Correlations of Social Media Performance and Electoral Results in Brazilian Presidential Elections

Kellyton dos Santos Brito[1]
Departamento de Computação, Universidade Federal Rural de Pernambuco, Recife, PE, Brazil

Silvio Romero de Lemos Meira
Centro de Informática, Universidade Federal de Pernambuco, Recife, PE, Brazil, silvio@meira.com

Paulo Jorge Leitão Adeodato
Centro de Informática, Universidade Federal de Pernambuco, Recife, PE, Brazil, pjla@cin.ufpe.br

ABSTRACT

The use of social media (SM) in modern political activities has reshaped how politicians run electoral campaigns. This study aims to improve the understanding of online campaigns and their correlation with electoral results. We focus on the 2018 Brazilian presidential campaign, which is well known for its strong online presence, and analyze how candidates used their SM profiles, as well as how citizens interacted with them. We propose a new set of metrics for modeling SM performance and identify statistical correlations between SM performance and votes received. For this, we analyzed more than 40,000 posts made by the 13 candidates on Brazil's three major social networks (Facebook, Twitter, and Instagram) from January to October 2018. Results indicate that candidates used SM heavily throughout the year but focused on engaging words and avoided contentious topics. The most voted-for candidate received more than half (55%) of the interactions received by all the candidates. Posts' interactions were highest on Instagram, where users were increasing the attention given to political content. Lastly, we found strong correlations between the proposed metrics and votes received. Thus, proposed metrics may support new models for predicting electoral results using combined data from many social networks.

Key points for practitioners:

- The most used approach for measuring SM performance based on sentiment analysis on Twitter presents many drawbacks.

- This study presents a new set of SM performance metrics, based on the exposition theory, that can use data from many different SM platforms.

- This study found strong correlations between the defined SM performance metrics and electoral results of the 2018 Brazilian Presidential Election.

- The new set of SM performance metrics may support new models for predicting electoral results using combined data from different SM platforms.

Keywords:

Social Media, Presidential Elections, Facebook, Twitter, Instagram, Brazil

[1] Corresponding author. Departamento de Computação, Universidade Federal Rural de Pernambuco, R. Dom Manuel de Medeiros, 52.171-900, Recife, PE, Brazil
kellyton@kellyton.com.br

1 Introduction

Social media (SM) has played a central role in politics and elections throughout this decade. We have entered a new era mediated by SM in which politicians conduct permanent campaigns without geographic or time constraints, and extra information about them can be obtained not only by the press but directly from their profiles on social networks and through other people sharing and amplifying their voices on SM. In this new scenario, SM is used extensively in campaigns, and an online campaign's success can even decide elections. As a consequence, much academic research has been devoted to this modern political campaign paradigm and its activities (Carlisle & Patton, 2013)(Jungherr, 2016), such as how well Facebook and Twitter users reflect the general voting public (Jungherr, 2016)(Mellon & Prosser, 2017), how the sentiment of conversations is connected to real-world events affecting a political campaign (Park, 2013) and whether it is possible to predict election results based on SM analysis (Brito, Silva Filho, & Adeodato, 2021; Tumasjan, Sprenger, Sandner, & Welpe, 2010).

SM also presents some new challenges: the popularity of fake news, in which false content against candidates is rapidly disseminated (Mustafaraj & Metaxas, 2017); the existence of social bubbles (Flaxman, Goel, & Rao, 2016), a phenomenon by which people are usually presented with content that mainly agrees with their personal convictions and imparts a sensation of majority or unanimity; and the use of automated software known as bots to spread true or fake news, whether supporting allies or defaming opponents (Filer & Fredheim, 2017). To deal with these problems, one approach is to directly reach candidates' SM profiles, which allows citizens to obtain official content instead of fake news, as well as to avoid the effects of a social bubble and massive exposure to bots. Moreover, by using their own SM profiles, candidates may actively engage with supporters of their campaigns, who can share and amplify candidates' voices. The potential of this engagement may be reflected in the number of votes received, as occurred in the 2016 U.S. presidential election when Donald Trump focused his campaign on free media marketing (Francia, 2018).

This scenario also occurred in the 2018 Brazilian presidential election. The candidate with more followers on SM (12.5 million followers on Facebook, Twitter and Instagram one day before elections), and almost no time on TV (8 seconds in public propaganda), ran his campaign almost entirely online and was elected (46.0% of valid votes in the first round), while the candidate with more time on TV (5' 32" in public propaganda) and fewer social network followers (2.2 million) received only the sixth most votes (4.7% of vote share).

Despite many initiatives aiming to study candidates' SM behavior and its correlation with electoral results, many of them are constrained by the technical challenges of collecting SM data, especially from Facebook and Instagram. As a result, as presented in very recent surveys (Brito et al., 2021; Chauhan, Sharma, & Sikka, 2020) most studies are limited to a small timeframe (some days before election day), and to Twitter posts, from which it is easier to collect data. Also, most studies focus on the sentiment analysis of citizens' posts about candidates, forgetting the valuable information that can be gathered from the candidates' networks.

In this context, this study aims to improve the understanding of online campaigning and its correlations with electoral results. It focuses on the first round of the 2018 Brazilian presidential campaign, well known for the strong online presence of the president who was elected, despite his small presence on traditional media and his absence from the debates before the first round. First, we analyzed how candidates used their SM profiles and how citizens interacted with them. Then, we present a new approach to find a correlation between candidates' SM performance and votes received. Instead of the traditional approach of counting the volume and/or sentiment of people talking about a candidate, we elaborate on Zajonc's exposure theory (ZAJONC, 1968) to consider how many people are paying attention to a candidate and amplifying his voice. Thus, we propose a new set of metrics based on attention to and engagement with model candidates' performance on SM. Using the proposed metrics, we identified statistical correlations between SM performance and votes received.

For this, we collected and analyzed data from more than 44,000 posts made by the 13 candidates from January 1, 2018, to October 6, 2018, one day before the first-round election day. Data was collected from the three most used social network sites in Brazil: Facebook, Twitter and Instagram.

The remainder of this paper is organized as follows: Section 2 presents the conceptual background and related works, followed by a brief overview of the 2018 Brazilian elections in Section 3. Section 4 presents the research approach, including the research questions, the reasoning behind and proposal of a new set of metrics for measuring performance on SM, data collection strategy, and analysis methodology. In Section 5, the results are presented and discussed, followed by Section 6, which presents concluding remarks and future work.

2 Background

Contemporary SM systems are new: Facebook launched for public access in 2006, Twitter debuted in 2006, and Instagram emerged in 2010. The use of SM in modern political activities is a new phenomenon that already presents promising results. As the background for this research, we first explore the role of SM in elections and the use of SM by politicians and candidates. Then, research on correlating SM and electoral performance is presented. Finally, the state of the art in this area is discussed.

2.1 The Use of Social Media in Elections

The impact of SM on politics and elections around the world is receiving attention. Smyth (DiGrazia, McKelvey, Bollen, & Rojas, 2013) studied how SM was used in the 2011 elections in West Africa, Nigeria and Liberia, concluding that SM helped to overcome a previous scarcity of information during the electoral process. In a study regarding the 2013 national election in Norway, Kalsnes (Kalsnes, 2016) described "social media interaction deadlock," a phenomenon that is increasing the disparity between the parties' expressed strategies and online performance. Moreover, it was determined that political parties identify three clear disadvantages when communicating with voters online: (a) online reputation risk; (b) negative media attention and (c) limited resources. In the 2014 Indian general elections, Jaidka (Jaidka & Ahmed, 2015) studied official Twitter accounts of the top ten political parties and identified the new paradigms created by political parties to engage and inform voters, driven by modern information and communications technology (ICT).

Concerning U.S. elections, in an analysis of 2012 U.S. presidential candidates' Facebook pages, Bronstein (Bronstein, 2013) showed that in addition to the mobilization of supporters, campaigns used to post information only on a small number of non-controversial subjects, discouraging dissent and encouraging affective allegiances between the candidate and his or her supporters. Regarding the same elections, Mascaro (Mascaro, Agosto, & Goggins, 2016) studied conversational features in Twitter and concluded that, although candidates and media are the most talked about and talked to, these interactions elicited no response.

More recently, Hall (Hall, Tinati, & Jennings, 2018) analyzed the role that SM played in the outcome of the 2016 U.S. presidential election and the Brexit referendum. His conclusions were different from those of previous studies, and he argued that discussions on SM only represent a small portion of the overall discussions in a political campaign and play a minor role in the overall ecosystem. However, in the same year, and regarding the 2016 U.S. presidential election, Morris' results (Morris, 2018) suggested that campaign messages about candidates sent via Twitter—regardless of the candidate of focus—resonated just as strongly with potential voters as those sent via traditional media, reinforcing the power of SM. In one of the very few studies including Instagram, Aminolroya (Aminolroaya & Katanforoush, 2017) highlighted that in 2016, the flow of information from followees to followers in Instagram played a significant role in the Iranian parliamentary election.

After finding the use of SM in campaigns worldwide, researchers naturally started to study the correlations of SM use and electoral performance, presented next.

2.2 The Use of Social Media and Electoral Performance

The correlation between SM performance and electoral performance has been the focus of research since 2010. Tumasjan et al. (Tumasjan et al., 2010) presented a seminal study in the context of the German federal election aiming to predict the results. They collected all tweets that contained the names of the six parties represented in the German parliament or selected prominent politicians of these parties and compared the volume of tweets with the election results. They observed that the relative volumes of tweets closely mirrored the results of the federal election. Then, they claimed that the mere number of tweets mentioning a political party has high correlations with votes, can be considered a plausible reflection of the vote share, and has predictive power even comes close to traditional election polls. In the same year, O'Connor (O'Connor, Balasubramanyan, Routledge, & Smith, 2010) found similar results with a similar approach improved by sentiment detection of tweets.

