CARACTERIZAÇÃO EM LARGA ESCALA E A LONGO PRAZO DE COMUNICAÇÕES POLÍTICAS NAS MÍDIAS SOCIAIS

LUCAS SANTOS DE OLIVEIRA

CARACTERIZAÇÃO EM LARGA ESCALA E A LONGO PRAZO DE COMUNICAÇÕES POLÍTICAS NAS MÍDIAS SOCIAIS

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Doutor em Ciência da Computação.

Orientador: Pedro O. S. Vaz de Melo

Belo Horizonte

Fevereiro de 2021

LUCAS SANTOS DE OLIVEIRA

LARGE-SCALE AND LONG-TERM CHARACTERIZATION OF POLITICAL COMMUNICATIONS ON SOCIAL MEDIA

Thesis presented to the Graduate Program in Computer Science of the Federal University of Minas Gerais in partial fulfillment of the requirements for the degree of Doctor in Computer Science.

Advisor: Pedro O. S. Vaz de Melo

Belo Horizonte

February 2021

Oliveira, Lucas Santos de.

O48l Large-scale and long-term characterization of political communications on social media[ma nuscrito] dinâmicos aleatórios/ Lucas Santos de Oliveira– 2021.
xxii, 76 f. il.

Orientador: Pedro Olmo Stancioli Vaz de Melo.
Tese (doutorado) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Ciência da Computação.
Referências: f.65-76.

1. Computação – Teses. 2. Mídias sociais – Teses. . Redes sociais on-line– Teses. 4. Redes sociais on-line – Aspectos políticos – Teses. I. Melo, Pedro Olmo Stancioli Vaz de. II. Carvalho, Sônia Pinto de. II. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Ciência da Computação. III.Título.

CDU 519.6*04 (043)

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa
CRB 6ª Região nº 1510

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Large-Scale and Long-Term Characterization of Political Communications on Social Media

LUCAS SANTOS DE OLIVEIRA

Tese defendida e aprovada pela banca examinadora constituída pelos Senhores:

Pedro Olmo

PROF. PEDRO OLMO STANCIOLI VAZ DE MELO - Orientador
Departamento de Ciência da Computação - UFMG

Jussara Marques de Almeida Gonçalves

PROFA. JUSSARA MARQUES DE ALMEIDA GONÇALVES
Departamento de Ciência da Computação - UFMG

Aline Carneiro Viana

PROFA. ALINE CARNEIRO VIANA
Inria - França

Fabrício Benevenuto

PROF. FABRÍCIO BENEVENUTO DE SOUZA
Departamento de Ciência da Computação - UFMG

Artur Ziviani

PROF. ARTUR ZIVIANI
Ciência da Computação - Laboratório Nacional de Computação Científica

Aline Marins Paes Carvalho

PROFA. ALINE MARINS PAES CARVALHO
Instituto de Computação - Universidade Federal Fluminense

Belo Horizonte, 9 de Fevereiro de 2021.

Acknowledgments

As Isaac Newton wrote, paraphrasing Bernard de Chartres: "If I have seen further it is by standing on the shoulders of Giants". This sentence summarizes my thanks to the people with whom I lived during these doctoral years. In particular, to my advisor, Pedro Olmo Stancioli Vaz de Melo. His attitude, dedication, advice, and even reprimands made me grow and see life and science in a more mature way. To Aline Carneiro Viana who welcomed me into her INRIA team and made me see science with new eyes. She made me admire her even more with her way of mother and friend who always wants to welcome, help, and, at the same time, proud of the achievements of the people around her. I appreciate your advice, guidance, and frankness. Uncountable times, when I needed some advice and confidence to overcome the research challenges that seemed insuperable, all of you were always willing to guide me to find a solution and both of you always intervened when the problem was bigger than I could handle. Thank you all.

I also thank Marcelo Amaral for his promptness and goodwill whenever I needed it. Without his help in research, this thesis would become much more difficult to achieve. To Professor Fabrício Benevenuto for the ideas and research that inspired me even before the beginning of the doctorate. I thank him and Márcio Silva for their collaboration on the project, which was of fundamental importance for this thesis.

