Do Politicians Talk about Politics? Assessing Online Communication Patterns of Brazilian Politicians

LUCAS SANTOS DE OLIVEIRA, Federal University of Minas Gerais, Brazil
MARCELO S. AMARAL, State University of Southwestern of Bahia, Brazil
JOSÉ ANTÔNIO G. PINHO, Federal University of Bahia, Brazil
PEDRO O.S. VAZ-DE-MELO, Federal University of Minas Gerais, Brazil

Politicians need to decide how to communicate with their voters in order to build their reputations. This problem is especially complicated during important political events such as the elections when politicians must decide whether to confront and share their thoughts about controversial topics or to simply communicate non-political messages. Aware of these communication behaviors, our goal is to analyze how politicians present themselves in the digital environment and how the public reacts to them. We also investigate whether they change their communication and if there is a typical pattern that is chosen by the majority of politicians over time. To address these problems, we collected 751,117 public tweets of 692 Brazilian deputies from October 2013 to October 2015. Furthermore, we propose a methodology for identifying Twitter messages about political issues at a large scale. We use this methodology to characterize the communication behavior of Brazilian congresspeople in a two-year span. We found that Brazilian congresspeople changed their communication behavior as the election approaches and as they were elected or not. Moreover, we showed that while most of the politicians increased the number of non-political messages during elections, the audience tends to favorite and retweet political messages more.

CCS Concepts: • Human-centered computing → Empirical studies in collaborative and social computing; • Information systems → Web searching and information discovery; Social networking sites; • Computing methodologies → Neural networks.

Additional Key Words and Phrases: Brazil, politics, social media, congresspeople, twitter

ACM Reference Format:

1 INTRODUCTION
In Brazil, democracy is facing a difficult time, where a political crisis is in course. It all started in 2013 when Brazil faced the largest and most significant mass protests in a generation [71]. One year after the protests, a new general election took place, when Dilma Rousseff was re-elected president of Brazil in midst of corruption scandals that involved not only the executive power but...
also legislators and companies in a corruption scheme of bribes, kickbacks, and inflated contracts [86]. This troubled political scenario is not a surprise. In 2001, Scott Mainwaring described several factors that possibly contributed to the current crisis [54]. In short, Brazil has a plethora of political parties, three levels of government and an open-list election system for legislators. Consequently, no single party has ever come close to a commanding majority in Congress, so support is bought with cabinet posts and/or cash, and the election process always leads to a relationship between constituents and politicians based on charisma and rhetorical style [59].

This problematic scenario tends to be more prominent in countries where the average voter is poorer [79] and, as expected, during elections [54, 73, 74], characterizing what Samuels [73] calls Candidate-centric Electoral Systems. Under such systems, party cohesion is weaker [7]; politicians switch parties more frequently [35]; congresspeople strive to give local and individual patronage [37]; and politicians are thus more prone to corrupt behavior [9, 66].

In Candidate-centric Electoral Systems, social media play a determinant role because they can shape the image of candidates [5, 53, 82]. One way to identify and quantify how this is done is by analyzing their public communications in this virtual and democratic environment. Although politicians post messages about a wide range of subjects [16, 20, 41], some works group these subjects into only two disjoint sets, political and non-political [10, 63, 64]. Messages of the first group explicitly describe political agendas, opinions, and activities, helping the public to know what to expect from politicians when they occupy public offices. On the other hand, messages of the second group are related to the private life of politicians or are not being directly linked to their activities and ideals as a politician. If we are able to automatically group these messages, we can characterize all politicians that have (and use) social media by the amount of non-political and political messages in their communications. With that, tools could be easily provided for the public to, for instance, identify the political agenda of politicians or identify politicians who are not transparent about their political views and activities.

In this direction, we propose a supervised machine learning method that classifies every message posted by politicians as political or non-political. After classifying all messages, we numerically characterize the politicians by the proportion of political communications they post, which we call the Political Communication Index (PCI). On one extreme, when PCI is 0, politicians only share non-political messages such as “I read in the Veja magazine that President Dilma will stop dyeing her hair...”. On the other extreme, when PCI is 1, politicians only share political messages such as “On yesterday’s election program, I talked about my commitment to education”. With that in hand, we analyze how politicians present themselves in the digital environment and how the public reacts to them. We identify if they change their communication behavior over time, e.g. around elections or when they take over or leave public office. Finally, and more importantly, our methodology allows us to identify which communication behavior evokes, in general, more (and less) engagement with the public on Twitter, both in terms of social media popularity and in terms of votes in elections.

Although many studies have analyzed the content of social media messages posted by politicians [44], the problem of identifying non-political and political messages, in large scale, is not trivial. In short, the proposed solutions are very difficult to be generalized or to keep accurate over a long period of time. Existing efforts ignore the content of the message [1, 13, 32, 47, 89], or focus on manual inspection of small sets of messages [24, 41, 49, 63], or propose aggregate functions (e.g. count, frequency, etc.) based on specific keywords or hashtags (e.g. “abortion”) to quantify how

---

1A short version of this work is described in Lucas S. Oliveira, Pedro O. S. Vaz de Melo, Marcelo Amaral, and José Antônio Pinho. 2018. When Politicians Talk About Politics: Identifying Political Tweets of Brazilian congresspeople. International AAAI Conference on Web and Social Media (2018).
much politicians are dedicating their communications to a specific topic [20, 22, 31, 36, 51, 67, 77]. Moreover, even works that perform a classification task do not characterize the communication of politicians over time and barely consider the nature of these messages, i.e., whether they are communicating about political issues or about non-political and irrelevant topics [10, 21, 64].

To demonstrate the usefulness of our proposed methodology, we analyze a collection of tweets posted by Brazilian politicians from one year before the 2014 Elections to one year after. We collected all tweets posted by all 692 congresspeople who were active on Twitter and worked in the Brazilian parliament from October 2013 to October 2015. The congresspeople were labeled as (1) newcomer (NC), if they were elected in 2014 but were not in Congress in the previous political term; (2) reelected (RE), if they were elected in 2014 and were also in Congress in the previous political cycle; and (3) loser (LS), if they were in Congress in the previous political term and were not able to be elected in 2014. This labeling is useful to separate politicians according to their success in the 2014 Elections and according to their position as a congressperson in the year before and/or after the elections. This work extends the findings of previous research in three-fold contributions, which are described as follows:

- A computational methodology for identifying political and non-political messages that, afterward, numerically characterize the communication behavior. This methodology can be applied to messages in different languages.
- A parsimonious characterization of Brazilian congresspeople communications over time and during the 2014 electoral campaign. We analyze the nature of this communication and how it may be linked to concepts that characterize Brazilian political relations.
- To the best of our knowledge, this is the first work in the literature that shows quantitatively and qualitatively at a large scale that non-political discourse arises and dominates the campaigns during the elections. Surprisingly, in spite of that, we showed that political messages are fair more popular among the public.

The rest of this paper is structured as follows. In Section 2 we describe the related work. In Section 3 we describe the political dataset characteristics. Thereafter, in Section 4 we present a methodology to identify tweets that have political content and in Section 5 we apply this methodology in the data set to characterize the communication behavior of Brazilian congresspeople over time. In Section 6, we summarize and compare our findings with other works in the literature. In addition, we show the limitations of our work. Finally, in Section 7 we conclude our work.

2 RELATED WORK

Political communication characterization. Computational communication science, social media, and big data are remaking and revolutionizing interpersonal communication [6]. In fact, computational approaches have the capacity to gather and process large quantities of information quickly to serve the public good [76]. Through these computational approaches, several studies analyzed and characterize the communication of politicians in online social networks.