Kruikemeier (Kruikemeier, 2014) studied content characteristics and candidates' styles of online campaigning during the Dutch national elections of 2010. His findings showed that candidates who used Twitter during the campaign received more votes than those who did not, and that using

Twitter in an interactive way had a positive impact as well. Effing (Effing, Van Hillegersberg, & Huibers, 2011) also studied the impact of SM usage in elections in the Netherlands, showing that during the national elections in 2010, politicians with higher SM engagement received relatively more votes within most political parties. In 2013, DiGrazia et al. (DiGrazia et al., 2013) showed a statistically significant association between tweets that mention a candidate for the U.S. House of Representatives and his or her subsequent electoral performance, indicating that data about political behavior can be extracted from SM. Later, Ramadhan (Ramadhan, Nurhadryani, & Hermadi, 2014) analyzed SM utilization in the 2014 Jakarta legislative election, showing that the usage of SM, especially Facebook and Twitter, is strongly correlated with the number of votes gained by the candidate.

Following these studies, a vast literature was published aiming to correlate SM and election performance, and ultimately trying to predict elections' results. Two very recent papers summarize these initiatives. Chauhan et al. (Chauhan et al., 2020), surveyed 38 papers, and Brito et al. (Brito et al., 2021) performed a systematic review of 83 relevant studies. Both reviews highlight that most studies have used Twitter as a corpus for correlating SM and election results based on volume and sentiment analysis. However, some studies challenge this approach (Gayo-Avello, Metaxas, & Mustafaraj, 2011; Jungherr, Jürgens, & Schoen, 2012; Jungherr, Schoen, Posegga, & Jürgens, 2017), and Brito et al. also found that only 55% of the 64 studies that used this approach obtained success. Both survey studies (Brito et al., 2021; Chauhan et al., 2020) also highlight that the existence of irrelevant, junk, fake, or spam posts can affect the results, as well as the challenges of accurate sentiment analysis on tweets, which are small texts. Moreover, (Brito et al., 2021) highlighted the difficulties in using data from multiple networks besides Twitter, and the positive effect of collecting data for long periods (from 45 to 120 days before elections). Finally, both studies call for metrics and models capable of exploring data from other social networks, useful in more than one electoral context (such as in different countries and continents), and not dependent on an arbitrary choice of words for filtering posts.

2.3 State-of-the-Art Discussion

Based on the conclusions of the presented studies and literature reviews, we can surmise that SM analysis, especially Twitter analysis, already plays an important role in democracies worldwide. Further, politicians and parties have already moved to online candidatures. Indeed, contemporaneous political activity is strongly based on the concept of the "permanent campaign" having a permanent nature, including the execution of campaign-like activities by the political actors during non-election periods. In addition, many studies correlating SM data and election outcomes are also being performed. However, some limitations can be highlighted.

a) Data-gathering barriers: SM platforms have substantial restrictions to access their data through their application programming interfaces (APIs). For instance, Twitter's standard API (Twitter Inc., 2020a) only searches against tweets published in last seven days, and open queries do not guarantee that all tweets are returned. Furthermore, due to the Cambridge Analytica data scandal (Isaak & Hanna, 2018), the process of collecting data from Facebook and Instagram became more difficult and now requires Facebook's explicit consent after it has analyzed the system requesting the information. This barrier leads to other study limitations, presented next.

b) Focus on Twitter data: Most studies focus on Twitter not because it is the more relevant social network, but because it is easier to collect data from their API than from other social networks, such as the Facebook/Instagram API. For instance, it is possible to perform an open search for posts containing a word on Twitter but not on Facebook. As a result, large sets of data and indicators on other social networks are simply being ignored.

c) Temporality: A great deal of research effort is spent only during campaign periods, even though there are currently "permanent campaign" activities. Considering that presidential candidates are usually also members of the parliament or are trying to be reelected, analysis of their online activities over longer periods can lead to better understanding. Moreover, most studies also fail to obtain data from the entire campaign period, restricting data collection to an arbitrary choice between one day and a few weeks prior to elections.

Agreeing with (Kreiss, Lawrence, & McGregor, 2018), who stated that "researchers should refrain from automatically generalizing the results of single-platform studies to 'social media' as a whole," we conclude that studies regarding politicians' behavior on SM covering other networks besides Twitter, as well as studying the correlation of this behavior and election results, is very necessary to better frame and understand this new scenario. Also, new metrics considering these networks should be investigated and proposed. These metrics must be well defined, generalizable, and applicable in several electoral contexts, such as across different countries and years.

3 Brief Overview of the 2018 Brazilian Elections

Thirteen candidates ran for president in a controversial campaign that was polarized by two main candidates: Fernando Haddad and Jair Bolsonaro.

Despite ex-President Lula's imprisonment in April 2018, polls pointed to him as the favorite to win the election in all scenarios. He was officially launched as a candidate, but after the campaigns started, Lula's candidacy was denied by the Superior Electoral Court; he was replaced by former São Paulo mayor Fernando Haddad, who used the slogan "Haddad is Lula." Both Lula and Haddad are from the Workers' Party (PT), which won the last four presidential elections. The party is left-wing oriented.

The second most prominent candidate (elected as president after winning the first and second election rounds) was Jair Bolsonaro. He had been a federal deputy since 1991 and is well known as a veteran and for his non-politically correct opinions and speeches. He moved to a small party (PSL) in 2018 to get support for his candidature. In contrast to Lula/Haddad's campaign, Bolsonaro presented a right-wing proposal. At the beginning of the campaign, he was stabbed in the stomach while interacting with supporters. His condition prevented him from returning to public activities and debates for the remainder of the first round.

Many candidates presented themselves as third options: Ciro Gomes (center-left) and Geraldo Alckmin (center-right) presented themselves as moderate options for left and right-wing voters. João Amoêdo, a right-wing businessman, was the "non-political candidate." Cabo Daciolo (far-right) was often the "comic candidate." Henrique Meirelles (center-right) represented the current government, which was very unpopular because of the impeachment of the last president. Marina Silva (center-left), the third most voted for candidate in 2014, and Alvaro Dias (center-right) completed the list of "third way" candidates. Guilherme Boulos (ultra-left), Vera Lúcia (ultra-left), Eymael (center-right) and João Goulart Filho (center-left) composed a group of candidates with very few supporters.

It is important to note that initially, the two most popular candidates were the Workers' Party candidate (Lula/Fernando Haddad) representing the left and having the second most time on TV (2'23"), and Geraldo Alckmin, representing the right and with the most time on TV (5'32"), similar to previous elections in 2014. Geraldo Alckmin was confident that after the beginning of his campaign on TV he would perform better; this was shown to be incorrect (he came in fourth). Most of the other candidates had little time on TV and had to concentrate their campaigns on the internet. Jair Bolsonaro had only 8 seconds of TV time. As he was also prevented from participating in debates and public events, his campaign was mostly based on social networks—he even published from the hospital moments after his surgery.

Table 1 lists the candidates with the number of votes received and the duration of their official propaganda on TV, showing no relationship between exposure time on TV and votes received.

Table 1 – Number of votes and TV airtime of each candidate

Candidate	Votes	Votes (%)	Time on TV	Time on TV (%)
Jair Bolsonaro	49,276,990	46%	0'08''	1%
Fernando Haddad	31,342,005	29%	2'23''	19%
Ciro Gomes	13,344,366	12%	0'38''	5%
Geraldo Alckmin	5,096,349	5%	5'32''	45%
João Amoêdo	2,679,744	3%	0'05''	1%
Cabo Daciolo	1,348,323	1%	0'08''	1%
Henrique Meirelles	1,288,948	1%	1'55''	16%
Marina Silva	1,069,577	1%	0'21''	3%
Alvaro Dias	859,601	1%	0'40''	5%
Guilherme Boulos	617,122	1%	0'13''	2%
Vera Lúcia	55,762	0%	0'05''	1%
Eymael	41,710	0%	0'08''	1%
João Goulart Filho	30,176	0%	0'05''	1%
Total	107,050,673	100%	12'21'	100%

4 Research Approach

This research studies the relationship between SM and the electoral performance of candidates running in Brazil's 2018 presidential election by focusing on the candidates' use of SM and the impacts of this use. From this objective, we derived the following research questions:

RQ1: How did candidates use social media in 2018?

RQ2: How did citizens interact with the official profiles of candidates during the year and during the campaign?

RQ3: Is there a correlation between social media performance and votes received by candidates?

The approach for this research is based on gathering all data regarding candidates' activities on the most used SM platforms in Brazil (Facebook, Twitter and Instagram) during 2018 from January until the first round of the elections. This data was collected and will be analyzed according to the following.

4.1 Measuring Social Media Performance

As mentioned in Section 2, most studies measure performance on SM as the volume (sometimes considering sentiment) of posts made by ordinary people talking about a candidate (usually on Twitter). Such studies are based on the seminal paper by Tumasjan (Tumasjan et al., 2010), who claimed that "the mere number of tweets mentioning a political party can be considered a plausible reflection of the vote share and its predictive power even comes close to traditional election polls." However, as discussed in Section 2.3, this approach has several drawbacks. This study presents an alternative.

In 1968 (ZAJONC, 1968) and beyond (Murphy & Zajonc, 1993; Zajonc, 1980, 2001) Zajonc studies on human psychology hypothesized that "mere repeated exposure of the individual to a stimulus object enhances his attitude toward it." This effect, also called the familiarity principle, has been demonstrated in many different contexts, such as paintings, sounds, geometric figures, and affective reactions. In agreement with this theory, Swap (Swap, 1977) indicated that "overall, more frequently viewed others were preferred to those less frequently seen." In other words, people tend to have better attitudes toward others whom they are used to seeing. Applying these theories in the electoral context, in 1986 Oppenheimer (Oppenheimer, Stimson, & Waterman, 1986) found a correlation between politicians' exposure and electoral performance, and Mondak (Mondak, 1995) found that "media exposure fuels political discussion."

In a way different from most common hypotheses, we based our performance measurement on Zajonc's mere-exposure theory by analyzing how many people are paying attention to a candidate by interacting with their content and propagating their presence, regardless of whether people are talking about them or not, in the context of SM. For this, we consider two sets of metrics. The first is the number of followers in each social network, and the second is the number of interactions on candidates' posts.