The staff and faculty members of the Universidade Federal de Minas Gerais (UFMG), undoubtedly, were cooperative and aided me to go throughout this process. I would like to thank everyone there for their support. I also thank my Lab colleagues from INRIA and WiseMap for the conversations and exchanges of experiences that made the Lab a pleasant and fun environment to work with. I also thank Luciene, Jersiene, and the entire Moreira family that received us in Belo Horizonte making our stay even more pleasant. Special thanks to my colleagues at UESB who, in this collaborative environment, made this doctorate possible by assuming my disciplines.

Definitely, I would like to express my deepest gratitude to my dear wife Stefhani Campos Souza. Firstly, for giving me Cecília as a gift. Then, for taking care of our family and supporting me, since the master's degree, in my decisions and personal growth. You are certainly a co-author of this thesis. Without your help and companionship, none of this would be possible. We love you!

To my parents Ariosvaldo Silva Oliveira and Eliete Santos de Oliveira who, in their simplicity, gave me all the support and security to follow my path. I am very happy to make their eyes shine with this victory, which, of course, is also theirs. To my brothers Milena and Tiago, who in mutual love and complicity we are always happy with the victories of each one.

Finally and most importantly to God, whom I thank with this prayer of Saint Augustine: "Too late have I loved you, O Beauty so ancient, O Beauty so new. Too late have I loved you! You were within me but I was outside myself, and there I sought you! In my weakness, I ran after the beauty of the things you have made. You were with me, and I was not with you. The things you have made kept me from you - the things which would have no being unless they existed in you! You have called, you have cried, and you have pierced my deafness. You have radiated forth, you have shined out brightly, and you have dispelled my blindness. You have sent forth your fragrance, and I have breathed it in, and I long for you. I have tasted you, and I hunger and thirst for you. You have touched me, and I ardently desire your peace."

"A man was meant to be doubtful about himself, but undoubting about the truth: this has been exactly reversed. Nowadays the part of a man that a man does assert is exactly the part he ought not to assert-himself. The part he doubts is exactly the part he ought not to doubt"

(G. K. Chesterton)

Resumo

As mídias sociais desempenham um papel importante na formação do discurso político, criando uma esfera pública que possibilita discussões, debates e deliberações. Cientes dessa importância, os políticos utilizam as redes sociais para se autopromoverem e como forma de influenciar pessoas e votos. Como exemplo dessa afirmação, em 2018, os brasileiros elegeram democraticamente para presidente o candidato Jair Bolsonaro. Um dos feitos mais surpreendentes desse desfecho é que seu partido, o PSL, quase não teve tempo de televisão. Sua vitória só foi possível devido ao engajamento e ao ativismo de seus apoiadores nas plataformas de mídias sociais, como Twitter, Facebook e WhatsApp.

Nesse contexto, os políticos precisam decidir como se comunicarem com seus eleitores para construirem suas reputações. Enquanto alguns políticos compartilham apenas comunicações profissionais sobre suas agendas e atividades políticas, outros preferem uma abordagem não-política e informal, compartilhando comunicações sobre os mais diversos assuntos, como religião, esportes e família. Outros, no entanto, fazem mau uso das plataformas, espalhando mensagens políticas que violam os termos e condições de uso das redes sociais e as leis eleitorais. Ciente desses problemas, propomos um classificador supervisionado baseado em aprendizado de máquina que rotula todas as mensagens textuais de diferentes plataformas de mídias sociais como políticas e não-políticas. O classificador é utilizado em larga escala e é robusto a mudanças de conceito ao longo do tempo, exigindo poucas novas mensagens rotuladas a cada ano.