The communication network structure among politicians was analyzed by Yoon and Park [89], who revealed that politicians follow each other as a social ritual based on dyadic reciprocity, and mention each other according to how popular they are with the public on Twitter. Conover et al. [10] used a combination of network clustering algorithms and manually annotated data to exhibit, in a politician’s network, a highly segregated partisan structure. Lietz et al. [51] characterized the online conversational practices of political parties in the 2013 German federal elections from several metrics rooted in theoretical constructs from relational sociology. They found that all parties concentrate their communications on a few hashtags during elections and drastically diverge afterward. Furthermore, the political communications of the public were characterized by Rori
Automatic classification of social media messages into categories is usually a necessary task involved in the process of large-scale analysis, being also used to characterize political communication. Paul et. al. [64] proposed a semi-supervised methodology in which they created a dataset of political keywords by training a topic model over a collection of news articles and then selecting the topics related to politics. Thus, they used an embedding word model to enrich the dataset with other similar political words. Finally, they labeled each tweet as political if it contained words from this dataset. Similarly, Conover et. al. [10] created two initial disjoint sets of tweets containing political and non-political hashtags. Thus, using the Jaccard coefficient, they labeled each tweet by assigning it to one of the two classes and then updating the dataset with the new hashtags. In a comparable approach, Gao et. al. [21] created an initial seed slur dataset by scoring each unique word that appears more than 10 times in a dataset of hateful tweets. Thus a tweet was classified as hateful if it contained one of the seed slurs and then the slur dataset was updated with the new slur terms.

Despite the similarity in some aspects to this work (see Section 4), there are constraints that make these approaches very difficult to be generalized or to be accurate over a long period of time. The approach of Paul et. al. [64] requires an external database of news articles for topic model training. Moreover, in the Paul et. al. [64], Conover et. al. [10] and Gao et. al. [21] approaches there is substantial overlap between streams associated with different political hashtags because many tweets contain multiple hashtags. Furthermore, as shown in Section 4, this approach suffers from some restrictions as politicians use the same hashtags for both political and non-political tweets. Additionally, a simple occurrence of any term present in this final list in a tweet is enough to label such a tweet as relevant which could lead to misclassification.

**Political messages content analysis.** With the rise of social media, many researchers got interested in exploring their role in politics. According to Karlsen and Enjolras [45], in an electoral system based on proportional representation, candidates can use social media in election campaigns with two goals: to mobilize the electorate for their parties or to invest in building their reputations. Also, Grimaldi [31] shows that tweets extraction and analysis could be used to predict the ranking of candidates in elections and also determine how their images are spread amongst the public. In fact, Hemphill et. al. [36] showed that U.S. politicians predominantly used Twitter to provide information and to position themselves on issues.

Concerning the role of social media on the general public, there is evidence that social media can create a public sphere that enables discussions and deliberations [56, 72]. Grant et. al. [30] analyzed the utilization of Twitter by Australian politicians and suggested that politicians are attempting to use Twitter for political engagement, though some are more successful in this than others. In the specific case of Latin America, which is our object of study, it was shown that social media is used to engage the electorate in campaigns even after elections [39], to spread misinformation [19] and even to incentive criminality [75]. Moreover, humans tend to homophily [23]². When it comes to politics or culture, homophily can amplify tribal mindsets and produce “echo chambers” that degrade the quality, safety, and diversity of the online discourse. Debois and Blank [13] showed that people who are both not politically interested and who do not use diverse media are more likely to be in an echo chamber.

Regarding the presence of world political leaders in the social media, Barberá and Thomas [2] suggest that leaders from many countries have social media accounts (e.g. Argentina, France,

---

²Tendency to surround ourselves with others who share our perspectives and opinions about the world.
Ukraine, Tunisia, South Africa, Philippines, Japan, etc), even in those countries with limited press freedom, such as Iran, Kyrgyzstan, or Cuba. Additionally, those accounts assume mostly two forms: either a personal account for the head of government, with messages that at least appear to be written by the world leader herself, or an institutional account for the presidency or prime ministry. In fact, the amateur and seemingly more authentic style of U.S. President Trump’s Twitter account points to deprofessionalization and amateurism as a counter-trend in political communication [17].

According to Pain and Chen [62], Trump may portray himself as the lone outsider who can save the country, but he maintains no balance in populism and civility, using rhetorical devices like capital letters associated with incivility frequently in his tweets, retweeting only his supporters while being extremely insulting to detractors. Apparently, this behavior is not unique to the US president and can be observed in other leaders, such as Jair Bolsonaro in Brazil [50].

Similar findings were presented by Gonawela et al. [28] in the analyses of the social media messages from Donald Trump, Narendra Modi, Nigel Farage, and Geert Wilders. They spent significant shares of their communications in making critical comments and creating enemies. Moreover, Pain and Chen [62] indicate that social media like Facebook and Twitter place the focus on the individual politician rather than the political party, thereby expanding the political arena for increased personalized campaigning.

**Political vs. Non-Political messages.** Given the importance of social media messages in politics, many works attempted to identify non-political and political messages in social media. Ganious and Wagner [20] analyzed messages of all candidates of the 2010 election for the US Congress. From counting the presence of keywords in these messages, they found that ≈ 44% of the messages were related to campaign announcements, ≈ 19% were non-political messages, ≈ 18% were attacking to other politicians and ≈ 17% related with policies. A different methodology was employed by Glassman et al. [24], who manually analyzed collections of thousands of tweets posted by U.S. congresspeople. They revealed that politicians rarely provide new insights into government or the legislative process with the goal of revealing their political activities to their constituents. Instead, they use Twitter as a vehicle for self-promotion. From another manual analysis of U.S. congresspeople tweets, Goldbek et al. [25] found that informational posts about themselves in the news articles and their blog posts are the most common, accounting for over half of all posts, followed by posts about places and their daily activities.

Similar communication behavior was seen in other countries as well. Jackson and Lilleker [41] manually analyzed the messages posted by 51 British members of Parliament and showed that they mostly contain details about their personal lives, personal interests and sense of humor, promotion of self, constituency service, or promotion of their own party.

Pal [63] examined the tweets posted by Narendra Modi and verified that his online image is carefully crafted with a range of banal but mostly positive messages. This actually seems the typical behavior of politicians on Twitter, since similar conclusions were reached by several other studies [44, 82], that is, non-political content apparently dominates politicians’ social media messages. However, Marques and Mont’Alverne [55] investigated the Fortaleza’s city councilors’ tweets and found that most messages are related to the promotion of ideas, negative campaign, mobilization, and promoting campaign events. To the best of our knowledge, we are the first to analyze the content of social media messages posted by Brazilian politicians (in such a large and longitudinal scale).

Moreover, to the best of our knowledge, the rules that formally separate political and non-political messages are yet not established in the literature. Bracciale and Martella [5], while investigating political communication styles, divided communications into four dimensions, and one of them, the “Topic” dimension, identifies the main argument of the tweet, which can be either about political issues, policy issues, campaign issues, personal issues, and current affairs. While the first
three describe the political figure of the politician, the last two are clearly about her/his personal figure. Conversely, Pal [63], inspired by [41], defined a "banal" tweet "by its apparent innocuous nature – delivered as a feel-good missive, ritualized response, or casual musing, but weighed by its underlying meaning as part of a larger message of impression management". Inspired by these two definitions, we propose, in Section 4, our own functional definition what can be considered a political tweet.