The number of followers of candidates in each social network is a direct measure of how many people subscribed to directly receive candidates' content. It is expected that more subscribers lead to more people receiving content and paying attention to a candidate. Considering an arbitrary time period before elections, we defined as metrics the total number of followers at the beginning and end of the period, the increase in the absolute number of followers over the period, and the increase in the relative number of followers during the period (see Table 2). It is important to note that this metric may fail to express how many people are paying attention to a candidate because not all content for all accounts followed by a person are shown to them: SM algorithms prioritize showing users content with more engagement and more aligned with users' preferences (Lars Backstrom & The Facebook, 2013).

The number of interactions on the candidates' posts consists essentially of the likes, comments, and shares on each post. These actions indicate that the user has seen and paid attention to the content and actively acted. One like may be considered a quick, easy endorsement of the content; a comment demands more cognitive effort and may be positive or negative; and a share replicates the content to the user's own network, thereby actively helping to propagate it. In the case of Facebook, a like has subtypes, such as "Like", "Love", "Haha", "Wow", "Sad" and "Angry". However, in practice there are no distinctions among these interactions, that may be considered as just one. This is because even negative reactions, such as “Sad” and “Angry”, are usually negative regarding the content of the post, for example the reporting of a sad situation, and not a disagreement with whoever posted it.

Table 2 – Follower metrics for performance measurement

Social Network	Metric	Description
Facebook	FBFollowStart	Number of followers on the Facebook at the start of period
	FBFollowEnd	Number of followers on the Facebook at the end of period
	FBIncrease	Absolute increase of followers in the period
	FBIncrease%	Relative increase of followers in the period
Twitter	TTFollowStart	Number of followers on Twitter at the start of period
	TTFollowEnd	Number of followers on Twitter at the end of period
	TTIncrease	Absolute increase of followers in the period
	TTIncrease%	Relative increase of followers in the period
Instagram	IGFollowStart	Number of followers on Instagram at the start of period
	IGFollowEnd	Number of followers on Instagram at the end of period
	IGIncrease	Absolute increase of followers in the period
	IGIncrease%	Relative increase of followers in the period
All Networks	FollowStart	Total number of followers at the start of period
	FollowEnd	Total number of followers at the end of period
	FollowIncrease	Absolute increase of total followers in the period
	FollowIncrease%	Relative increase of total followers in the period

Indeed, all these actions, even negative comments, help to propagate a candidate’s presence online. As social network algorithms prioritize showing the content of users with more engagement (Lars Backstrom & The Facebook, 2013), this creates a snowball effect. As more people interact with a post, so it is shown to more people, leading to more people interacting with it. The end result of the exposure theory is that more engagement and more exposure may be correlated with a better attitude toward a candidate and more votes.

Thus, regardless of the social network, we consider the metrics related to number of likes, comments and shares (or similar items, such as Twitter retweets as synonyms of shares). We consider the absolute numbers in a period and the relative numbers per post. In the specific case of this study, we consider Facebook, Twitter and Instagram, and all defined metrics are presented in Table 3. It is important to note that if other relevant social networks would be created or identified as relevant in other elections, their metrics can also be added by following the same rationale of interactions. For example, considering YouTube, the number of visualizations, likes and comments on a video may be considered.

For our analysis, in order to avoid the selection of small arbitrary periods of time that would bias the results, we analyzed two periods: the campaign and the entire year leading up to the election (a period of 9 months).

Table 3 – Interaction metrics for performance measurement

Social Network	Metric	Description
Facebook	FBPosts	Sum of posts in the period
	FBLikes	Sum of likes in the period
	FBShares	Sum of shares in the period
	FBComments	Sum of comments in the period
	FBLikesPPost	Average of likes per post in the period
	FBSharesPPost	Average of shares per post in the period
	FBCommentsPPost	Average of comments per post in the period
Twitter	TTPosts	Sum of posts in the period
	TTLikes	Sum of likes in the period
	TTRetweets	Sum of retweets in the period
	TTLikesPPost	Average of likes per post in the period
	TTRetweetsPPost	Average of retweets per post in the period
Instagram	IGPosts	Sum of posts in the period
	IGLikes	Sum of likes in the period
	IGComments	Sum of comments in the period
	IGLikesPPost	Average of likes per post in the period
	IGCommentsPPost	Average of comments per post in the period

4.2 Data Collection

Data was collected from the period of January 1, 2018, to October 6, 2018—one day before the election. An information system was developed entirely for this collection and passed the verification process for access to Facebook/Instagram and Twitter APIs according to the official guidelines of each platform (Facebook Inc., 2020)(Twitter Inc., 2020b). The following data was collected:

Followers: The number of followers of all candidates' public accounts on a daily basis;

Posts and Interactions: Posts performed by candidates and their ensuing interactions, which consisted of:

- From Facebook: number of likes (including subcategories such as sad, wow and lol), shares and comments;

- From Twitter: number of likes and retweets;

- From Instagram: number of likes and comments;

Social networks' APIs allow for the gathering of data about past posts. Then, when a candidate was included in the system, all of their posts since January 1 were collected. In addition, considering that these metrics change in real time, the strategy consisted of updating data from the last 200 posts of all candidates every day. Then, the system was able to keep posts updated for 2 months after publishing, on average, without overloading the system or overcoming the APIs' limits.

Data collection faced some limitations. As these networks' APIs do not provide the number of followers for previous days, this information must be gathered on a day-by-day basis. Then, data about some candidates, such as Fernando Haddad, was not gathered from January 1 because they were not yet considered possible candidates. Thus, data started to be collected at least from the beginning of candidates' campaigns. In addition, at the beginning of data collection, the accounts of Cabo Daciolo, Eymael and João Goulart Filho on Instagram were *personal accounts*, and it is only possible to automate data gathering from *business accounts*. Therefore, some of their data about Instagram followers were projected according to Facebook and Twitter variance. Finally, Instagram's official API does not allow data collection of IGTV posts. Thus, data from this kind of post was ignored.

Data presented in this paper may present small differences in presented numbers from a preliminary version of the study presented in (Hidden, 2019). The differences in followers are due to the aforementioned projection of followers of minor candidates. Small differences in recorded interactions are due to adjustments to the time zones for data filtering. In (Hidden, 2019), we considered UTC, but in this paper we considered the Brazilian capital local time. This difference does not impact the results or conclusions of either of the papers.

4.3 Data Analysis

Data analysis aimed to answer the research questions directly. Quantitative and statistical analyses were performed.

For "*RQ1: How did candidates use social media in 2018?",* we performed quantitative analysis regarding the total number of posts by day and by platform, as well as the most used hashtags and words. We also analyzed the number of posts related to contentious topics at that moment in Brazil, such as healthcare, unemployment, education, corruption, public security, and social security.

For "*RQ2: How did citizens interact with the official profiles of candidates during the year and during the campaign?",* the analysis is focused on the variation of candidates' followers in each network and quantitative analysis of citizens' interactions (likes, shares/retweets and comments) regarding the candidates' posts.

For "*RQ3: Is there a correlation between social media performance and votes received by candidates?",* we performed a statistical analysis to correlate data regarding candidates' activities and votes received, as well as data pertaining to citizens' interactions and votes, according to performance metrics already presented. The analysis was performed in two steps: (i) correlation analysis between each metric as defined in Section 4.1 (for example, likes on Facebook versus votes received) in order to find the strength of the relationship between votes and the variables related to SM performance; and (ii) linear regression models were created and tested for a preliminary prediction function.

5 Study Results

This section presents the analysis and discussion of the collected data. First, we provide an overall summary of the results. Then, the findings and answers to the defined research questions are presented and discussed.

5.1 Overall Results

Thirteen candidates ran for the presidency. During the campaign, all of them had accounts on Facebook, Twitter, and Instagram (see Table 4).

Table 4 – Candidates' official accounts

Candidate	Party	Facebook Account	Twitter Account	Instagram Account
Alvaro Dias	PODE	ad.alvarodias	alvarodias_	ad.alvarodias
Cabo Daciolo	Patriota	depudadocabodaciolo	CaboDaciolo	cabodaciolo
Ciro Gomes	PDT	cirogomesoficial	cirogomes	cirogomes
Eymael	PSDC	eymaelOficial	Eymaeloficial	eymael_presidente27
Fernando Haddad	PT	fernandohaddad	Haddad_Fernando	fernandohaddadoficial
Geraldo Alckmin	PSDB	geraldoalckmin	geraldoalckmin	geraldoalckmin_
Guilherme Boulos	PSOL	guilhermeboulos.oficial	GuilhermeBoulos	guilhermeboulos.oficial
Henrique Meirelles	MDB	hmeirellesoficial	meirelles	henriquemeirelles.real
Jair Bolsonaro	PSL	jairmessias.bolsonaro	jairbolsonaro	jairmessiasbolsonaro
João Amoêdo	Novo	JoaoAmoedoNOVO	joaoamoedonovo	joaoamoedonovo
João Goulart Filho	PPL	joaogoulart54	joaogoulart54	joaogoulartfilho_oficial
Marina Silva	REDE	marinasilva.oficial	MarinaSilva	_marinasilva_
Vera Lúcia	PSTU	verapstu	verapstu	vera_pstu

The night before election day, the presidential candidates had, in total, 30.2 million followers of their SM accounts. The candidate with the most followers was Jair Bolsonaro, with a total of 12.5 million, and the candidate with the fewest followers was João Goulart Filho, with 20,500 followers.

In total, the candidates published 44,265 posts, with 12,776 (29%) on Facebook, 23,312 (53%) on Twitter and 8,177 (18%) on Instagram. These posts generated 290 million interactions—143 million on Facebook (49%), 32 million on Twitter (11%) and almost 116 million on Instagram (40%). Detailed results and discussions are presented next.

5.2 RQ1: How did candidates use social media in 2018?

This research question aims to identify how candidates used their SM profiles throughout 2018 with regard to their total number of posts by day and by platform; most used hashtags; most used words; and their posts about contentious topics of the moment in Brazil, such as healthcare, unemployment, education, corruption, public security and social security.

The majority of the posts made by candidates were issued on Twitter (53%), followed by Facebook (29%) and Instagram (18%). It is relevant to note that the main candidates, Jair Bolsonaro and Fernando Haddad, were not the candidates who posted most often. Considering the quantity of days before the election (279), each candidate performed an average of 4.1 daily posts on each platform, as shown in Fig. 1.

