatísticas de usuários únicos De acordo com a metodologia proposta e a partir das novas análises apresentadas na Seção 4.1 foi possível estabelecer que, devido à baixa diferença dos níveis de aceitação, neutralidade e rejeição em relação aos resultados originais, a indisponibilização de uma fração da base de comentários não foi suficiente para impactar de forma significativa as análises realizadas em [9]. Consequentemente, pode-se considerar que mesmo se todos

os comentários que estão atualmente indisponíveis tiverem sido removidos por motivo de detecção de *bots*, o objetivo de manipulação política por parte desses comentários não foi alcançado.

Contudo, ainda existe a possibilidade de autores dos comentários ainda disponíveis terem compartilhado uma quantidade muito grande de comentários a favor ou contra determinado candidato, o que também caracterizaria a ocorrência de *bots* durante o período eleitoral. Como foi possível obter o identificador que distingue unicamente os usuários da base de comentários após a operação apresentada na Seção 3.4, tornou-se viável abranger, adicionalmente, essa possibilidade.

Para essa análise foram considerados de forma individual todas as três classes de similaridade dos comentários, ou seja, relacionados ao Bolsonaro, relacionados ao Haddad e relacionados a ambos, que corresponde aos comentários que possuem mesma similaridade quanto aos dois candidatos. Vale ressaltar que medidas estatísticas tradicionais [10] de valor mínimo, valor máximo, média, variância, desvio padrão e quantis foram consideradas na análise da quantidade de comentários publicados por cada um dos 1.527.075 usuários únicos identificados.

Tratando-se do conjunto de comentários relacionados com o candidato Jair Messias Bolsonaro, foram identificados 582.590 autores, sendo possível observar que a maior quantidade de comentários publicados por um único usuário foi de 326. Em média, usuários únicos dessa classe de comentários postaram 2,68 comentários, com variância e desvio padrão aproximados de 32,29 e 5,68 comentários, respectivamente. Sendo assim, percebe-se que o valor máximo obtido é consideravelmente maior que a variância, que corresponde à distância aproximada das quantidades quanto à média do conjunto. Entretanto, analisando os quantis do conjunto, 50% dos usuários postaram 1 comentário, 75% postaram até 2 comentários, 95% postaram até 9 comentários e 99% postaram no máximo 26 comentários.

Quanto aos comentários relacionados ao Haddad, foi possível verificar 344.739 autores, que postaram, em média, 2,36 comentários. As medidas de dispersão variância e desvio padrão se mostraram um pouco menores ao se comparar com as que foram obtidas no conjunto de comentários relacionados ao Bolsonaro, sendo 20,73 e 4,55 comentários, respectivamente. Embora a quantidade máxima de comentários por usuário tenha sido de 228, 50% dos autores publicaram 1 comentário, 75% publicaram até 2 comentários, 95% publicaram até 7 comentários e 99% deles publicaram no máximo 21 comentários.

Por fim, o conjunto comentários que obtiveram a mesma similaridade para os dois candidatos englobou 1.117.180 usuários, que compartilharam no máximo 393 e em média 1,97 (aproximadamente) comentários. Dentre todas as classes de similaridade, essa foi a que obteve as menores medidas de dispersão: aproximadamente 12 comentários para variância e 3,46 para desvio padrão. De forma similar às outras análises, também consideramos os quantis do conjunto. 50% dos autores compartilharam 1 comentário, 75% compartilharam até 2 comentários, 95% compartilharam até 6 comentários e 99% até 16 comentários.

Levando em consideração, como análise adicional, o percentual de usuários únicos dentro o total de comentários referentes a cada classe de similaridade, temos, aproximadamente, 37,31% para o conjunto relacionado ao Bolsonaro, 42,32% para o conjunto relacionado ao Haddad e 50,7% para o conjunto de mesma similaridade. Todas as

Tabela 1: Estatísticas tradicionais da quantidade de comentários publicados por usuários únicos

Similaridade	Usuários	MAX	AVG	VAR	STD
Bolsonaro	582.590	326	2,68	32,286	5,682
Haddad	344.739	228	2,363	20,731	4,553
Ambos	1.117.180	393	1,972	11,995	3,463

Tabela 2: Quantis da quantidade de comentários publicados por usuários únicos

Similaridade	50%	75%	95%	99%
Bolsonaro	1	2	9	26
Haddad	1	2	7	21
Ambos	1	2	6	16

informações referentes às estatísticas tradicionais e aos quantis da quantidade de comentários publicados por usuários únicos podem ser conferidos nas Tabelas 1 e 2, respectivamente.

Após a obtenção de todos esses resultados ficou evidente que não houve indícios de manipulação política a partir de *bots* devido à ínfima porcentagem de usuários que publicaram uma quantidade muito superior de comentários do que a média, o que implicou em baixas variâncias e desvios padrão. Visto que o período eleitoral teve início em 16 de agosto de 2018 e se encerrou no dia 28 de outubro do mesmo ano, mesmo que se cada usuário único da base de comentários tivesse publicado 393 comentários, que é o maior valor máximo dos conjuntos, isso corresponderia a, aproximadamente, 5,35 comentários por dia, o que ainda assim seria muito pouco para indicar a ocorrência de *bots*.