To conclude, the aforementioned works show the role of social media use during elections and how it can enable debate and interaction between citizens and politicians, promoting democracy. More specific, some works investigate the content of these messages and strive to classify those messages related to state and public administration issues as political, and trivial or banal messages as non-political. Different from the studies described in this section, in the present work we use a supervised machine learning methodology to identify how much of the communications posted by politicians are devoted to political issues, e.g. reforms, and to non-political subjects, e.g. football messages [60]. Moreover, as far as we know, our analyzed data set is the largest in terms of the number of deputies (692) and time span (2 years). We also analyzed the behavior of different groups of politicians before, during, and after elections. Politicians were grouped according to their situation before and after the election, i.e., if they had or not a seat in Congress during these periods.

3 DATA COLLECTION AND PROCESSING

We collected 751, 117 public tweets of 692 Brazilian deputies from October 2013 to October 2015 by means of the Twitter Search and Standard API (Application Programming Interface), available at http://doi.org/10.6084/m9.figshare.7615760. The names of the active congresspeople during this period were retrieved and validated by a researcher in March 2015 using the Chamber of Deputies Open Data website³. The list of the Twitter accounts associated with the congresspeople was collected from the personal profile pages of each congressperson. After this process, each account was manually validated and the collection of messages was performed in November 2015. We prepared the text of the tweets for processing by removing duplicated tweets, punctuation, words with less than 2 characters, Portuguese stop words, URLs, and mentions.

Considering the time span of our analysis, three sets of politicians exist according to their position before and after the 2014 elections: reelected (RE), loser (LS) and newcomer (NC). By separating the politicians into these three groups, we isolate any confounding effect that may arise from being elected in 2014 or not. It is natural to expect that losers will behave differently than newcomers in the two years around Election Day. More important, we will be able to verify whether a communication behavior is more present in the group of successful politicians (NC and RE) or in the unsuccessful ones (LS).

Table 1 summarizes our data set. Note that the reelected form the largest group of deputies, followed by the losers and, finally, by the newcomers. Also, observe that the reelected are, in average, the most active congresspeople in Twitter, with an average of 1,302 messages posted during this period, followed by the losers, with 977 messages, and the newcomers, with 908 messages in average.

4 IDENTIFICATION OF POLITICAL AND NON-POLITICAL TWEETS

Federal deputies are elected by the population of a country and their duty is to propose, discuss and pass laws, which can change even the Constitution. It is also the federal deputies who approve or not the provisional measures proposed by the president and the country’s annual budget. Given their importance, social media can serve as a valuable tool for them to account for the service they

³https://dadosabertos.camara.leg.br
are providing to the country. Thus, in this section, we present a methodology to classify tweets into two categories: political and non-political. Congresspeople that post mostly non-political tweets are failing to account for the citizens who elected them.

Inspired by what Pal [63] defined as a “banal” tweet and the arguments that compose a political communication identified by Braciale and Martella [5], we propose the following definition of a “political tweet”:

**Definition 1.** A **political tweet** is a message posted on Twitter by a politician whose content expresses subjects related to fundamental issues about state, politics, governance and justice. More specifically, such messages have to cover one or more of the following topics: government, state and/or nation; public programs or policies; projects and laws; political campaign; congress or congressperson’s agenda; government taxes or subsidies; court decisions; budget or public expenditure; corruption or crimes against the public administration; actions and positions on civil society movements.

From Definition 1, we modeled the problem of identifying political and non-political tweets as a supervised binary classification problem. Figure 1 shows the methodology, which can be divided into four steps. First, we sample a large set of tweets posted by deputies that are evenly distributed across deputies and across time. For this work, our sample consisted of 2,000 tweets. Second, we manually label the sampled set of tweets according to Definition 1. Third, we use a text embedding technique, namely Word2Vec C-BoW [57], to transform every tweet in a sequence of numerical n-dimensional vectors that represent each word in that tweet. Fourth, we used a Convolutional Neural Network [60], which is a supervised machine learning model, to automatically label the unlabeled tweets as political or non-political. These four decisions were based on a careful empirical evaluation, which is the described next.

**Fig. 1.** Overview of the methodology to identify political and non-political tweets.

### 4.1 Meta Parameters

The process of identifying political tweets using a classification approach involves several methodological decisions. These decisions can be thought of as the meta parameters of the methodology, which affect the speed and quality of the learning process and cannot be estimated from data. They are related to the following challenges: (i) the number and the selection of instances to manually label; (ii) the text embedding method to be used to transform tweets into vectors; (iii) the
selection of the classification method. Table 2 describes all the meta parameters and their possible values. During our experiments, we verified that the meta parameters are independent among themselves, e.g., changing the text embedding technique does not alter the relative performance of the classification methods. Because of that, the configuration used to generate the results is: 2000 labeled tweets; Word2Vec C-BoW with 300 dimensions as an embedding technique, where number of dimensions represents the vector size to which each word of the text is mapped to; and the Convolutional Neural Network (CNN) architecture as the classification method.

<table>
<thead>
<tr>
<th>labeled tweets</th>
<th>period</th>
<th>embedding</th>
<th>embedding size</th>
<th>classification method</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>random</td>
<td>Word2Vec C-BoW</td>
<td>100</td>
<td>CNN</td>
</tr>
<tr>
<td>500</td>
<td>few months</td>
<td>Word2Vec Skip-Gram</td>
<td>300</td>
<td>LSTM</td>
</tr>
<tr>
<td>1000</td>
<td>few deputies</td>
<td>Glove</td>
<td></td>
<td>FastText</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>Word2Vec C-BoW over our dataset</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 4.2 Sampling and labeling tweets

The first challenge is to select the tweets to be manually labeled as *political* and, conversely, *non-political*. Then, we generate classification results for the following number of labeled instances: 100, 500, 1000 and 2000. For all cases, half of the tweets are manually labeled as *political* and half as *non-political*. To do that, we adopt the following process. Given the whole collection of tweets \( \mathcal{U} \), we create a subset \( \mathcal{P} \) of tweets labeled as *political*, such that \( \mathcal{P} \subseteq \mathcal{U} \). For each tweet \( t \in \mathcal{U} \), \( t \) is labeled as *political* and assigned to the subset \( \mathcal{P} \) if and only if there is a political position on \( t \) according to Definition 1. In case a political position is not clearly stated in the tweet, \( t \) is labeled as *non-political*. Note that this process is different from labeling a tweet as speaking well or badly about a subject, in which the two things can happen at the same time in different degrees, or not happen at all. In our case, if this political position exists, regardless of the rest of the content, the tweet is considered *political*.

The initial and most comprehensive sample of tweets to be labeled was randomly selected from the whole collection. In total, we labeled 2,814 tweets, where 1510 were labeled as *political* and 1304 as *non-political*. Then, we filter this sample to create a fully balanced training set that represents the monthly distribution of the original data. To do that, for each label (*political* and *non-political*), we compute the number of tweets that must be sampled each month from this collection so that the final set has size of 1,000 tweets with this label and that their distribution over the months matches the original data. In order to check whether the distributions match, we calculated the KL divergence [10] between the two samples of 1,000 *political* and *non-political* tweets and the original data. The KL divergence is a measure of how different two sample distributions are, where values closer to zero indicate similar distributions. We obtained 0.002 for the *political* sample and 0.00006 for the *non-political*, which indicates that the distribution of tweets per month in these samples is very similar with the original data. Then, the other samples of 100, 500, and 1,000 tweets were randomly sampled from this collection of 2,000 labeled tweets. The other 814 tweets that were left out of this balanced set is used as a hold-out test set of arbitrary distribution.

Figure 2a shows the F1 score for the classification task when the size of the manually labeled data is varied. The F1 is a single score that balances both the concerns of precision and recall in one number. Having high precision means that the majority of tweets which the classifier labeled as *political* are in fact *political*. Having high recall means that from the total number of *political* tweets, the classifier correctly labeled the majority as *political*. Moreover, the F1 score reaches its best value at 1 and worst score at 0.