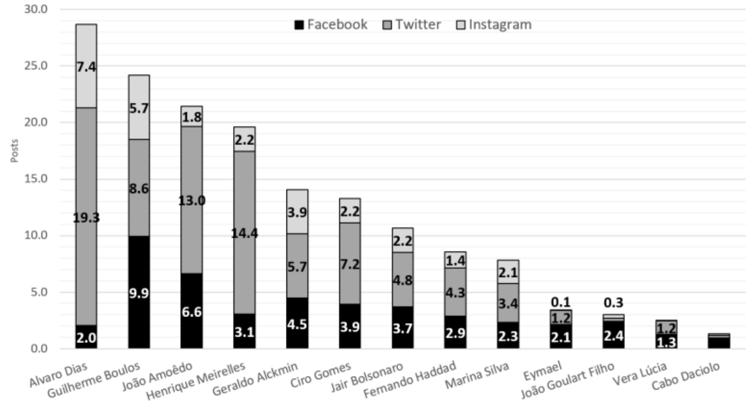

Fig. 1: Number of daily posts made by candidates

Regarding content, the most used hashtags and most used words (excluding stop words) are presented in Table 5. From the list of the 30 most used hashtags, it is clear that institutional hashtags prevailed in one of two categories: (i) a “team” post, indicating that the post was made by a candidate’s team but not the candidate himself (e.g., #ADCOMUNICAÇÃO, which refers to the Alvaro Dias team, or #EQUIPEHM, referring to the Henrique Meirelles team) or (ii) a slogan created for the campaign, such as #VoteSemMedo, meaning “vote without fear,” used by Guilherme Boulos. Only two hashtags do not fall into this category, #AoVivo, which was related to live content, and #Eleicoes2018, a general hashtag referring to elections.

Table 5 – Most-used hashtags

	Hashtag	Facebook	Twitter	Instagram	Total
1	#ADCOMUNICAÇÃO	217	1,786	1,025	**3,028**
2	#EQUIPEHM	17	2,811	15	**2,843**
3	#CHAMAOMEIRELLES	302	2,025	309	**2,636**
4	#ALVARODIAS	120	775	1,052	**1,947**
5	#CIRO12	397	721	312	**1,430**
6	#MEIRELLES15	110	1,205	88	**1,403**
7	#ALVARODIAS19	188	592	489	**1,269**
8	#VOTEMEIRELLES15	187	898	170	**1,255**
9	#BOULOSESONIA	622	39	474	**1,135**
10	#BOULOSESONIA50	459	133	442	**1,034**
11	#PSOL2018	592	29	382	**1,003**
12	#VAMOSSEMMEDO	547	28	399	**974**
13	#PSOL	445	26	377	**848**
14	#EQUIPEJA	28	789	9	**826**
15	#CHAMAOMEIRELLES15	68	681	75	**824**
16	#GERALDO45	316	239	226	**781**
17	#VEMCOMJOÃO30	184	449	61	**694**
18	#EQUIPEGA	376	109	196	**681**
19	#AOVIVO	462	157	5	**624**
20	#PREPARADOPARAOBRASIL	211	137	211	**559**
21	#HADDADPRESIDENTE	235	125	112	**472**
22	#VAMOSRENOVARTUDO	33	391	46	**470**
23	#VOTEMARINA18	127	174	169	**470**
24	#CIROPRESIDENTE	169	168	132	**469**
25	#CHAMAOMEIRELLES💪		423		**423**
26	#CIROSIM	106	197	77	**380**
27	#ELEICOES2018	59	169	119	**347**
28	#GERALDOPRESIDENTE	65	53	226	**344**
29	#VAMOS2018	194	125		**319**
30	#VOTE13	145	80	86	**311**

Fig. 2 presents a word cloud related to the frequency of the 100 most used words, excluding most basic stop words. The most frequently used words were **Brazil** (in Portuguese *Brasil*; 9,459 occurrences) and **country** (*País*; 3,999 occurrences). The analysis of the other 98 most frequently used words indicates some conclusions:

- Candidates used plural engaging words, such as **all of us** (*todos*, 3rd most frequent), **we go** (*vamos*, 4th most frequent), **we** (*nós*, 14th), **our** (*nosso*, 21st), and **together** (*juntos*, 22nd), among others.
- Ex-President Lula, who started serving a 12-year jail sentence during the campaign but was considered a candidate until August 31, was the 32nd most frequent word, with 1,567 occurrences.
- Contentious topics were avoided by the candidates, appearing only after the 50th position, such as **employment** and **jobs** (*trabalho*, 54th and *empregos* 73rd), **education** (*educação*, 55th), **healthcare** (*saúde*, 65th), **economy** (*economia*, 76th), **public security** (*segurança*, 81st) and **corruption** (*corrupção*, 89th).

Going further in the content analysis, we assessed how each candidate posted about the contentious topics: healthcare, employment, education, economy, corruption, public security and social security.

Fig. 2: Word cloud of most-used words

Overall, the number of posts was well balanced among topics, with two exceptions: employment was the most discussed, while social security reform was almost forgotten, as shown in Fig. 3 and Fig. 4. Analysis indicates that the main candidates, Jair Bolsonaro and Fernando Haddad, posted little about contentious topics. Fernando Haddad focused on education and employment, and Bolsonaro on corruption, public security, education, and the economy. In fact, we now know that one of Bolsonaro's first actions at the beginning of his term was social security reform, but this topic was not discussed at all in his SM posts. On the extremes, Alvaro Dias focused mainly on corruption, and candidates associated with economic liberalism (Geraldo Alckmin and Henrique Meirelles) focused on employment and the economy.

This result is compliant with (Bronstein, 2013), which concluded that "campaigns wanted to retain control… by posting information on a small number of non-controversial subjects," and (Ouédraogo, Séré, Rouamba, & Safiatou, 2018), which stated that "campaign debates have been focused on campaign slogans and on the candidates rather than their political programs." In fact, the analysis of hashtags indicates a focus on campaign slogans. Analysis of sensitive words indicates that some less-expressive candidates were more willing to discuss contentious topics, while the main candidates avoided these subjects. The exception was the very sensitive topic of social security, with very few mentions by any candidate.

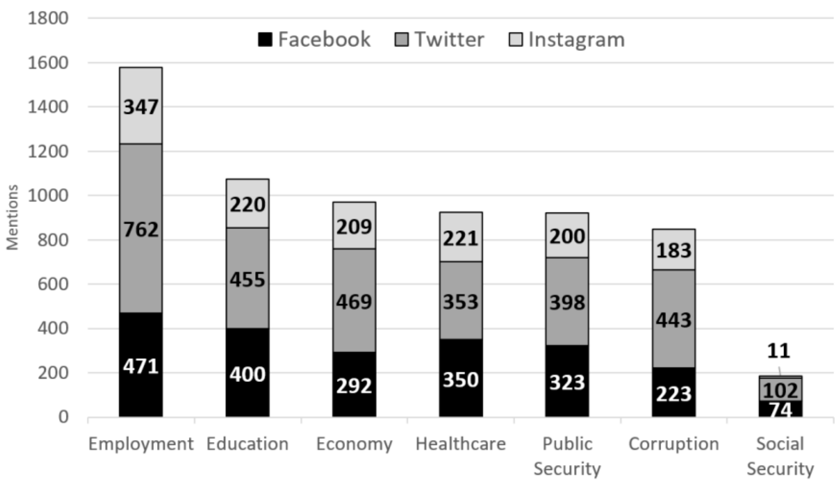

Fig. 3: Total posts about contentious topics

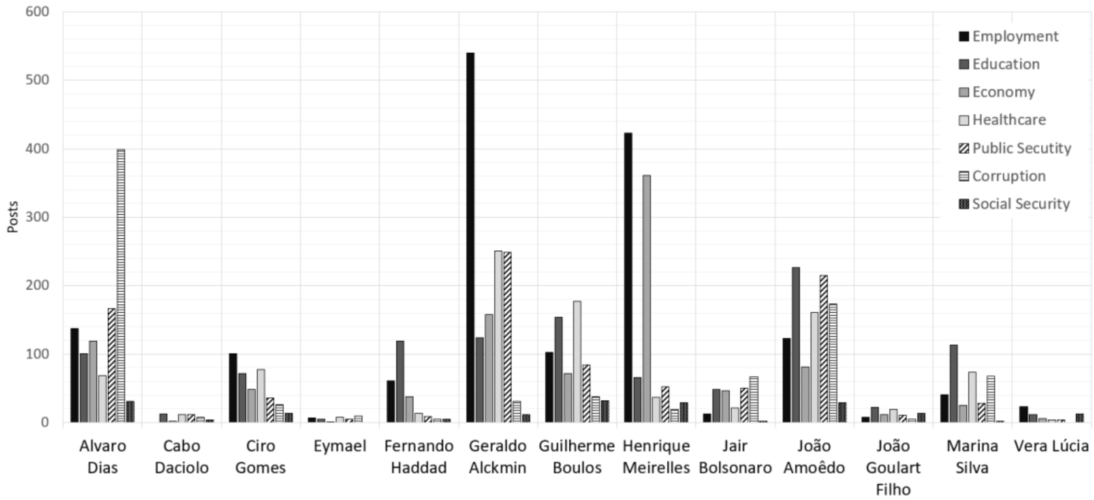

Fig. 4. Posts about contentious topics, by candidate

5.3 RQ2: How did citizens interact with the official profiles of candidates during the year and during the campaign?

This research question aims to identify how Brazilian citizens interacted with candidates' official profiles in two ways: (i) by the number of followers in each social network, and (ii) by the reach of posts issued by each candidate, as a means to measure their impact and ability to mobilize voters. As described in Section 4, we collected the number of candidates' followers from the first day of campaigning, as well as the number of citizens' interactions (likes, shares, comments and related metrics) on posts that the candidates made since January.