A partir das mensagens classificadas, pudemos caracterizar as comunicações dos políticos ao longo do tempo e fazer novas descobertas: (i) os parlamentares brasileiros mudaram seus comportamentos de comunicação ao longo do tempo; (ii) mudanças de conceito ocorreram durante eventos importantes da política brasileira; (iii) uma ascensão explosiva da direita vista pouco antes do Eleições de 2018; (iv) uma participação da direita mais ampla e mais bem distribuída do que a da esquerda e, por fim, (v) o aumento do engajamento do público ao longo do tempo.

palavras-chave: política, mídias sociais, comunicação, caracterização

Abstract

Social media play an important role in shaping political discourse, creating a public sphere that enables discussions, debates, and deliberations. Aware of this importance, politicians use social media for self-promotion and as a means of influencing people and votes. As an example of this assertion, in 2018, Brazilians democratically elected for president the far-right candidate Jair Bolsonaro. One of the most surprising feats of this outcome is that his party, PSL, had almost no television time. His victory was only possible because of his supporters' engagement and activism on social media platforms, such as on Twitter, Facebook, and WhatsApp.

In this context, politicians need to decide how to communicate with their voters to build their reputations. While some politicians only share professional communications about their political agenda and activities, others prefer a more non-political and informal approach, sharing communications about the most varied subjects, such as religion, sports, and their families. Others, however, misuse platforms by spreading political messages that violate policies and circumvent electoral laws.

Aware of these problems, we propose a supervised machine learning classifier that labels all textual messages from different social media platforms as political and non-political. The classifier runs on a large scale and it is robust to concept drifts over time, requiring few new labeled messages each year. From the classified messages, we were able to characterize the communication of politicians over time and identified new findings: (i) Brazilian congresspeople changed their communication behavior over time; (ii) concept drifts occurred during important events of Brazilian politics; (iii) the explosive rise of the right seen just before the 2018 elections; (iv) a broader and more evenly distributed right-wing participation than the left-wing, and, finally, (v) the increase of public engagement over time.

keywords: political, social media, communication, characterization

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Motivation

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 [Saad-Filho, 2013]. 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 [Watts, 2016]. This troubled political scenario is not a surprise. In 2001, Scott Mainwaring described several factors that possibly contributed to the current crisis [Mainwaring, 2001]. 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 can be 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 [Moisés, 2011].

This problematic scenario tends to be more prominent in countries where the average voter is poorer [Shin, 2017] and, as expected, during elections [Mainwaring, 2001; Samuels, 1999; Samuels and Zucco, 2014], characterizing what Samuels [1999] calls *Candidate-centric Electoral Systems*. Under such systems, party cohesion is weaker [Carey, 2007]; politicians switch parties more frequently [Heller and Mershon, 2005]; congresspeople strive to give local and individual patronage [Hicken and Simmons, 2008]; and politicians are thus more prone to corrupt behavior [Chang and Golden, 2007; Persson et al., 2001].

In *Candidate-centric Electoral Systems*, social media play an important role in shaping political scenario [Caetano et al., 2018; Tan et al., 2018; Tumasjan et al., 2010].

This the case in Brazil, where 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 [Machado et al., 2018]. In this scenario, the elections have been marked by the heavy usage of social media during the campaign [Recuero et al., 2020], by electoral laws violation [Silva et al., 2020] and by the attempt to influence voters to change the outcome of the elections [Machado et al., 2018; Marques and MontÁlverne, 2016].

The first politician to massive use social media in an electoral campaign was the former U.S. president Barack Obama [Hughes et al., 2010]. His Chicago-based campaign team used social media and technology as an integral part of their campaign strategy, not only to raise money but also, more importantly, to develop a groundswell of empowered volunteers. In Brazil, the political use of social media gained prominence during the 2010 presidential elections, when the Green Party candidate, Marina Silva, used social media as the main platform and obtained a surprisingly high number of votes [Pereira, 2011].