Observe that, as expected, the model F1 score grows as we increase the training set size. Also observe that, despite the F1 scores from the results stabilize with a training set of 500 instances,
it grows significantly up to 2000 instances for the hold-out test set, with a F1 score of 99% in the training set and 91% in the hold-out test set.

![Graphs showing classification results](image)

**Fig. 2.** Classification results. F1 scores for different configurations.

### 4.3 Dispersion of labeled tweets

It is also worth mentioning the importance of having an unbiased training set in terms of time. To show that, we compared the performance of the classifier when three different training sets are used: (i) the previous randomly and unbiased collection of 500 manually labeled tweets, (ii) a biased collection of 500 labeled tweets in the time dimension and (iii) a biased collection of 500 labeled tweets in the deputy dimension. In these two biased collections, we artificially made the frequency of tweets more skewed towards a few months and, for the second case, a few deputies. In such cases, more than half of the labeled tweets are from only a few months, for the first biased collection, and from a few deputies, for the second biased collection.

In Figure 2b we show the F1 score for these three collections of training sets. Observe that for the biased collections, the F1 score grows, revealing overfitting that happens when the trained model captures the noise along with the pattern in data and loses generalization capacity. On the other hand, the F1 score in the hold-out test set for the unbiased collection decay, what is expected.

### 4.4 Text embedding technique

Before running the classification methods, we execute a text-embedding technique to transform every word into a numerical vector. We compare four text embedding techniques. The first three word-vectors are publicly available and were trained over a large Portuguese data set [34], which is able to produce an embedding matrix for a vocabulary of 1.3 trillion terms. These vectors were produced using the following methods: \textit{Word2Vec C-BoW} [57], \textit{Word2Vec Skip-Gram} [57] and \textit{Glove} [65]. Additionally, we trained the \textit{Word2Vec C-BoW} model using the collection of tweets of deputies described in this work. We also used the pre-trained \textit{Word2Vec C-BoW} [34] weights and vocabulary to train a new model to recognize the hashtags of our dataset.

Thereafter, we evaluated the different embedding techniques using the parameters described in Table 2. Figure 2c exhibits that \textit{Word2Vec C-BoW} and \textit{Glove} have the same 99% of F1 score in the training set. On the other hand, the result in the test set shows that \textit{Word2Vec C-BoW} achieved a higher F1 score than \textit{Glove}.

Also, it is important to note that the \textit{Word2Vec} model trained using our dataset and the other using hashtags obtained the worst results. The main reason is that in the first case, our corpus of tweets is not as large as \textit{Word2Vec C-BoW} from [34]. Moreover, despite hashtags are good predictors of political tweets [4, 10, 67], in this work, they made the classifier’s performance worse. This is because some politicians often use the same hashtags for political and non-political tweets according to Definition 1. This is the case for the following tweets:
“We continued walking around Rio Grande, spreading the ideas of the well-prepared pre-candidate for governor in order to make our state strong again. #heinze #luiscarlosheinze #oriograndefortoutravez #oriograndetemjeito.”

“On July 25 we celebrate the Settler and Driver’s Day. Congratulations to all the settlers and drivers! #oheinzefaz #oriograndefortoutravez #luiscarlosheinze #oriograndetemjeito.”

### 4.5 Classification method

The last decision is to choose which Neural Network architecture to use. More specifically, we evaluated three different architectures: Convolutional Neural Networks (CNN) [46], Long Short Term Memory Networks (LSTM) [38] and FastText [43]. The evaluation was done through a 10-Fold Cross-Validation, where the training dataset is divided into 10 disjointed sets of approximately equal size. Each set is selected in turn as the testing data, whereas the remaining sets are used as the training data, after that, we calculated the F1 scores. In addition, we also validated the result in the hold-out test set using the same F1 score.

For comparison purposes, we standardize the neural network input layer and an output layer. In the input layer, each word in a congressperson tweet is represented as a dense numerical vector with 300 dimensions learned by Word2Vec C-BoW. In case the word is not present in the vocabulary, we replaced it by a special symbol UNK (unknown) and get its embedding representation. Thus, we have a matrix of words and embeddings with vocabulary size \( \times 300 \) dimensions that we provide as the embedded input layer. Additionally, for the output layer, we use a single neuron with a sigmoid activation function, which outputs a continuous range of values between 0 and 1.

Figure 2d shows the performance of the different Neural Networks. Observe that CNN has the highest F1 score in both data sets, achieving 99% in the training set and 97% in the test set, followed by LSTM with 98% in the training set and 95% in the test set. Finally, the FastText neural network achieved an 86% F1 score in the training set and 95% in the test set.

### 4.6 CNN Architecture

Given that CNN had the best performance among tested Neural Network approaches, in Figure 3 we show the architecture of the Convolutional Neural Network for political and non-political tweet classification, which is similar to [46] architecture, however, with the following adjustments and improvements in the network parameters. In the input layer, we represented each word in a congressperson tweet as a dense vector retrieved from Word2Vec C-BoW [57] with 300 vector positions. Subsequently, there is a 25% rate dropout regularization layer, for reducing overfitting in the neural networks by preventing complex co-adaptations on training data, connected to a convolutional layer with 120 different filters and sizes (3,4,5), activated by a ReLU function, which is a piecewise linear function that outputs the input directly if is positive, otherwise, it will output zero. Thus, the output of the previous layer is connected to a global max-polling layer, which is sample-based discretization process that down-samples an input representation, reducing its dimensionality. Additionally, the previous output is fully connected to a ReLU activation and to another 25% rate dropout layer. Finally, the last dense layer is a single neuron with a sigmoid activation function that outputs 1 if the tweet is political and 0 if non-political. Moreover, we optimized the neural network by means of cross entropy loss function using RMSProp [14] optimization algorithm.

It is also important to describe the reasons why we decided to use a neural network approach. For over a decade, core natural language processing (NLP) techniques were dominated by linear modeling approaches to supervised learning, trained over very high dimensional yet very sparse feature vectors [27, 33]. Such vectors, also called bag-of-words or bag-of-n-grams [33], are attractive due to its simplicity, efficiency, and often surprising accuracy. In this direction, recent work in
learning dense vector representations of words [57] using neural networks [3, 11, 27, 58, 65, 84] were proposed. In all these studies, including Word2Vec [57], the one we use in this work, the main idea is that each word is represented by a vector representing the context in which the word is usually used, being constructed from co-occurrences of words in a given text training data.

In fact, representing features as dense vectors is an integral part of the neural network framework [26], whose resurgence greatly impacted in text classification tasks [88]. In particular, Convolutional Neural Networks [46] are specialized architectures that excel at extracting local patterns in the data. They are fed arbitrarily sized inputs and can extract meaningful local patterns that are sensitive to word order, regardless of where they appear in the input. Despite little tuning of hyper-parameters, a simple CNN with one layer of convolution proved to perform remarkably well in text classification tasks [46].

4.7 Evaluation and Validation

Before running the classifier, we labeled 2,000 tweets evenly distributed across time and congresspeople and further validated by other three independent researchers. In this set of tweets, 1,000 tweets were labeled as political and 1,000 as non-political according to the Definition 1 of political tweets.

The idea of evenly distributing the labeled set across deputies and across time is to make the classifier able to accurately classify political tweets independently of the deputy who posted it and of the time it was posted. Therefore, we evaluated quantitatively the F1 score of our method by grouping the training set by month and by deputy. After that, we classified the tweets in each group using our method and calculated the F1 score. Figure 4 shows the box-plot of F1 scores per deputy and per month (outliers were not removed). Observe that the median in deputy distribution is 0.95 and the minimum is 0.84. Even outliers obtained a good performance (greater than a random classifier). Similarly, the result per month also got a very good performance in general, with a median of 0.97, a minimum of 0.83 and outliers getting better results than a random classifier as well.