The number of followers of candidates' accounts increased from 21 million on the first campaign day to 30 million on the last campaign day, an increase of 44%. Instagram presented the highest rate of increase (147%), followed by Facebook (34%) and Twitter (16%). Ciro Gomes (136%) and João Amoêdo (128%) received the most noticeable general increases in followers in relative numbers, while the president-elect, Jair Bolsonaro, increased his number of followers by 49% and Fernando Haddad by 67%. However, in absolute numbers Jair Bolsonaro was the candidate with the most new followers (4.1 million); Fernando Haddad had only the fourth-greatest total amount of new followers (733,000). João Goulart also presented a high increase in percentage terms (213%), but he started from a very small voter base compared to the main candidates.

The most noticeable network-specific increases occurred on Instagram for João Amoêdo (326%), Fernando Haddad (282%), Ciro Gomes (230%) and Jair Bolsonaro (143%). This data suggests the beginning of a behavioral change in Brazil, with people who use Instagram becoming more interested in political content. At the other end, the candidate with the second most followers in total, Marina Silva, presented an increase rate of only 3%, and Alvaro Dias also presented a small increase rate (4%). Their lesser performances may be explained because both of them already had large bodies of followers at the beginning of campaign (Marina Silva was the 3rd most voted for in 2014 presidential elections, and Alvaro Dias was serving as a senator of the republic in 2018), but their online campaigns did not engage their voters. Table 6 presents the metrics related to the number of followers of each candidate at the beginning and end of the campaign. Table 7 shows metrics related to the changes in numbers of followers.

Candidates' posts generated 290 million interactions, by considering the sum of the number of likes, shares, and comments—143 million on Facebook (49%), 32 million on Twitter (11%), and almost 116 million on Instagram (40%), showing that Facebook was the social network with the most impact. Despite the higher number of posts on Twitter, as indicated in Section 5.2, the performance of all candidates was low on this social network, obtaining a maximum of 20% of interactions. Moreover, two candidates performed better on Instagram: Jair Bolsonaro and Ciro Gomes. All other candidates performed better on Facebook. Nevertheless, the fact that the first and third most voted-for candidates performed better on Instagram supports the previous conclusion that Instagram may be gaining greater relevance in the Brazilian political context.

Table 6 – Followers at the beginning and end of the campaign

Candidate	Begin of Campaign – August 16				End of Campaign – October 6			
	FBFollowStart	TTFollowStart	IGFollowStart	FollowStart	FBFollowEnd	TTFollowEnd	IGFollowEnd	FollowEnd
Alvaro Dias	1,154,175	353,671	34,032	1,541,878	1,189,736	358,484	53,567	1,601,787
Cabo Daciolo	220,595	54,357	161,158	436,110	405,137	99,830	295,977	800,944
Ciro Gomes	318,175	198,944	179,927	697,046	672,327	376,030	594,619	1,642,976
Eymael	13,465	22,036	1,319	36,820	22,716	23,757	2,170	48,643
Fernando Haddad	364,882	617,853	111,124	1,093,859	691,049	711,891	424,418	1,827,358
Geraldo Alckmin	933,559	970,833	119,134	2,023,526	1,106,053	991,959	137,874	2,235,886
Guilherme Boulos	345,237	98,287	140,817	584,341	520,523	163,586	260,821	944,930
Henrique Meirelles	198,235	54,393	21,018	273,646	249,672	65,896	31,328	346,896
Jair Bolsonaro	5,496,048	1,265,397	1,596,822	8,358,267	6,995,358	1,606,036	3,886,599	12,487,993
João Amoêdo	1,399,838	110,658	147,602	1,658,098	2,932,508	220,645	628,153	3,781,306
João Goulart Filho	5,469	555	585	6,609	17,099	1,612	1,764	20,475
Marina Silva	2,331,855	1,877,026	108,419	4,317,300	2,386,091	1,905,002	155,484	4,446,577
Vera Lúcia	16,680	489	632	17,801	23,729	1,815	1,129	26,673
Total	12,798,213	5,624,499	2,622,589	21,045,301	17,211,998	6,526,543	6,473,903	30,212,444

Table 7 – Variations in the number of followers during campaign

Candidate	FBIncrease	FBIncrease%	TTIncrease	TTIncrease%	IGIncrease	IGIncrease%	FollowIncrease	FollowIncrease%
Alvaro Dias	35,561	3%	4,813	1%	19,535	57%	59,909	4%
Cabo Daciolo	184,542	84%	45,473	84%	134,819	84%	364,834	84%
Ciro Gomes	354,152	111%	177,086	89%	414,692	230%	945,930	136%
Eymael	9,251	69%	1,721	8%	851	65%	11,823	32%
Fernando Haddad	326,167	89%	94,038	15%	313,294	282%	733,499	67%
Geraldo Alckmin	172,494	18%	21,126	2%	18,740	16%	212,360	10%
Guilherme Boulos	175,286	51%	65,299	66%	120,004	85%	360,589	62%
Henrique Meirelles	51,437	26%	11,503	21%	10,310	49%	73,250	27%
Jair Bolsonaro	1,499,310	27%	340,639	27%	2,289,777	143%	4,129,726	49%
João Amoêdo	1,532,670	109%	109,987	99%	480,551	326%	2,123,208	128%
João Goulart Filho	11,630	213%	1,057	190%	1,179	202%	11,630	213%
Marina Silva	54,236	2%	27,976	1%	47,065	43%	129,277	3%
Vera Lúcia	7,049	42%	1,326	271%	497	79%	8,872	50%
Total	4,413,785	34%	902,044	16%	3,851,314	147%	9,164,907	44%

The most important finding in this analysis is that the profiles of the most voted-for candidate, Jair Bolsonaro, were responsible for receiving more than half (55%) of the interactions received by all the candidates. Since he had almost no time on official TV propaganda (1% of the time, as shown in Table 1), this data supports the assumption that his online campaign was the main determinant of his election.

Table 8 shows the sum of interactions received by the posts of candidates. All detailed metrics related to interactions throughout the year and during the campaign are shown in Appendix.

Among 290 million interactions on candidates' posts, 193 million (67%) took place during the campaign, as illustrated in Fig. 5 (which shows the total number of interactions in both periods). From this data, we can observe that (i) the impact of candidates Fernando Haddad (92%), Ciro Gomes (87%), and Eymael (75%) mainly occurred during the campaign, while other candidates, such as the winner Jair Bolsonaro (64%), started their campaigns and mobilized their networks beforehand. It is also worth highlighting the results of Alvaro Dias, whose campaign interactions were only 39% of the total. As with the small variation in the number of followers, this finding can be explained by the fact that he was already a senator, and his SM impact while in that office was not very different from his impact during the campaign.

Table 8 – Number and percentages of interactions in each network

Candidate	Facebook	FB%	Twitter	TT%	Instagram	IG%	Total	% of Total
Alvaro Dias	1,846,898	51%	462,750	13%	1,280,158	36%	3,589,806	1%
Cabo Daciolo	3,727,720	79%	298,174	6%	665,078	14%	4,690,972	2%
Ciro Gomes	7,517,984	34%	3,575,586	16%	11,022,358	50%	22,115,928	8%
Eymael	230,851	92%	15,947	6%	3,691	1%	250,489	0%
Fernando Haddad	9,385,394	48%	2,071,855	11%	7,948,762	41%	19,406,011	7%
Geraldo Alckmin	3,740,682	62%	519,785	9%	1,740,811	29%	6,001,278	2%
Guilherme Boulos	6,861,079	41%	3,198,877	19%	6,550,757	39%	16,610,713	6%
Henrique Meirelles	2,539,289	84%	196,407	7%	278,910	9%	3,014,606	1%
Jair Bolsonaro	68,572,956	43%	16,038,521	10%	74,889,520	47%	159,500,997	55%
João Amoêdo	31,863,570	71%	4,121,313	9%	8,591,849	19%	44,576,732	15%
João Goulart Filho	103,488	86%	5,036	4%	11,981	10%	120,505	0%
Marina Silva	6,216,476	63%	993,453	10%	2,612,881	27%	9,822,810	3%
Vera Lúcia	94,434	93%	5,245	5%	1,844	2%	101,523	0%
Total	142,700,821	49%	31,502,949	11%	115,598,600	40%	289,802,370	100%

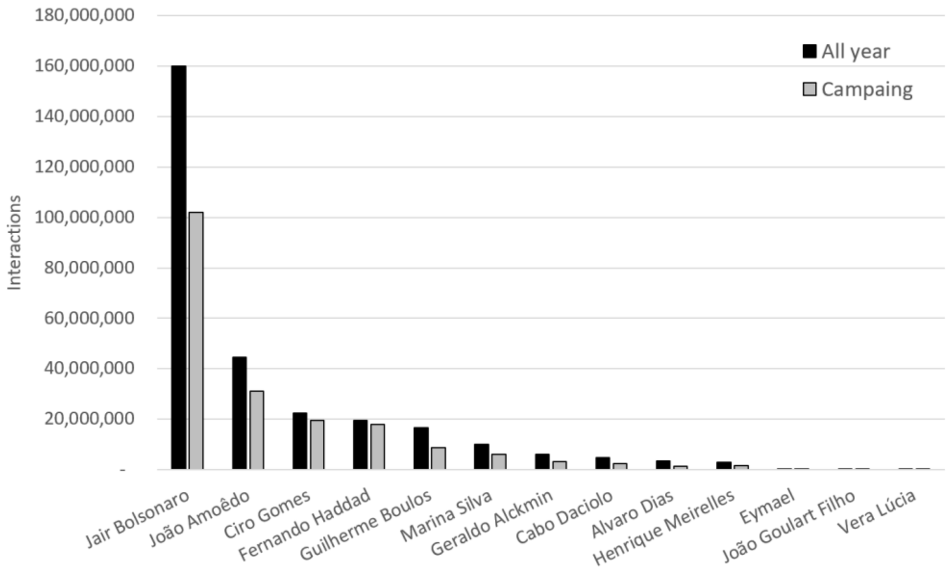

Fig. 5: Interactions only during the campaign and throughout the entire year

Table 9 shows the number of interactions per post in each platform and the average number considering all posts of candidates. Because the number of posts made by candidates is different in each platform, the average value is a weighted average. This data indicates that for the most prominent candidates, the impact of a post on Instagram was higher than any other type of post; that is, although the use of Twitter was higher than any other SM platform, Instagram was more effective. Another unexpected result was the interaction rate of Cabo Daciolo. He was the candidate with the fewest number of posts (as shown in Fig. 1), but he received the second-best interaction rate by post. This outcome can be explained by the fact that Cabo Daciolo was perceived as a "comic candidate." Most of his posts were also humorous, and people often viewed his posts as a pleasant escape from the extremist duality observed in the election period. Many of Daciolo's posts became memes (Taecharungroj & Nueangjamnong, 2015).