The peak of the social media influence during elections in Brazil occurred in 2018 when Brazilians democratically elected a candidate for president: Jair Bolsonaro [Recuero et al., 2020]. His party, PSL, which in the 2014 elections obtained only $808,404$ votes, was the party that received the highest number of votes in the 2018 House election, $11,640,033$, electing 52 federal deputies. One of the most surprising and fascinating feats of this outcome is that all candidates affiliated to PSL, including Jair Bolsonaro, had almost no television time [Council on Foreign Relations, 2018]. Before 2018, television time was considered a fundamental requirement to those aspiring to the presidential chair [Council on Foreign Relations, 2018]. Many analysts claim that his victory was only possible because of his supporters' engagement and activism in social media platforms, such as Twitter, Facebook, and WhatsApp [Recuero et al., 2020; Council on Foreign Relations, 2018; The Economist, 2019].

1.2 Problem Statement

Aware of the social media importance, politicians use social media to build their reputations and increase political engagement [Grant et al., 2010], though some are more successful in this than others [Tumasjan et al., 2010]. Part of this success is related to the communication strategy of the politicians. While some politicians only share professional communications about their political agenda and activities, others prefer a more personal and informal approach, sharing communications about the most varied subjects, such as religion, sports, and their families [Oliveira et al., 2020; Jackson and Lilleker, 2011; Gainous and Wagner, 2014]. Others, however, misuse social media platforms by spreading political messages that violate policies and circumvent electoral laws [Silva et al., 2020].

One way to investigate these problems and characterize the communication of these politicians is through the identification of *political* and *non-political* textual content posted on different social media [Caetano et al., 2018; Glassman et al., 2010; Oliveira et al., 2018, 2020; Silva et al., 2020; Golbeck et al., 2010]. Contents of the first group explicitly describe political opinions and activities, helping the public to know what to expect from politicians when they occupy public offices, e.g., "Today I attended a meeting with Minister Ricardo Veléz at the Education Commission ...". On the other hand, the contents of the second group are related to their private life or are not directly linked to their political activities and ideals, e.g., "Me and my daughter Maria Laura preparing our Christmas dinner..."

Both messages above are real and were very popular on Twitter during the 2018 elections, showing that the public might like these two types of communication. Therefore, understanding how politicians use these two types of communication leads us to the following research questions:

(R0): Can we say that, in general, the public appreciates and react to *political* and *non-political* content equally?

(R1): Are there differences in behavior depending on the electoral success?

(R2): Do politicians and their followers behave differently depending on the ideological spectrum?

(R3): Do politicians change their communication behavior on social media based on the public response?

We hypothesize that the answers to these questions can provide valuable insights for understanding the Brazilian political scenario and the communication of politicians over the years. If we are able to automatically group these posts, we could characterize all politicians that have (and use) social media by the amount of *non-political* and *political* content in their communications. With that, tools could be easily provided for the public to, for instance, identify the political agenda of politicians, identify politicians who are not transparent about their political views, and identify candidates that misuse social media platforms.

Although many studies have analyzed the content of social media messages posted by politicians [Jungherr, 2016], the problem of identifying *non-political* and *political* content, on a 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 post [Klinger, 2013; Yoon and Park, 2014; Amaral and Pinho, 2016; Dubois and Blank, 2018; Hampton et al., 2017], or focus on manual inspection of small sets of messages [Lee and Shin, 2014; Pal, 2015; Glassman et al., 2010; Jackson and Lilleker, 2011], or propose aggregate functions (e.g. count, frequency, etc.) based on specific keywords or hashtags (e.g. "abortion") to quantify how much politicians are dedicating their communications to a specific topic [Grimaldi, 2019; Garimella et al., 2018; Pond and Lewis, 2019; Shapiro and Hemphill, 2017; Gainous and Wagner, 2014; Hemphill et al., 2013; Lietz et al., 2014]. Moreover, even works that perform a classification task do not characterize the communication of politicians over time and barely consider the nature of these contents, i.e., whether they are communicating about *political* or *non-political* topics [Paul et al., 2017; Conover et al., 2011; Gao et al., 2017].

Furthermore, when the topics covered in these posts change significantly, due to noise disturbance or long-term evolution, for instance, a concept drift [Sethi and Kantardzic, 2017; He et al., 2018] is characterized. Under this circumstance, a classifier responsible for identifying political contents loses its generalization capability, failing to capture nuances in the texts from different periods [Krawczyk et al., 2017]. In this case, it is necessary to adopt mechanisms for the classifier to understand this new context.