4.3 Análise de possíveis eleitores

Como não foram encontrados indícios de manipulação política por meio de *bots* de acordo com as análises realizadas, surgiu a oportunidade de identificar os possíveis eleitores de cada candidato e, conseqüentemente, todos os usuários indecisos a partir da base de comentários disponíveis, a qual foi apresentada na Seção 3.5.

Para que essa análise fosse possível, foi necessário definir, inicialmente, as condições a serem atendidas para que cada usuário único fosse classificado como uma das três classes: possíveis eleitores do Bolsonaro, possíveis eleitores do Haddad e usuários indecisos. Com isso, foi definido que um usuário só poderia ser classificado como possível eleitor de um determinado candidato caso ele tivesse postado ao menos um comentário a favor do candidato ou ao menos um comentário contra o concorrente do candidato e nunca tivesse publicado comentários contra o candidato ou a favor do adversário do candidato. Analogamente, um usuário seria considerado indeciso caso ele não se encaixasse em nenhuma das outras classes.

A fim de classificarmos todos os usuários únicos da base de comentários de acordo com as condições estabelecidas, adotamos a estratégia de dividir, inicialmente, os autores em 9 conjuntos distintos, sendo cada um deles referentes à uma classe de similaridade e à uma classe de sentimento, como por exemplo usuários que postaram comentários positivos sobre o Bolsonaro (<positivo, bolsonaro>).

Após isso, como um usuário só pode ser considerado um possível eleitor caso ele não tenha postado apenas comentários neutros e não tenha publicado somente comentários de mesma similaridade com o Bolsonaro e o Haddad, foram criados conjuntos auxiliares adicionais para cada uma dessas duas possibilidades.

Por fim, a partir dos conjuntos criados, foi possível estabelecer a classificação de cada um dos usuários únicos de acordo com as seguintes considerações:

- Possíveis eleitores do Bolsonaro → usuários contidos nos conjuntos <positivo, bolsonaro>, <neutro, bolsonaro>, <negativo, haddad>, <neutro, haddad>, <negativo, ambos>, <neutro, ambos> ou <positivo, ambos> mas que não estão contidos nos conjuntos <negativo, bolsonaro>, <positivo, haddad>, <somente_neutro> e <somente_ambos>.
- Possíveis eleitores do Haddad → usuários contidos nos conjuntos <positivo, haddad>, <neutro, haddad>, <negativo, bolsonaro>, <neutro, bolsonaro>, <negativo, ambos>, <neutro, ambos> ou <positivo, ambos> mas que não estão contidos nos conjuntos <negativo, haddad>, <positivo, bolsonaro>, <somente_neutro> e <somente_ambos>.
- Usuários indecisos → todos os usuários únicos da base de comentários, exceto os possíveis eleitores dos candidatos.

Após a classificação dos usuários foi possível observar que, do total de usuários únicos da base de comentários, 17,7% deles foram classificados como possíveis eleitores do Bolsonaro, 10,9% como possíveis eleitores do Haddad e, por consequência, 71,4% como usuários indecisos. Os resultados dessa análise podem ser observados de forma gráfica na Figura 6.

É importante mencionar que, apesar da porcentagem de usuários indecisos ter sido muito superior à porcentagem das demais classes, os resultados obtidos são relevantes devido ao alto nível de restrição estabelecido para que os usuários fossem considerados possíveis eleitores. Com isso, além de já ter sido verificado uma vantagem de aceitação do Bolsonaro em comparação com o Haddad em [9], pode ser observado, adicionalmente, que o Bolsonaro também obteve um número maior de possíveis eleitores que o seu adversário de acordo com usuários do YouTube.

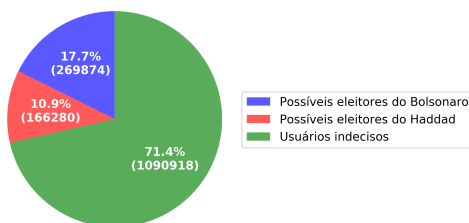


Figura 6: Possíveis eleitores e usuários indecisos

5 CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foram realizadas novas coletas de vídeos e comentários do YouTube referentes ao período eleitoral brasileiro de 2018 que são relacionados aos dois candidatos à presidência que passaram para o segundo turno, Jair Messias Bolsonaro e Fernando Haddad, como forma de verificar quais comentários ainda estão

disponíveis e, como um acréscimo, identificar unicamente seus respectivos autores. Após essa verificação foram realizadas análises com o objetivo de identificar se houve indícios de manipulação política por meio de *bots* e, adicionalmente, avaliar quantos usuários do YouTube podem ser considerados indecisos ou possíveis eleitores dos candidatos.