A common hypothesis regarding a direct relationship between interactions on SM and received votes cannot be easily observed. Although the most voted-for candidate was also the one with more interactions on SM than others, the candidate with the second most interactions (João Amoêdo) only received the fifth most votes. Also, the second most interacted-with candidate by post (Cabo Daciolo) was the sixth most voted for. This lack of direct correlation, but other possible correlations, is better presented and discussed in the next subsection.

Table 9 – Number of interactions per post

Candidate	Facebook	Twitter	Instagram	Average
Alvaro Dias	3,246	86	619	**448**
Cabo Daciolo	13,705	4,970	17,502	**12,678**
Ciro Gomes	6,828	1,787	18,129	**5,961**
Eymael	386	47	264	**263**
Fernando Haddad	11,702	1,740	20,226	**8,133**
Geraldo Alckmin	3,002	326	1,606	**1,530**
Guilherme Boulos	2,479	1,337	4,110	**2,459**
Henrique Meirelles	2,963	49	463	**551**
Jair Bolsonaro	66,254	11,978	123,580	**53,524**
João Amoêdo	17,177	1,134	17,570	**7,456**
João Goulart Filho	155	62	129	**143**
Marina Silva	9,520	1,039	4,505	**4,487**
Vera Lúcia	268	16	263	**146**

5.4 RQ3: Is there a correlation between social media performance and votes received by candidates?

To find possible correlations between SM performance and electoral results, we performed a correlation analysis by calculating the Pearson correlation coefficient (PCC) of all measures presented in Section 4.1 and the numbers of votes received. In addition to the defined metrics, we also calculated the coefficient considering posts with mentions contentious topics, as discussed in Section 5.2. We considered two periods: campaign only, and the entire year.

Table 10 presents the Pearson correlation coefficient regarding the candidates' behavior and votes. We found no correlation between the absolute number of posts on SM platforms and the number of received votes, as the resulting correlations (r) varied from r = -.11 to r = .09 on these metrics. In fact, the first and second most voted-for candidates were only ranked seventh and eighth in number of posts issued, as already presented. Regarding mentions of contentious topics, in general small negative correlations existed between mentions of these topics and votes, varying from r = -.22 to r = -0.01 for most metrics. Exceptions were mentions of education during the campaign, with a small positive correlation, and mentions of social security, with worse results of r = -.27 and r = .40 for mentions during the campaign and during the year, respectively. As already discussed, social security was a topic avoided by the main candidates because they already knew it was an unpopular topic, but we now know that one of the elected president's first actions was social security reform. This analysis again reinforces existing theories that candidates prefer to focus on campaign slogans and non-controversial subjects (Bronstein, 2013; Ouédraogo et al., 2018).

The correlations between followers' metrics and received votes are presented in Table 11. Despite higher number of followers on Facebook, the highest correlations with votes were related to the absolute increase of followers on Instagram, as well as to the number of followers at the end and beginning of the campaign in this SM platform. Thus, this data once again reinforces the importance of the Instagram platform in Brazilian elections. The absolute increase of followers on Twitter also presented high correlations, but all other metrics regarding Twitter presented small correlations. In addition, the rates of follower increases had no, small, or even negative correlations, varying from r = -.0.25 to r = 0.36. This pattern can be explained by the fact that it is more difficult to increase the percent share of an already large base, as the two most voted-for candidates performed only seventh and fifth in this metric, respectively. On the other hand, for minor candidates, just a few new followers had a greater impact on this metric.

Table 10: Pearson correlations between candidates' behavior and votes

Metric	Correlations	
	Campaign	All Year
Posts on Facebook (FBPosts)	0.09	0.01
Posts on Twitter (TTPosts)	-0.10	-0.11
Posts on Instagram (IGPosts)	0.01	-0.05
Mentions to Employment	-0.14	-0.17
Mentions to Education	0.22	-0.01
Mentions to Economy	-0.08	-0.12
Mentions to Healthcare	-0.13	-0.22
Mentions to Public Secutity	-0.15	-0.14
Mentions to Corruption	-0.11	-0.10
Mentions to Social Security	-0.27	-0.40

Table 11: Pearson correlations between followers' metrics and received votes

Campaign	
Metric	Correlation
FBFollowStart	0.70
FBFollowEnd	0.72
FBIncrease	0.60
FBIncrease%	-0.07
TTFollowStart	0.43
TTFollowEnd	0.53
TTIncrease	0.86
TTIncrease%	-0.25
IGFollowStart	0.84
IGFollowEnd	0.86
IGIncrease	0.87
IGIncrease%	0.36
FollowStart	0.71
FollowEnd	0.78
FollowIncrease	0.80
FollowIncrease%	-0.02

Regarding citizens' interactions, we found strong correlations among all the defined interaction metrics and votes, as shown in Table 12. All metrics presented correlations equal to or higher than r = .75, both during the year as well as considering only the campaign period. During the campaign, metrics related to Instagram and Facebook presented the highest correlations, especially those related to comments on both platforms (1st and 3rd highest correlations) and Instagram likes (2nd). But Twitter's absolute number of shares and likes also presented high correlations (4th and 5th highest correlations). By considering the entire year, the five higher correlations were found for metrics related to Instagram (1st, 3rd and 4th highest correlations) and Facebook (2nd and 5th). One difference in the periods is that by considering only the campaign period, the absolute number of interactions presented higher coefficients. When considering only the campaign, the relative number of likes and comments per post was also important. Also, the metrics related to absolute number of comments on candidates' posts generally yielded higher correlations. The extremes occurred with Facebook metrics: Facebook comments resulted in the highest and second highest correlations in campaign and for the whole year, respectively, but Facebook shares and likes were the metrics with lower correlations in both scenarios.

It is important to highlight that correlations do not mean causality. Although the theory that inspired the definition of the set of performance metrics used in this study suggests a causality relation between exposure and enhancing of attitudes regarding an individual, the objective of this study was to find whether correlations—not causality—existed. Finding correlations among SM metrics and votes does not necessarily mean that SM impacts votes. Offline events, the behavior of candidates in debates, the effectiveness of their propaganda or many other facts may equally impact both electoral results and SM performance, leading to such correlation. In this sense, we think that measuring SM performance may be a quick and easy way to measure public opinion, complementing traditional polling methods.

As a natural consequence of finding these correlations, the next step is trying to predict electoral results based on SM data. As presented in Section 2, there are already many studies trying to correlate SM performance and electoral results (Chauhan et al., 2020), but most of them try to correlate the volume of people talking about a candidate (by measuring the number of posts on Twitter mentioning a candidate) and electoral results. In this study, we tested the defined performance metrics with a linear regression model in a very preliminary approach.

Despite the high correlations, it was not possible to find a linear model to precisely describe election results based on defined metrics. By applying a linear regression algorithm (Yan & Gang Su, 2009) on all metrics combined, the best model produced a mean absolute error (MAE) of 24 million votes (22.5% of total votes). By applying the same linear regression using each metric individually, the best model presented an MAE of 9 million votes (8.8% of total votes). These results are far from acceptable, but such margins were expected due to the characteristics and simplicity of the tested model. The model with all metrics contained more than 50 highly correlated features and only 13 samples. We did not expect good results with this setup. When each variable was considered individually, we had already concluded during data analysis that the high correlations between SM and electoral performance were not linear correlations. Thus, nonlinear approaches should be defined for an adequate modeling of this problem.

Table 12: Pearson correlations between interaction metrics and received votes

Campaign		All year	
Metric	Correlation (r)	Metric	Correlation (r)
FBComments	0.92	IGLikesPPost	0.88
IGLikes	0.88	FBComments	0.88
IGComments	0.87	IGLikes	0.87
TTShares	0.86	IGComments	0.86
TTLikes	0.86	FBCommentsPPost	0.85
FBCommentsPPost	0.85	TTLikes	0.84
IGLikesPPost	0.85	FBLikesPPost	0.84
IGCommentsPPost	0.83	TTShares	0.83
FBSharesPPost	0.83	IGCommentsPPost	0.83
FBLikesPPost	0.82	FBSharesPPost	0.82
TTLikesPPost	0.82	TTLikesPPost	0.80
TTSharesPPost	0.82	TTSharesPPost	0.79
FBShares	0.76	FBLikes	0.77
FBLikes	0.75	FBShares	0.76

Our analysis indicates that, despite high correlations between SM performance and received votes, further studies are needed to create a prediction model based on these metrics, especially nonlinear models. In this sense, our proposed metrics may be used as input data for such future models.

6 Conclusion and Future Works

This paper presented a study on the relationship between SM and the electoral performance of candidates running in the 2018 Brazilian presidential election by analyzing how candidates used their SM profiles and the ways in which citizens interacted with them. We tried a new approach to find a correlation between candidates' SM performance and votes received by using metrics from the three major social networks: Facebook, Twitter and Instagram. For this effort, we collected data about all 44,265 posts from candidates within these networks from January 1, 2018, to October 6, 2018, one day before election day. This study is novel in that it defines a new set of SM performance metrics based on data of the three major social networks, which contrasts with most studies' focus on only Twitter data, and we used a wider data collection period.

Regarding *RQ1* (How did candidates use social media in 2018?) we summarize our findings as: (i) The candidates used SM very heavily, with an average of 4.1 posts every day on each platform, totaling 12.3 posts per day. (ii) The most-used platform was Twitter, receiving 53% of total posts. (iii) The candidates who received the most votes were not the candidates with the most posts. (iv) The main hashtags used were identified as "team posts," or campaign slogans. (iv) The most-used words were engaging words. Contentious topics (e.g., employment, education, and healthcare) were not prominent, and the controversial topic of social security was almost forgotten, especially by the candidates who received the most votes.