In our first attempt [Oliveira and Vaz Melo, 2017], we compared different approaches to tf-idf (term frequency, document inverse frequency) in order to identify relevant words in the context of Brazilian politics. Therefore, we propose an alternative to the traditional idf, using Shannon's entropy [MacKay and Mac Kay, 2003] as an auxiliary feature in the idf calculation to improve the identification of these words in the tweets data dictionary of Brazilian parliamentarians. Although our proposal outperformed traditional tf-idf approaches, there was no clear distinction between the topics identified.

Therefore, in this thesis, we propose a supervised machine learning classifier that labels all textual messages from social media platforms as *political* or *non-political*. Moreover, the classifier is robust to concept drifts, where topic changes are identified over the years by an unsupervised drift detection method [Costa et al., 2018] that applies an Active Learning [Zliobaite et al., 2014] technique in which the true labels of the messages are required only after drift detection. Finally, we use a transfer learning approach [He et al., 2018] to retrain the classifier after drift detection that maintains the high sensitivity to the long-term data changes. From the classified messages, we were able to characterize the communication of politicians over time. We could identify, for instance, if they change their communication behavior over time, e.g. around elections or when they take over or leave public office.

1.3 Contributions

To demonstrate the usefulness of our proposed methodology, we applied it in two case studies related to Brazilian political communications.

(C1) Objective: To Analyze whether the nature of politicians' communications changes around the elections and whether this behavior is related to electoral success.This objective answers the research questions **R0** and **R1**, and partially **R3**.

For this task, 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, tried to re-elect 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. 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). With that in hand, we identify if they change their communication behavior over time, e.g. around elections or when they take over or leave public office. Moreover, 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.

(C2) Objective: To Analyze if there are differences in the communication strategies used by the *right*, *left*, and *center* politicians over time. Also, try to understand the growth of the right-wing in recent years. This objective answers the research questions **R0**, **R2** and **R3**.

For this task, we collected more than $3M$ tweets from the 914 politicians with a valid Twitter account and who occupied a public office in Congress for at least one year from October 2013 to October 2019. These congresspeople were labeled as *right* (R), *left* (L), and *center* (C) based on their party ideology. By separating the deputies' tweets according to these labels, we were able to analyze the communication behavior of politicians over 6 years. We reveal striking and even abrupt changes in the behavior of politicians and the public. While before 2016 most of the communications posted by politicians were *non-political* after 2016 *political* communications started to heavily dominate. Also, although total left-wing politicians post more on social media, the participation of right-wing politicians is broader, more evenly distributed, and increased significantly after the 2018 elections. The overall engagement of the public with politicians grew constantly since 2013, especially with *political* communications. Differences in engagement with respect to ideology and content increased greatly after 2015, which is possibly catalyzing the increase of polarization in Brazil [Recuero et al., 2020; Weber et al., 2013]. Politicians from the center became much less popular than those from the left and right. Also, after 2017 the public engagement with right-wing communications became much higher than with those from the left, even considering the usually less popular *non-political* messages. In summary, our results corroborate with the hypothesis that social media was fundamental for the recent rise of the *right* in Brazil and its success in the 2018 elections.

This thesis extends the findings of previous research in four-fold contributions, which are described as follows:

- A computational methodology for identifying *political* and *non-political* textual content that is robust to concept drifts and that can be applied in different languages.

- Two parsimonious characterizations of Brazilian congresspeople communications that comprise the 2014 and 2018 elections. We analyze the nature of these communications and how they may be linked to concepts that characterize Brazilian political relations. To the best of our knowledge, we are the first to analyze such a large number of messages ($3,337,744$ tweets) from a significant number of deputies (914) in such a long period (6 years span).

- To the best of our knowledge, we are the first that shows quantitatively and qualitatively at a large scale that *non-political* discourse arises and dominates the campaigns during the 2014 elections. Surprisingly, in spite of that, we showed that *political* messages are far more popular among the public. Also, we are the first to characterize and demonstrate quantitatively the rise of the right in Brazil in this decade from social media data.