Vale destacar que, como o objetivo principal do trabalho corresponde à análise do impacto de *bots* do YouTube durante o período eleitoral de 2018, nenhuma alteração foi realizada nas estratégias de detecção de posicionamentos e análise de sentimentos que foram estabelecidas na metodologia de nosso trabalho anterior [9]. Dessa forma, foi possível utilizar os resultados originais de aceitação obtidos em 2018 como base de comparação para a nova análise de aceitação realizada no presente trabalho.

Os resultados obtidos demonstram que mesmo se todos os comentários que estão indisponíveis tiverem sido removidos pelo YouTube apenas por motivo de detecção de atividades caracterizadas por *bots*, os mesmos, quando analisados separadamente, não foram suficientes para impactar diretamente nos níveis de aceitação, neutralidade e rejeição dos candidatos. Tratando-se dos comentários que ainda estão disponíveis, ou seja, que foi possível obter o identificador único de cada um de seus respectivos autores, foi possível observar que, de acordo com as medidas estatísticas de média, variância, desvio padrão e valor máximo quanto a quantidade de vezes que cada usuário único publicou comentários, não houve indícios de atividades de *bots*. Por fim, após definirmos as condições a serem atendidas para que os usuários pudessem ser classificados como possíveis eleitores ou como indecisos, pôde-se identificar que o Bolsonaro obteve 103.504 (6,8%) a mais de possíveis eleitores do que o seu adversário, Fernando Haddad.

Apesar de ter sido mostrado que os comentários removidos não causaram impacto direto na aceitação dos candidatos, é importante mencionar a seguinte limitação do trabalho: como não foi possível obter o identificador único de seus autores (atributo *author_channel_id*), não foi possível verificar se os mesmos estavam compartilhando uma quantidade muito grande de comentários.

Outro ponto que vale a pena destacar é que, apesar de não terem sido encontrados indícios de manipulação política por *bots* nas análises realizadas, isso não significa necessariamente que não houve a ocorrência de bots durante o período eleitoral de 2018 no YouTube. Sendo assim, como forma de comprovarmos ainda mais que os resultados originais de aceitação dos candidatos não foram manipulados politicamente de alguma maneira, pretendemos, como trabalho futuro, considerar adicionalmente a distância de edição do texto dos comentários publicados por cada usuário único.

ACKNOWLEDGMENTS

Este trabalho obteve apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação Arthur Bernardes (FUNARBE).

REFERÊNCIAS

- [1] Ehsan M. Ardehaly and Aron Culotta. 2017. Mining the Demographics of Political Sentiment from Twitter Using Learning from Label Proportions. In *2017 IEEE International Conference on Data Mining (ICDM '17)*.
- [2] Adam Badawy, Emilio Ferrara, and Kristina Lerman. 2018. Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign.

- In *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '18)*.
- [3] Marija Anna Bekafigo and Allan McBride. 2013. Who Tweets About Politics?: Political Participation of Twitter Users During the 2011 Gubernatorial Elections. *Social Science Computer Review* 31, 5 (2013), 625–643.
 - [4] Josemar A. Caetano, Jussara Almeida, and Humberto T. Marques-Neto. 2018. Characterizing Politically Engaged Users' Behavior during the 2016 US Presidential Campaign. In *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '18)*.
 - [5] William Christie, Julio C. S. Reis, Fabrício Benevenuto, Mirella M. Moro, and Virgílio Almeida. 2018. Detecção de Posicionamento em Tweets sobre Política no Contexto Brasileiro. In *Proceedings of the VII Brazilian Workshop on Social Network Analysis and Mining (BraSNAM '18)*.
 - [6] Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. BotOrNot: A System to Evaluate Social Bots. In *Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion)*. 273–274.
 - [7] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The Rise of Social Bots. *Commun. ACM* 59, 7 (2016), 96–104.
 - [8] Debjyoti Paul, Feifei Li, Murali K. Teja, Xin Yu, and Richie Frost. 2017. Compass: Spatio Temporal Sentiment Analysis of US Election What Twitter Says!. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17)*. 1585–1594.
 - [9] Cristian Amaral Silva and Daniel Mendes Barbosa. 2019. Analyzing the Acceptance of the 2018 Brazilian Presidential Election' Main Candidates Based on YouTube Comments. In *Proceedings of the 25th Brazilian Symposium on Multimedia and the Web (WebMedia '19)*. 377–384.
 - [10] Marina Sokolova and Guy Lapalme. 2009. A Systematic Analysis of Performance Measures for Classification Tasks. *Inf. Process. Manage.* (2009), 427–437.
 - [11] V.S. Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. The DARPA Twitter Bot Challenge. *Computer* 49, 6 (2016), 38–46.
 - [12] Mike Thelwall. 2017. The Heart and Soul of the Web? Sentiment Strength Detection in the Social Web with SentiStrength. In *Understanding Complex Systems*. 119–134.
 - [13] Savvas Zannettou, Tristan Caulfield, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. 2018. Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. arXiv:cs.SI/1801.09288