In order to qualitatively validate the manually labeled set of 2,000 tweets and the classifier, we created two validation sets: labeled validation set and classified validation set. The first contains a sample of 200 randomly selected tweets (100 labeled as political and 100 as non-political) from the initially labeled set. The second contains 200 randomly selected tweets (100 political and 100 non-political) that were labeled by the classifier and not labeled previously. Then, we asked six
independent researchers to manually classify each tweet into political and non-political categories according to Definition 1. To avoid bias, three researchers labeled the labeled validation set and the other three labeled the classified validation set. We evaluate the agreement among our researchers, the classifier and the six independent researchers using the agreement percentage and the Cohen’s Kappa coefficient ($\kappa$) [48, 61, 75], which measures the agreement between two raters. If the raters are in complete agreement then $\kappa = 1$. If there is no agreement among the raters then $\kappa = 0$, otherwise values vary between 0 and 1. According to Landis and Koch [48], when $\kappa$ is in the range between 0 and 0.2, the agreement is considered poor; between 0.21 and 0.4 is considered fair; between 0.41 and 0.6 is considered moderate; between 0.61 and 0.8 is considered substantial; and finally between 0.81 and 1 is considered almost perfect.

Table 3 shows the result of the agreement over the labeled validation set. Observe that the three researchers have almost the same level of agreement as ours, 85% for Researchers 1 and 2 with $\kappa = 0.70$, 84% agreement with $\kappa = 0.68$ for Researcher 3 and 86% agreement with $\kappa = 0.73$ for majority vote. According to [48], these Kappa scores fall into the range of scores referred to as “substantial” agreement, which validates the labeled set used to train our classifier. Moreover, the kappa scores between the Researchers are 0.78 (R1 and R2), 0.73 (R1 and R3) and 0.80 (R2 and R3). These values are not much higher than those found between the researchers and our classifier, and all values are still in the category of “substantial” agreement [48]. These values show that even among humans there are divergences, which clearly shows that the classification of political tweets is not a trivial task.

Table 3. Cohen’s Kappa and Agreement percentage among the researchers who labeled our set of 2,000 tweets used to train the classifier and three independent researchers over the labeled validation set.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa</td>
<td>0.70</td>
<td>0.70</td>
<td>0.68</td>
<td>0.73</td>
<td>0.78</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>Agreement</td>
<td>85%</td>
<td>85%</td>
<td>84%</td>
<td>86%</td>
<td>90%</td>
<td>88%</td>
<td>91%</td>
</tr>
</tbody>
</table>

The agreement results between the three researchers and the classifier in the classified validation set are described in Table 4. While Researcher 4 agrees with the classifier in 77% of the tweets, with $\kappa = 0.55$, Researcher 5 and 6 agreed with the classifier in 74% of tweets, with $\kappa = 0.49$ and the majority vote has a slightly high agreement with the classifier, with 77% of agreement and $\kappa = 0.54$. 

Fig. 4. F1 score per deputy and per month.
Observe that the majority vote has the same $\kappa$ and agreement percentage as Researcher 4 and higher values than Researchers 5 and 6. Moreover, the kappa scores between the Researchers are 0.60 (R4 and R5), 0.64 (R5 and R6) and 0.59 (R4 and R6). These values are not much higher than those found between the researchers and the classifier, and two of these values are still in the category of “moderate agreement” [48]. Moreover, in recent work and under a similar context, Resende et al. [69] asked researchers to label WhatsApp messages as political or not, and they obtained a kappa of 0.42 between labelers, lower than the one we obtained. Again, this clearly shows that the classification of political tweets is not a trivial task, even for humans. We exemplify this difficulty by showing as follows some examples of tweets in which the researchers and the classifier diverged. We translated them from Portuguese for better understanding.

The following tweet was (incorrectly) labeled as non-political by two researchers and as political by one researcher and by the classifier: “RESTART OF BR 101 WORKS AND DUPLICATION OF AIRPORT TRACK ARE RELEASED BY THE PRESIDENT.” Next, for the following tweet, two researchers (correctly) labeled it as political, and one researcher and the classifier (incorrectly) labeled it as non-political: “Itaperuçu is one of the poorest cities in the metropolitan region, we need to give new groups the opportunity to administer it.” For the following tweet, one researcher labeled it as political, and two researchers and the classifier labeled it as non-political: “Femininity is neither modern nor old! It is simply a woman’s issue. Help share this idea.” This is, in fact, a difficult tweet to label. Note that even among humans there is no consensus. Finally, for the following tweet, all three researchers correctly labeled it as non-political, but the classifier labeled it as political: “on the birthday of district deputy candidate Marcias Sidarta.” Probably the classifier took into account the amount of words related to politics present in this tweet and could not understand the context of these words, i.e., this was just a tweet about a politician’s birthday.

Table 4. Cohen’s Kappa and Agreement percentage among our classifier and three independent researchers over the classifier validation set.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa coefficient $\kappa$</td>
<td>0.55</td>
<td>0.49</td>
<td>0.49</td>
<td>0.55</td>
<td>0.60</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>Agreement %</td>
<td>77%</td>
<td>74%</td>
<td>74%</td>
<td>77%</td>
<td>80%</td>
<td>82%</td>
<td>79%</td>
</tr>
</tbody>
</table>

To illustrate the performance of our classifier, we used the same methodology employed by [30] to summarize the content of different sets of tweets. In Figure 5, we show the word clouds for the sets of political and non-political tweets during the elections period. In Figure 5a we can identify the most popular words related to politics, which include “debate”, “federal”, “support”, “government” and “policy”. In contrast, Figure 5b exhibits words related to a wide range of topics, but some are clearly linked to a discourse that seeks to reinforce the image of the politician, their individual actions and their relationship with their voters, such as “God”, “I published”\(^4\), “Facebook”, “together”, “congratulations”, “family” and “thank you”. Also note that some terms such as “Brazil”, “campaign” and “dilma” appear in both clouds, as they are used in political and non-political messages.

5 RESULTS

From the classification method presented in the previous section, we can characterize each politician according to the number of political and non-political messages they posted over time. After the classification of all messages from our dataset we obtained that reelected have around 51% of their posts classified as political, followed by losers with 48% and newcomers with 43%. This suggests that having a position in the Chamber of Deputies of Brazil apparently increases the participation in

\(^4\)In Portuguese, “I published” can be written using a single word: “publiquei”.

Twitter and the propensity to post more about political issues. From now on, we call the decision to post more or less political tweets as her/his communication behavior. In the following sections, we analyze and describe the communication behavior of Brazilian congresspeople.

5.1 Quantifying political communication

To quantify and characterize the communication behavior of a politician, we propose the Political Communication Index (PCI), a simple ratio between their number of political tweets and their total number of tweets. For a set of tweets $T_i$ of a given politician $i$, this set can be divided into two disjoint sets: $P_i$, containing her/his political tweets, and $NP_i$, containing her/his non-political tweets. With that, we define the PCI of politician $i$ as:

$$PCI_i = \frac{|P_i|}{|NP_i| + |P_i| + 1}.$$  

In other words, if some congresspeople posted the same amount of political and non-political tweets for a given set of tweets, then the PCI is equal to 0.5. If they posted political messages only, then the PCI approaches 1. Conversely, if they posted only non-political messages, then the PCI is 0. Otherwise, the PCI varies between $[0, 1]$. We add 1 in the denominator for cases where the deputy has not posted any message in the given period.