- Finally, we provide two datasets for the research community: one containing more than 3M tweets from 914 Brazilian politicians; other containing a set of 3, 116 tweets labeled as *political* and *non-political* comprising the years from 2013 to 2019;

The following publications are the current results of this thesis:

- **How to Find the Relevant Words Politicians Use in Twitter?**. Oliveira, L. S. D.; and Vaz Melo, P. O. In Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web - WebMedia '17, pages 465–468, New York, New York, USA, 2017. ACM Press.
- **When Politicians Talk About Politics: Identifying Political Tweets of Brazilian Congressmen**. Oliveira, L. S.; de Melo, P. O. S. V.; Amaral, M. S.; and Pinho, J. A. G. , International AAAI Conference on Web and Social Media (ICWSM): 664–667. 2018.
- **Facebook Ads Monitor : An Independent Auditing System for Political Ads on Facebook**. Silva, M.; Oliveira, L. S. D.; Vaz de Melo, P. O. S.; Benevenuto, F.; Andreau, A.; and Goga, O. In Proceedings of The Web Conference 2020, Taipei, Taiwan, 2020. ACM **[Honorable Mention]**
- **Do Politicians Talk about Politics? Assessing Online Communication Patterns of Brazilian Politicians**. Oliveira, L. S. D.; Vaz-de-Melo, P. O. S.; Amaral, M. S.; and Pinho, J. A. G. ACM Transactions on Social Computing (TSC), 3(4). September 2020.

1.4 Work Organization

The rest of this thesis is structured as follows. In Chapter 2, we describe the related work and show an overview of the Twitter datasets. In Chapter 3, we present the methodology for *political* message classification and we characterize the political communication during the elections. Thereafter, in Chapter 4, we present the methodology to keep the classifier accurate over time and a characterization of the political communication over the years. Finally, in Chapter 5, we summarize and compare our findings with other works in the literature; show the limitations of this thesis and the future directions; and conclude our thesis.

Chapter 2

Related Work

2.1 Political communication characterization

Computational communication science, social media, and big data are remaking and revolutionizing interpersonal communication [Cappella, 2017]. In fact, computational approaches have the capacity to gather and process large quantities of information quickly to serve the public good [Shah et al., 2015]. 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 [2014], 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. [2011] 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. [2014] 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 and Richards [2017] by means of an algorithm that assigns citizens to political spaces based on their social media communications. They concluded that political networks are not static neither solid, as actors consider turbulence and volatility as structures of political opportunity to gain visibility.

There is already evidence that the action and interaction of voters in social media can influence their inclination to vote or not for a candidate. Maruyama et al. [2014] found a relation between Twitter use and the voting choice. They investigated how using a social network while watching a political event could influence the experience of a voter, especially when the user actively participates by posting messages about the event. Pal et al. [2018] examined the function and public reception of critical tweeting in online campaigns of four nationalist-populist politicians during major national election campaigns. They found that cultural and political differences impact how each politician employs their tactics. In South Korea, Lee and Shin [2012] and Lee and Shin [2014] designed experiments to investigate how the level of interactivity in politicians' Twitter communication affects the public's cognitive and affective reactions. They found that exposure to high-interactivity Twitter pages induces a stronger sense of direct conversation with the candidate, which, in turn, led to more positive overall evaluations of the candidate and a stronger intention to vote for him. However, politicians are not only the ones who try to influence the elections. Hemphill and Roback [2014] examined common strategies of lobbying on Twitter and found that assumed citizens used Twitter to merely shout out their opinions on issues and utilize a variety of sophisticated techniques to impact political outcomes. Kim et al. [2018] used an ad tracking app that enabled them to trace the sponsors/sources of political campaigns and unpack targeting patterns. Their empirical analysis identified "suspicious" groups, including foreign entities, and operating divisive issue campaigns on Facebook.