Considering *RQ2* (How did citizens interact with the official profiles of candidates during the year and during the campaign?) we conclude that: (i) Instagram users are increasing the attention given to political content on that platform. During the campaign, the rate at which the number of followers increased was 147%, while on Facebook it was 34% and on Twitter only 16%. (ii) Most interactions occurred on Facebook, because there were more posts on this social network than on Instagram. However, considering reactions by post, Instagram was more relevant for all main candidates, even those with fewer followers. (iii) Although Twitter was the most-used network for posting, its impact was very low, with the lowest rate of interactions. (iv) Some candidates' performance was very concentrated during the campaign period, including the candidate who received the second highest number of votes. Others, including the winning candidate, got citizens' attention early. (v) The most voted-for candidate, Jair Bolsonaro, received more than half (55%) of the interactions received by all the candidates.

Finally, for *RQ3* (Is there a correlation between social media performance and votes received by candidates?), we defined a set of metrics to measure SM performance, as well was seven specific measures for posts related to contentious topics. Our statistical analysis found several notable relations: (i) There were no correlations between the number of posts and received votes. (ii) There was a small negative correlation for posts about contentious topics. (iii) There was a strong correlation with respect to candidates' numbers of followers, especially on Instagram. (iii) There were strong correlations with all variables related to interactions with posts, both during the entire year and only on the campaign. (iv) Despite strong correlations, further studies are necessary in order to create a nonlinear model to describe these relations.

This study analyzed an election held before the COVID-19 pandemic. The pandemic led to restrictive measures being adopted worldwide, such as lockdowns and social distancing (Bonaccorsi et al., 2020). Popular rallies and the concentration of supporters have not been allowed in many places. Thus, the online campaign was the main, and in some cases the only, way to campaign in 2020 and 2021, speeding up the adoption of SM by candidates. Consequently, it is expected to see the increase of SM use by candidates in coming years, as well as the correlations between citizens' online behavior and electoral results, and this study is one of the pioneers in this subject.

In terms of future work, we highlight the objective of forecasting electoral results based on SM data. In this sense, the proposed metrics may be used as input data for future models. Also, new datasets may be included, such as campaign pools and demographic data, and specific approaches using data mining and machine learn methods may be promising. Further, as this study can only draw conclusions about one election that occurred in Brazil, future work is suggested to replicate this study with data from other elections around the world, especially the elections that occurred during the COVID-19 pandemic.

APPENDIX – All Collected Metrics and Pearson Correlation with votes

	Metric	Alvaro Dias	Cabo Daciolo	Ciro Gomes	Eymael	Fernando Haddad	Geraldo Alckmin	Guilherme Boulos	Henrique Meirelles	Jair Bolsonaro	João Amoêdo	João Goulart	Marina Silva	Vera Lúcia	Pearson (r)
Mentions to sensitive topics – All year	**Employment**	137	-	101	7	61	540	103	423	13	123	8	41	23	-0.17
	Education	101	13	72	5	119	124	154	66	48	226	22	113	12	-0.01
	Economy	119	2	48	1	38	158	72	361	47	81	12	25	6	-0.12
	Healthcare	69	12	77	8	14	251	177	37	21	161	19	74	4	-0.22
	Public Security	166	12	36	5	9	249	84	52	50	215	11	28	4	-0.14
	Corruption	399	8	26	10	5	31	38	19	67	173	5	68	-	-0.10
	Social Security	31	4	14	-	5	12	32	29	2	29	14	2	13	-0.40
Mentions to sensitive topics – Campaign	**Employment**	72	-	62	1	54	157	66	307	10	41	4	27	10	-0.14
	Education	34	8	42	1	102	38	59	44	20	76	19	57	7	0.22
	Economy	32	1	20	-	24	61	19	142	17	15	10	13	6	-0.08
	Healthcare	29	7	62	-	9	72	44	25	16	41	15	51	2	-0.13
	Public Security	44	7	24	-	4	63	39	37	21	51	10	14	4	-0.15
	Corruption	135	3	13	-	3	15	12	15	22	41	3	35	-	-0.11
	Social Security	9	2	9	-	3	4	11	4	-	4	13	-	-	-0.27
Followers on Facebook	**FBFollowStart**	1,154,175	220,595	318,175	13,465	364,882	933,559	345,237	198,235	5,496,048	1,399,838	5,469	2,331,855	16,680	0.70
	FBFollowEnd	1,189,736	405,137	672,327	22,716	691,049	1,106,053	520,523	249,672	6,995,358	2,932,508	17,099	2,386,091	23,729	0.72
	FBIncrease	35,561	184,542	354,152	9,251	326,167	172,494	175,286	51,437	1,499,310	1,532,670	11,630	54,236	7,049	0.60
	FBIncrease%	3%	84%	111%	69%	89%	18%	51%	26%	27%	109%	213%	2%	42%	-0.07
Followers on Twitter	**TTFollowStart**	353,671	54,357	198,944	22,036	617,853	970,833	98,287	54,393	1,265,397	110,658	555	1,877,026	489	0.43
	TTFollowEnd	358,484	99,830	376,030	23,757	711,891	991,959	163,586	65,896	1,606,036	220,645	1,612	1,905,002	1,815	0.53
	TTIncrease	4,813	45,473	177,086	1,721	94,038	21,126	65,299	11,503	340,639	109,987	1,057	27,976	1,326	0.86
	TTIncrease%	1%	84%	89%	8%	15%	2%	66%	21%	27%	99%	190%	1%	271%	-0.25
Followers on Instagram	**IGFollowStart**	34,032	161,158	179,927	1,319	111,124	119,134	140,817	21,018	1,596,822	147,602	585	108,419	632	0.84
	IGFollowEnd	53,567	295,977	594,619	2,170	424,418	137,874	260,821	31,328	3,886,599	628,153	1,764	155,484	1,129	0.86
	IGIncrease	19,535	134,819	414,692	851	313,294	18,740	120,004	10,310	2,289,777	480,551	1,179	47,065	497	0.87
	IGIncrease%	57%	84%	230%	65%	282%	16%	85%	49%	143%	326%	202%	43%	79%	0.36
Followers on All Networks	**FollowStart**	1,541,878	436,110	697,046	36,820	1,093,859	2,023,526	584,341	273,646	8,358,267	1,658,098	6,609	4,317,300	17,801	0.71
	FollowEnd	1,601,787	800,944	1,642,976	48,643	1,827,358	2,235,886	944,930	346,896	12,487,993	3,781,306	20,475	4,446,577	26,673	0.78
	FollowIncrease	59,909	364,834	945,930	11,823	733,499	212,360	360,589	73,250	4,129,726	2,123,208	11,630	129,277	8,872	0.80
	FollowIncrease%	4%	84%	136%	32%	67%	10%	62%	27%	49%	128%	213%	3%	50%	-0.02
	Votes	859,601	1,348,323	13,344,366	41,710	31,342,005	5,096,349	617,122	1,288,948	49,276,990	2,679,744	30,176	1,069,577	55,762	1.00

	Metric	Alvaro Dias	Cabo Daciolo	Ciro Gomes	Eymael	Fernando Haddad	Geraldo Alckmin	Guilherme Boulos	Henrique Meirelles	Jair Bolsonaro	João Amoêdo	João Goulart	Marina Silva	Vera Lúcia	Pearson (r)
Posts and Interactions – All Year	**FBPosts**	569	272	1,101	598	802	1,246	2,768	857	1,035	1,855	668	653	352	0.01
	FBLikes	1,049,107	1,725,553	5,198,149	116,026	5,520,652	2,130,438	3,987,155	1,997,480	46,158,936	22,386,426	63,996	4,429,417	44,184	0.77
	FBComments	286,925	747,312	715,216	60,385	1,609,825	1,068,070	1,037,414	386,955	7,708,532	1,459,616	10,760	982,931	13,168	0.88
	FBShares	510,866	1,254,855	1,604,619	54,440	2,254,917	542,174	1,836,510	154,854	14,705,488	8,017,528	28,732	804,128	37,082	0.76
	FBLikesPPost	1,844	6,344	4,721	194	6,884	1,710	1,440	2,331	44,598	12,068	96	6,783	126	0.84
	FBCommentsPPost	504	2,747	650	101	2,007	857	375	452	7,448	787	16	1,505	37	0.85
	FBSharesPPost	898	4,613	1,457	91	2,812	435	663	181	14,208	4,322	43	1,231	105	0.82
	TTPosts	5,373	60	2,001	341	1,191	1,592	2,393	4,013	1,339	3,635	81	956	337	-0.11
	TTLikes	352,345	232,376	2,992,142	11,001	1,578,327	428,410	2,434,091	162,987	12,822,854	3,191,205	4,040	828,985	4,023	0.84
	TTShares	110,405	65,798	583,444	4,946	493,528	91,375	764,786	33,420	3,215,667	930,108	996	164,468	1,222	0.83
	TTLikesPPost	66	3,873	1,495	32	1,325	269	1,017	41	9,576	878	50	867	12	0.80
	TTSharesPPost	21	1,097	292	15	414	57	320	8	2,402	256	12	172	4	0.79
	IGPosts	2,068	38	608	14	393	1,084	1,594	603	606	489	93	580	7	-0.05
	IGLikes	1,225,090	614,781	10,676,211	3,015	7,725,549	1,545,329	6,366,526	267,850	72,435,297	8,394,506	11,610	2,485,601	1,738	0.87
	IGComments	55,068	50,297	346,147	676	223,213	195,482	184,231	11,060	2,454,223	197,343	371	127,280	106	0.86
	IGLikesPPost	592	16,178	17,560	215	19,658	1,426	3,994	444	119,530	17,167	125	4,286	248	0.88
	IGCommentsPPost	27	1,324	569	48	568	180	116	18	4,050	404	4	219	15	0.83
Posts and Interactions – Campaign	**FBPosts**	222	113	629	153	642	459	995	343	282	720	323	341	108	0.09
	FBLikes	377,606	745,751	4,100,829	81,575	5,034,744	1,035,887	1,836,192	1,174,641	23,624,026	15,233,935	45,092	2,130,576	22,934	0.75
	FBComments	155,202	342,684	492,015	46,420	1,544,815	690,244	486,109	228,188	5,365,812	983,274	8,763	568,967	5,893	0.92
	FBShares	208,829	392,036	1,273,852	46,233	2,170,493	289,773	681,368	62,601	7,174,620	5,137,060	19,239	462,596	16,906	0.76
	FBLikesPPost	1,701	6,600	6,520	533	7,842	2,257	1,845	3,425	83,773	21,158	140	6,248	212	0.82
	FBCommentsPPost	699	3,033	782	303	2,406	1,504	489	665	19,028	1,366	27	1,669	55	0.85
	FBSharesPPost	941	3,469	2,025	302	3,381	631	685	183	25,442	7,135	60	1,357	157	0.83
	TTPosts	1,426	51	1,145	31	797	709	1,007	3,492	446	1,523	79	543	129	-0.10
	TTLikes	95,446	229,701	2,695,973	6,777	1,234,742	244,223	1,190,647	46,197	8,306,372	1,907,609	4,029	616,890	3,242	0.86
	TTShares	32,582	64,997	530,285	2,628	338,144	56,932	348,720	10,383	2,084,709	527,257	992	128,091	848	0.86
	TTLikesPPost	67	4,504	2,355	219	1,549	344	1,182	13	18,624	1,253	51	1,136	25	0.82
	TTSharesPPost	23	1,274	463	85	424	80	346	3	4,674	346	13	236	7	0.82
	IGPosts	520	32	463	13	323	340	783	293	189	192	58	331	7	0.01
	IGLikes	499,316	576,317	9,871,318	2,905	7,237,936	681,728	4,053,099	128,140	53,219,796	7,085,494	8,293	1,894,246	1,738	0.88
	IGComments	26,760	45,890	314,151	667	212,039	52,352	103,302	7,198	1,955,432	166,875	249	100,508	106	0.87
	IGLikesPPost	960	18,010	21,320	223	22,408	2,005	5,176	437	281,586	36,904	143	5,723	248	0.85
	IGCommentsPPost	51	1,434	679	51	656	154	132	25	10,346	869	4	304	15	0.83
	Votes	859,601	1,348,323	13,344,366	41,710	31,342,005	5,096,349	617,122	1,288,948	49,276,990	2,679,744	30,176	1,069,577	55,762	1.00