Figure 6 shows the scatter plot of the PCI and the number of tweets for all politicians of our dataset, considering their whole set of tweets, i.e., 2 years of data. Note that is not possible to differentiate the congresspeople by their label. Moreover, most deputies have a significant number of tweets in this period, with PCI varying between $0.1$ and $0.8$, which indicates diversity in communication behavior. It is interesting that for most deputies 0.8 is an apparent upper bound for the PCI, i.e., even politicians who decide to disclose their political views very often (high PCI) have a 20% share of non-political messages. Particularly, note that congresspeople who post more about politics are the majority ($57\%$).

Additionally, there are also divergent behaviors, congresspeople with high PCI and a low number of tweets and congresspeople with low PCI and a high number of tweets. However, on average, politicians that post few or many tweets seem to have no significant difference in their PCI. This behavior can be seen in Figure 7, which shows a box-plot of Political Communication Index for each quartile of the distribution of the number of tweets. Observe that all box-plots are quite similar and concentrate most PCI values between 40% and 70% with the median around 50%. While the works of [20, 29] also pointed for at least 20% of non-political messages, to the best of our knowledge

Adding 1 to the PCI denominator equation makes it impossible that in the extreme case, where a politician posts only political tweets, the division of the number of political tweets by the total would be equals to 1.
no previous work showed that this fraction is surprisingly invariant with respect to how active a politician is.

5.2 Consistent use of Twitter
As depicted in Figure 6, a significant portion of deputies have a small number of tweets over the two years of our analysis. In order to make solid considerations about how deputies behave on Twitter, we must exclude from analyzes those deputies who do not use Twitter consistently. But what is the quantity $n_t$ that defines a deputy who does and does not use Twitter consistently? To answer this question, we propose a simple methodology, which is described as follows.

In a conservative way, we consider all deputies who posted a quantity $n_t = 87$ tweets or less in the two years of our analysis as politicians who do not use Twitter consistently. In other words, a deputy is active on Twitter if he has posted several messages that could be sampled from a Poisson
distribution with $\lambda = 1$ tweet per week or $\lambda = 104$ tweets per two years. Using a significance level of 0.05 and a one-tailed hypothesis test, all deputies who tweeted less than 88 tweets in two years have less than 5% of probability to behave like a deputy who tweets, in average, 1 tweet per week. Thus, we consider all deputies who posted $n_t = 87$ tweets or less in the two years of our analysis as politicians who do not use Twitter consistently.

The solid vertical red line in Figure 6 denotes this threshold $n_t$. Observe that 30% of the deputies have not posted messages consistently in Twitter over the two years of our analysis. For the remainder of this paper, we call such deputies as inactive. The other 70% of deputies, which represents the majority, are called, from now on, as actives. It is important to point out that other thresholds $n_t$ could have been used to separate inactive from actives. Nevertheless, we tested different values of $n_t$ and the following results are very similar for large values of $n_t$ (e.g., $n_t > 50$). All the results in the next sections consider only active congresspeople.

5.3 Talking about politics over time

In order to verify the communication behavior of congresspeople over time, we compute, for each active congresspeople, their PCI for six different periods: Oct 2013 to Feb 2014 ($P_1$), Mar 2014 to Jun 2014 ($P_2$), Jul 2014 to Sep 2014 ($P_3$), Oct 2014 to Dec 2014 ($P_4$), Jan 2015 to Apr 2015 ($P_5$) and May 2015 to Sep 2015 ($P_6$). Then, for each deputy, we created a six-dimension numerical vector containing the PCI for each period. Our conjecture is that there is typical deputy behavior over time and during elections. In order to find this typical behavior, we performed a dimension reduction in the 483 × 6 matrix composed of all these vectors by means of Principal Component Analysis (PCA) [42].

If our conjecture is correct, the first PCs of the transformation will carry most of the information contained in the six-dimensional vectors. With that, we will be able to visualize the typical (or normal) behavior of the deputies and, if it is the case, who are the outliers.

Figure 8 shows the scatter plot for the two principal components (PCs) of the PCA for all deputies and also the kernel density estimation [81] for each class. First, note that the first two PCs were able to explain most of the variance in the data (75%). After analyzing the values of each of the two PCs and corresponding PCI vectors, we found an intuitive explanation for what they mean. Concerning the first PC, we found that positive values are associated with the tendency of a deputy to communicate non-political messages (low PCI), while negative values are to the tendency of a deputy to communicate political messages (high PCI) along the six periods. Concerning the second PC, negative values are associated with the tendency of a deputy to increase the ratio of political messages over time (increasing PCI), while positive values are associated with the tendency of a deputy to decrease the ratio of political messages (decreasing PCI).

Regarding how the deputies are distributed along the first two PCs, first observe that the kernels resemble a bivariate Normal distribution, which suggests the existence of a typical (or normal) behavior for each class. Second, observe that the center of the kernel for losers and reelected is near the ($-0.5, 0$) coordinate, while for newcomers is closer to (0, 0). This suggests that newcomers have, in general, a higher tendency to communicate non-political tweets than losers and reelected. Third, note that while losers tend to have more positive values along the second PC, newcomers tend to have more negative values along this dimension. This suggests that while losers tend to reduce the number of political tweets posted over time, newcomers tend to increase this amount.

In order to show another evidence that our intuition behind the first two PCs is correct, congresspeople LS1, RE1, and NC1 have similar coordinates among themselves and very different in

---

6Each deputy $i$ is characterized by a vector $\langle PCI^P_{i1}, ..., PCI^P_{i6} \rangle$, where $PCI^P_{ij}$ is calculated using only the tweets deputy $i$ posted during $P_j$. This simple temporal vector characterizes how each deputy communicates in terms of political and non-political messages over time (see Figure 9 for some examples).
Fig. 8. Principal Component Analysis for PCI

comparison with the other ones in the PCA. The same is true for the triples (LS2, RE2, NC2), (LS3, RE3, NC3) and (LS4, RE4, NC4). In Figure 9 we plot the vectors $(PCI_{P1}, ..., PCI_{P6})$ used to generate the PCA for all these twelve deputies. Observe that for all congresspeople classes in Figures 9a, 9b, 9c and 9d, the deputies with the same index (e.g. LS1, RE1 and NC1) have practically the same behavior along the time. Congresspeople LS1, RE1, and NC1 have high PCI over the entire period, i.e. they have a tendency to talk about political topics over time. Conversely, LS2, RE2, and NC2 maintain low PCI values, which means that they talk mostly about non-political topics over the entire period. Moreover, LS3, RE3, and NC3 have low PCI values before elections and high PCI values after. On the other hand, LS4, RE4 and NC4 have high values of PCI in the initial months and decrease over time. Finally, note that how the PCI of these 12 different congresspeople varies over time is coherent with their principal component coordinates and the intuition behind them as we described earlier.

When analyzing the number of political and non-political tweets over time, i.e., two years span, it is also possible to note similar behavior among the congresspeople classes. Figure 10 shows the number of political and non-political tweets over time for the three classes of congresspeople. First, observe that the number of tweets around the elections increases significantly and, afterward, decreases drastically. Analyzing individually each group, newcomer’s congresspeople increased the number of political and non-political tweets as the elections approaches and maintained almost the same frequency after the election period. Observe that before the elections the number of newcomer’s non-political tweets is always greater than the political ones. Also, the reelected tweeted constantly about politics over time and increased the frequency during the elections. Note, however, that before elections the frequency of the non-political tweets increased more than the frequency of political tweets, reaching its peak in the month of the elections. After that period, the number of political tweets is closer to non-political, though political tweets are always higher. Conversely, losers decreased drastically the number of tweets after the election and had almost the same number of political and non-political tweets along the entire period. Also, note that, contrary to the reelected and the newcomers, they maintained almost the same frequency of political and non-political posts nearby election, showing a behavior totally different from the others.

5.4 Behavior change during elections

In the previous section, we showed how much congresspeople talk about political topics over time by means of their PCI vectors and the correspondent PCA transformation coordinates. We also showed that the aggregate number of tweets significantly increased near the election term. However, those results do not quantify how much and how many deputies increase their number of social media posts as the election approaches. To tackle this problem, we formulate two hypotheses:
H1: Deputies change their posting behavior during election term.
H0: Deputies do not change their posting behavior during election term.

In order to verify these hypotheses, we performed the following task. We consider as the pre-election period the four months from February/14 to June/14 and election period the four months from July/14 to October/14. Recall that the 2014 elections occurred in October/14. Then, we counted, for each deputy, how many tweets they posted in all 17 weeks in the pre-election period and in the election period, which resulted in two distributions of weekly posting rates. Next, for each deputy, we compare the two distributions using a two-sample Kolmogorov-Smirnov (KS) test, a statistical test that quantifies the distance between the empirical distribution functions of two samples, i.e., if two samples are significantly different from each other. The null hypothesis is that the samples are drawn from the same distribution. The objective of this test is to identify deputies that changed their behavior in these two periods. In other words, if the p-value of KS test is greater than 0.05 for a deputy, then we have no evidence that this deputy changed his posting rate from the pre-election period to the election period, i.e., we cannot reject hypothesis H0. On the other hand, if the p-value of the KS test is smaller than 0.05, then we have evidence that they changed their behavior and we reject H0. In this case, if the average number of posts per week in the election period is greater than in the pre-election period, then we consider that they increased the number of posted messages in the election period. Otherwise, we consider that they decreased the number of posts in the election term.

Table 5 shows how many deputies changed their behavior during elections for political and non-political tweets among the three classes: reelected, losers and newcomers. Note that the data are
separated into two non-complementary sets of political and non-political tweets, and then between the politician classes (NE, LS, and NC). Also, deputies are divided into three categories: deputies who increased, who maintained and who decreased the number of posts in election term. Thus, if we add, for example, the percentages of reelected politicians (RE) in the political category, we get 19% + 59% + 22% = 100%.

Observe that around 40% of deputies changed their behavior of posting political tweets, 24% increase the frequency and 16% decreased. Also note that newcomers are the ones who increased the most, 44 deputies in total, which represents 32% of this class. Surprisingly, 22% of reelected decreased their frequency of political posts and 19% increased, which highlights a significantly different behavior of reelected in comparison with the other classes.

Concerning non-political messages, around 42% of deputies changed their behavior in the election period, 32% increased and 10% decreased their frequencies. In this case, reelected are the ones who increased their posting frequency the most in absolute values, 61 deputies, which represents 32% of the deputies belonging to this class. However, in relative values, newcomers are also the ones who increased their post frequency of non-political messages the most (39%). In addition, reelected are the ones who decreased the most of their posting frequency in absolute values, 21 in total. Howbeit, in relative values, 13% of losers have decreased their non-political posting frequency instead of 11% of reelected and 7% of newcomers.

In summary, the majority of deputies do not change their posting behavior as election approaches. However, these results also reveal an antagonistic attitude among deputies of the same class. There were deputies that increased their posting frequency and others that decreased. While newcomers are the ones who changed their behavior the most, especially in terms of posting more messages, losers are the ones reacted the less, i.e., this is the class with more deputies with unaltered behavior.
and with a decreasing posting frequency. Additionally, as far as we know, no previous work analyze the politician’s behavior change, before and after the election for such a large period.

Table 5. How deputies changed their posting frequency as the 2014 elections approached.

<table>
<thead>
<tr>
<th></th>
<th>Political</th>
<th>Non-Political</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RE (%)</td>
<td>LS (%)</td>
</tr>
<tr>
<td>increased</td>
<td>37 (19%)</td>
<td>33 (21%)</td>
</tr>
<tr>
<td>maintained</td>
<td>113 (59%)</td>
<td>96 (63%)</td>
</tr>
<tr>
<td>decreased</td>
<td>42 (22%)</td>
<td>24 (16%)</td>
</tr>
</tbody>
</table>

5.5 Public engagement on congresspeople political tweets

In the last sections, we saw that politicians usually post more non-political messages in the election term. From that, a question arises: does the public following of these politicians prefer tweets that are non-political or political? To answer this question, we collected the number of favorites and retweets for each tweet of our data set.

For this task, we consider all tweets from the active deputies in our dataset without discards. In addition, we performed the data collection 1 month after the last day of our data set to ensure that we captured as many retweets and mentions as possible, since a tweet receives 75% of its retweets in the first 6 hours [90]. Moreover, given that each tweet is labeled as political or non-political, we can verify which class of tweet is more popular among users.

Figure 11 shows the cumulative distribution function (CDF) of the number of favorites (11a, 11b, 11c), and retweets (11d, 11e, 11f) received by the messages posted by each class of congresspeople. For a better visualization, values greater than 30 were grouped together. Observe in all figures that political tweets tend to be favorited more than non-political ones. The number of tweets with at most one favorite is between 68% and 73% for political and between 78% and 82% for non-political tweets. Similarly, for retweets we can observe that political messages also have more retweets than non-political ones. The number of messages with at most one retweet is between 64% and 78% for political and between 78% and 88% for non-political tweets. Also note that newcomers have a dissimilar behavior, with fewer favorites and retweets than other classes. A two-sample Kolmogorov-Smirnov test, which quantifies how much the two distributions are significantly different from each other, reveals that the number of favorites and retweets are statistically different for political and non-political tweets in all classes. For favorites, the KS-statistics are respectively 0.10, 0.09 and 0.08 with p-values 0.0, 0.0 and 2.2e-252. For retweets, the KS-statistics are respectively 0.16, 0.17 and 0.12 with p-values 0.0, 0.0 and 0.0.

Figure 12 shows the popularity of the tweets posted by politicians over time, i.e., two years span. For simplicity, we summed, for each tweet, the number of times it was retweeted and favorited. First, note that political and non-political tweets become significantly more popular in election term and decay drastically afterward. Moreover, while the differences between the popularity of political and non-political tweets are small before the election term, it becomes significant during elections. Again, political tweets are much more popular than non-political ones.

Analyzing individually each group, the popularity of tweets posted by reelected reaches its peak in the month of the elections and decay afterward but maintains a growing rate along the time. Figure 12 also suggests that reelected are the most popular politicians. Also, note that the popularity of political tweets is greater than non-political ones for almost all the time. Similarly, the popularity of tweets of newcomer also increases during elections and keep growing moderately afterward.

---

Twitter swaps favorites for likes in November 2015. However, our database is from a period prior to this swap.
Additionally, the popularity for both classes is almost the same over the entire period, with political tweets reaching its peak of popularity after the election, which is also a surprising result. However, newcomers have the smallest number of popularity among the congresspeople classes. Finally, the popularity of tweets posted by loser also reaches its peak in election term. Again, the popularity of political tweets is also always higher during the analyzed period.
6 DISCUSSION
In conformity to prior research [1, 2, 10, 44] that seek to understand the communication behavior of politicians on social media, especially during important periods such as the electoral period, this study proposes a supervised methodology to identify short messages with political and non-political content posted by parliamentarians. From that, we characterized the nature of social media communications posted by Brazilian parliamentarians from October 2013 to October 2015, one year before and after the 2014 elections. The findings presented here provide multiple contributions to the literature that examines communication behavior and the use of social media by politicians.