REFERENCES

Aminolroaya, Z., & Katanforoush, A. (2017). How Iranian Instagram users act for parliament election campaign A study based on followee network. In *2017 3rd International Conference on Web Research, ICWR 2017*. https://doi.org/10.1109/ICWR.2017.7959297

Brito, K. S., Silva Filho, R. L. C., & Adeodato, P. J. L. (2021). A Systematic Review of Predicting Elections Based on Social Media Data: Research Challenges and Future Directions. *IEEE Transactions on Computational Social Systems*, 1–25. https://doi.org/10.1109/TCSS.2021.3063660

Bronstein, J. (2013). Like me! Analyzing the 2012 presidential candidates' Facebook pages. *Online Information Review*, *37*(2), 173–192. https://doi.org/10.1108/OIR-01-2013-0002

Carlisle, J. E., & Patton, R. C. (2013). Is Social Media Changing How We Understand Political Engagement? An Analysis of Facebook and the 2008 Presidential Election. *Political Research Quarterly*, *66*(4), 883–895. https://doi.org/10.1177/1065912913482758

Chauhan, P., Sharma, N., & Sikka, G. (2020). The emergence of social media data and sentiment analysis in election prediction. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-020-02423-y

DiGrazia, J., McKelvey, K., Bollen, J., & Rojas, F. (2013). More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behavior. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079449

Effing, R., Van Hillegersberg, J., & Huibers, T. (2011). Social media and political participation: Are Facebook, Twitter and YouTube democratizing our political systems? In *Electronic Participation. ePart 2011. Lecture Notes in Computer Science*. https://doi.org/10.1007/978-3-642-23333-3_3

Facebook Inc. (2020). Facebook APP Review. Retrieved October 10, 2020, from https://developers.facebook.com/docs/app-review/

Filer, T., & Fredheim, R. (2017). Popular with the Robots: Accusation and Automation in the Argentine Presidential Elections, 2015. *International Journal of Politics, Culture and Society*, *30*(3), 259–274. https://doi.org/10.1007/s10767-016-9233-7

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*(S1), 298–320. https://doi.org/10.1093/poq/nfw006

Francia, P. L. (2018). Free Media and Twitter in the 2016 Presidential Election: The Unconventional Campaign of Donald Trump. *Social Science Computer Review*, *36*(4), 440–455. https://doi.org/10.1177/0894439317730302

Gayo-Avello, D., Metaxas, P. T., & Mustafaraj, E. (2011). Limits of Electoral Predictions using Social Media Data. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*.

Hall, W., Tinati, R., & Jennings, W. (2018). From brexit to trump: Social media's role in democracy. *Computer*, *51*(1), 18–27. https://doi.org/10.1109/MC.2018.1151005

Hidden. (2019). Hidden to avoid authors identification. In *Hidden Hidden*.

Isaak, J., & Hanna, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, *51*(8), 56–59. https://doi.org/10.1109/MC.2018.3191268

Jaidka, K., & Ahmed, S. (2015). The 2014 Indian general election on Twitter: an analysis of changing political traditions. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*.

Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology and Politics*. https://doi.org/10.1080/19331681.2015.1132401

Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the pirate party won the german election of 2009 or the trouble with predictions: A response to tumasjan, A., sprenger, T. O., sander, P. G., & welpe, I. M. "predicting elections with twitter: What 140 characters reveal about political sentiment." *Social Science Computer Review*. https://doi.org/10.1177/0894439311404119

Jungherr, A., Schoen, H., Posegga, O., & Jürgens, P. (2017). Digital Trace Data in the Study of Public

Opinion. *Social Science Computer Review*, *35*(3), 336–356. https://doi.org/10.1177/0894439316631043

Kalsnes, B. (2016). The Social Media Paradox Explained: Comparing Political Parties' Facebook Strategy Versus Practice. *Social Media and Society*, *2*(2). https://doi.org/10.1177/2056305116644616

Kreiss, D., Lawrence, R. G., & McGregor, S. C. (2018). In Their Own Words: Political Practitioner Accounts of Candidates, Audiences, Affordances, Genres, and Timing in Strategic Social Media Use. *Political Communication*. https://doi.org/10.1080/10584609.2017.1334727

Kruikemeier, S. (2014). How political candidates use Twitter and the impact on votes. *Computers in Human Behavior*, *34*, 131–139. https://doi.org/10.1016/j.chb.2014.01.025

Lars Backstrom, & The Facebook. (2013). News Feed FYI: A Window Into News Feed. Retrieved November 19, 2020, from https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed

Mascaro, C., Agosto, D., & Goggins, S. P. (2016). One-Sided Conversations: The 2012 Presidential Election on Twitter. *17th Annual International Conference on Digital Government Research (Dg.o 2016)*. https://doi.org/10.1145/2912160.2912185

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of british social media users. *Research and Politics*, *4*(3). https://doi.org/10.1177/2053168017720008

Mondak, J. J. (1995). Media exposure and political discussion in U.S. elections. *The Journal of Politics*, *57*(1). https://doi.org/10.2307/2960271

Morris, D. S. (2018). Twitter Versus the Traditional Media: A Survey Experiment Comparing Public Perceptions of Campaign Messages in the 2016 U.S. Presidential Election. *Social Science Computer Review*, *36*(4), 456–468. https://doi.org/10.1177/0894439317721441

Murphy, S. T., & Zajonc, R. B. (1993). Affect, Cognition, and Awareness: Affective Priming With Optimal and Suboptimal Stimulus Exposures. *Journal of Personality and Social Psychology*, *64*(5), 723–729. https://doi.org/10.1037/0022-3514.64.5.723

Mustafaraj, E., & Metaxas, P. T. (2017). The Fake News Spreading Plague: Was it Preventable? In *Proceedings of ACM Web Science Conference* (pp. 235–239).

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *4th International AAAI Conference on Weblogs and Social Media*.

Oppenheimer, B. I., Stimson, J. A., & Waterman, R. W. (1986). Interpreting U. S. Congressional Elections: The Exposure Thesis. *Legislative Studies Quarterly*, *11*(2), 227–247. https://doi.org/10.2307/439877

Ouédraogo, F. T., Séré, A., Rouamba, E., & Safiatou, S. (2018). Analysis of the 2015 Presidential Campaign of Burkina Faso Expressed on Facebook. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. https://doi.org/10.1007/978-3-319-66742-3_4

Park, C. S. (2013). Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. *Computers in Human Behavior*, *29*(4), 1641–1648. https://doi.org/10.1016/j.chb.2013.01.044

Ramadhan, D. A., Nurhadryani, Y., & Hermadi, I. (2014). Campaign 2.0: Analysis of social media utilization in 2014 Jakarta legislative election. In *Proceedings - ICACSIS 2014: 2014 International Conference on Advanced Computer Science and Information Systems*. https://doi.org/10.1109/ICACSIS.2014.7065881

Swap, W. C. (1977). Interpersonal Attraction and Repeated Exposure to Rewarders and Punishers. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/014616727700300219

Taecharungroj, V., & Nueangjamnong, P. (2015). Humour 2.0: Styles and Types of Humour and Virality of Memes on Facebook. *Journal of Creative Communications*, *10*(3), 288–302. https://doi.org/10.1177/0973258615614420

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *Fourth International AAAI Conference on Weblogs and Social Media*.

Twitter Inc. (2020a). Twitter API. Retrieved November 21, 2020, from https://developer.twitter.com/en/docs

Twitter Inc. (2020b). Twitter Apply for Access. Retrieved October 10, 2020, from https://developer.twitter.com/en/apply-for-access

Yan, X., & Gang Su, X. (2009). *Linear regression analysis: Theory and computing*. *Linear Regression Analysis: Theory and Computing*. https://doi.org/10.1142/6986

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, *10*(6), 224–228. https://doi.org/10.1111/1467-8721.00154

ZAJONC, R. B. (1968). Attitudinal Effects Of Mere Exposure. *Journal of Personality and Social Psychology*, *9*(2). https://doi.org/10.1037/h0025848