6.1 Contributions and findings
Using the proposed methodology, we numerically characterized the politicians by the proportion of political communications they post, which we call the Political Communication Index (PCI). On one extreme, when PCI is 0, politicians only share non-political messages. On the other extreme, when PCI is 1, politicians only share political messages. This approach offers a compact and parsimonious representation of how politicians communicate in the digital environment and, contrary to the previous work [10, 21, 64], the PCI allows the characterization of political communication at a large scale, as no manual effort is necessary after the classifier is trained. We observed that congresspeople who post more about politics are the majority (57%) in our dataset, what corroborates with [29], who showed that about 70% of UK politicians tweets were used for broadcasting political messages and, conversely, contrasts with other reports [44, 63, 82], which showed a predominance of non-political messages in politicians’ communications. To the best of our knowledge, no previous work showed that this fraction is surprisingly invariant concerning how active on Twitter a politician is.

These analyses considered all congresspeople, disregarding how active they are in Twitter. Because many deputies have a small number of tweets during the analyzed period, and in order to make more solid considerations about the behavior of deputies in social media, we classified the deputies into two disjoint classes: actives and inactives. Then, we discarded the inactive parliamentarians and kept only the active ones, which represent 70% of the deputies, a number which is similar to to the one reported by Lietz et. al. [51], who showed that the majority of deputies are active on Twitter.

After filtering out the inactive deputies, we analyze the parliamentarian communication behavior over time, focusing on the electoral period. The results showed that newcomers have, in general, a higher tendency to communicate non-political tweets than losers and reelected. Moreover, losers tend to reduce the number of political tweets posted during the elections period, while newcomers tend to increase the amount of this type of communication. This is not surprising since it is natural for losers (newcomers) to stop (to start) broadcasting political messages after losing (winning) their position in the Chamber of Deputies of Brazil after the 2014 elections.

Furthermore, when we analyze the behavior of the congresspeople over the two years that comprise the dataset, it is possible to note a slightly distinguishable behavior among and within the congresspeople classes. There is an increase in the number of tweets, both political and non-political, close to the elections period and a decline soon after. This behavior is partially corroborated by [51, 87]. However, it is also possible to note that the increase is higher for non-political than for political tweets, a behavior that was also reported by other works[29, 41, 44, 54].

We also analyzed the popularity of the tweets through the number of likes (or favorites) and retweets they receive by the public. We noticed that political and non-political tweets become significantly more popular in election terms, especially the political ones, which are much more preferred than the non-political. This result is antagonistic to [49], who showed that non-political tweets are more effective in attracting favorites and retweets. Nevertheless, our results corroborate
with previous works that analyzed the content posted by congresspeople [1, 12], and with the literature on the political behavior of Brazilian parliamentarians [54, 73, 74].

It is important to point out that the use of non-political rhetoric during elections to get votes is not exclusive to Brazilian politicians. According to Bracciale and Martella [5], politicians from different leanings in Italy focused their communication strategy mainly on self-promotion, endorsement, personal issues, and daily affairs. Similar behavior can be found in the UK Parliament [41], US Congress [25], candidates in Spain [31], mayors in Turkey [83] and party leaders in Canada [82].

Regarding the use of social media by Brazilians, according to Machado et al. [53], Brazilians are considered some of the most enthusiastic users of social networks. Online platforms remain the main source of news within urban Brazil with massive content consumption and share. In this scenario, the elections have been marked by the heavy usage of social media during the campaign [68] and by the attempt to influence voters to change the outcome of the elections [53, 55]. Consequently, many politicians try to increase their influence on the network. An effective way to measure this influence is through favorites and retweets, once popular tweets could propagate multiple hops away from the source before they are retweeted throughout the network [8]. Therefore, we show that during the election term congresspeople devote much of their communications to propagate non-political messages, which can be an erroneous practice. Our analyses revealed that people who follow the politicians of our data set are more likely to prefer political tweets. This is in line with the work of Hwang et al. [40], who showed that voters mostly expect that politicians actively share their candid opinions through the open public sphere of Twitter.

6.2 Limitations and future direction

Despite the importance of this work, we acknowledge that our study has several limitations. First, although the methodology can be replicated in other contexts, the results and conclusions found in this work are specific to the Brazilian scenario around the 2014 elections and are not necessarily valid for other countries and periods. Second, when labeling a tweet as political or not, we analyze only the content of the message and do not consider its subjective information. For instance, a deputy who posts a message about a family lunch on Sunday may want to reinforce her/his image with voters, i.e, there is a political intention in the post, but not in the message content. Third, we do not investigate an alternative explanation for the results. For instance, we do not investigate whether the politician’s variables such as age or gender are correlated to the behavior on social media or whether the differences observed may be caused by political party strategy and/or by economic and geographic factors. Fourth, we do not investigate the reasons why political tweets have more retweets/favorites. For instance, a disagreement among popular users might trigger a long discussion and increase the popularity of those tweets. Additionally, we do not investigate whether there is a correlation between the number of tweets posted by each deputy class and the number of favorites/retweets received. Fifth, ideally the training data should be as faithful as possible to the real-world data for the trained classifier to be able to generalize the results. However, the loss of information in the real data sampling is almost inevitable. Thus, as we saw in the validation dataset, there is a labeling error inherent in the classifier around 10%. This error is reflected in the kappa values in the validation set which is lower than in the training set. Despite being smaller, the classifier agreement with human labelers is in the category of “moderate agreement”, which makes us confident to trust the results.

In future work, we plan to address these limitations and also evaluate the proposed methodology in different languages and election contexts. For instance, we plan to use our methodology to evaluate the tweets posted by politicians in the previous U.S. election campaign, especially the ones posted by Donald Trump, which were very popular and controversial. Additionally, we plan to create a politician’s profile by showing how much he/she talks about political issues, when they
tend to post more about politics, which topics they prefer to disclose, and which are the main agendas defended in their messages. To tackle this problem, besides the automatic classifier, we intend to use other computational methods such as topic modeling [78], a statistical model for discovering the abstract subjects that occur in a collection of documents, and text summarization [52], a process of shortening a text document to create a summary of the major points of the original document.

Moreover, our proposed methodology opens a wide range of possibilities for new applications and research. For example, one could use it to identify illegal political messages and ads during elections. Brazil has a recent law that prohibits the publication of political ads in social media without the proper information. In addition, Twitter recently announced that it will no longer allow political ads on its website [18]. We hope that this work can inspire initiatives similar to the work of Silva et al. [80], who used our classifier [60] to identify political ads posted on Facebook during the 2018 Brazilian elections. Another possibility is to deploy an independent audit system to monitor and to provide transparency to the public concerning the parliamentarians’ activities. Although there are some efforts to make the elections in Brazil transparent [15, 85], having multiple independent systems would make the monitoring of political messages more robust to attackers, preventing unfair politicians from identifying and exploiting false negatives.

7 CONCLUSION

Our study is an important step in social computing literature, providing a computational methodology to identify short messages with political and non-political content. From that, we parsimoniously characterized the social media communications posted by Brazilian parliamentarians from October 2013 to October 2015, one year before and after the 2014 elections. From the perspective of political and non-political communications, we analyze how politicians present themselves in the digital environment and how the public reacts to them. Our methodology allowed us to investigate the behavior of politicians over time, i.e., whether they change their communication and if there is a typical temporal pattern that is chosen by the majority of politicians. We noticed that politicians changed their communication behavior as the election approaches and as they were elected or not. Moreover, we showed that political and non-political tweets become significantly more popular in election terms, especially the political ones, which are much more preferred than the non-political.

REFERENCES


Do Politicians Talk about Politics? Assessing Online Communication Patterns of Brazilian Politicians