Automatic classification of social media content 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. [2017] 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. [2011] 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. [2017] 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 our methodology (see Section 3.2), 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. [2017] requires an external database of news articles for topic model training. Moreover, in the Paul et al. [2017], Conover et al. [2011] and Gao et al. [2017] approaches there is substantial overlap between streams associated with different *political* hashtags because many tweets contain multiple hashtags. Furthermore, as shown in Section 3.2, this approach suffers from some restrictions as politicians use the same hashtags for both *political* and *non-political* messages. Additionally, a simple occurrence of any term present in this final list in a post is enough to label such a post as relevant which could lead to misclassification.

2.2 Political content analysis

With the rise of social media, many researchers got interested in exploring their role in politics. According to Karlsen and Enjolras [2016], 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 [2019] 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. [2013] 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 [Mascaro and Goggins, 2011; Saldivar et al., 2019]. Grant et al. [2010] 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. Differently, Etudo et al. [2019] investigated the effects of Russian ads and what is the relation with Black Lives Matter Protests and found that Russian ads related to police brutality were issued to coincide with periods of higher unrest. Similarly, Ribeiro et al. [2019] investigated how malicious Russian advertisers were able to run ads with divisive or polarizing topics (e.g., immigration, race-based policing) at vulnerable subpopulations.

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 [Howard et al., 2016], to spread misinformation [Forelle et al., 2015] and even to incentive criminality [Savage et al., 2015]. In the specific case of Brazil, researchers studied polarization, hyperpartisanship, and disinformation during the election [Recuero et al., 2020; Machado et al., 2018]; the influence of social networks in the electoral process [Marques and MontÁlverne, 2016]; the understanding of politicians behavior on social media [Oliveira et al., 2018, 2020]; and proposed systems to the increase of transparency during elections [Silva et al., 2020].

Regarding the presence of world political leaders in social media, Barberá and Zeitzoff [2018] suggest that leaders from many countries have social media accounts (e.g. Argentina, France, Ukraine, Tunisia, South Africa, Philippines, Japan, etc), even in those countries with limited press freedom, such as Iran, Kyrgyzstan, or Cuba. 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 [Enli, 2017]. According to Pain and Chen [Pain and Masullo Chen, 2019], 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 [Library, 2019]. Similar findings were presented by Gonawela et al. [2018] 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 Masullo Chen [2019] 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.

2.3 Political vs. Non-Political content

Given the importance of social media messages in politics, many works attempted to identify *non-political* and *political* content in social media. Gainous and Wagner [2014] 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 $\approx 44\%$ of the messages were related to campaign announcements, $\approx 19\%$ were *non-political* messages, $\approx 18\%$ were attacking to other politicians, and $\approx 17\%$ related with policies. A different methodology was employed by Glassman et al. [2010], 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 From another manual analysis of U.S. congresspeople tweets, Golbeck et al. [2010] 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 [2011] manually analyzed the content 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 [2015] 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 social media, since similar conclusions were reached by several other studies [Jungherr, 2016; Small, 2010], that is, *non-political* content apparently dominates politicians' social media messages. However, Marques and MontÁlverne [2016] investigated Fortaleza's city councilors' tweets and found that most messages are related to the promotion of ideas, negative campaign, mobilization, and promoting campaign events. Finally, Hwang [2013] analyzed how Korean young adults evaluate the use of Twitter by South Korean politicians and conclude that politicians who actively use Twitter are seen as more credible and, as a consequence, are more positively evaluated by young adults.

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 [2017], while investigating political communication styles, divided communications into four dimensions, and one of them, the "Topic" dimension, identifies the main argument of the message, 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 [2015], inspired by [Jackson and Lilleker, 2011], defined a "banal" message "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 3.2, our own functional definition of what can be considered a *political* message.

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 specifically, 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 chapter, in the present thesis 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 [Oliveira et al., 2018].

2.4 Twitter Datasets

In this section, we describe the two public tweet datasets of tweets posted by Brazilian deputies. The description and characteristics of these datasets are described in more details in Chapters 3 and 4. In Figure 2.1, we show an overview of the datasets used to validate the proposed methodology and to characterize the behavior of deputies over time.

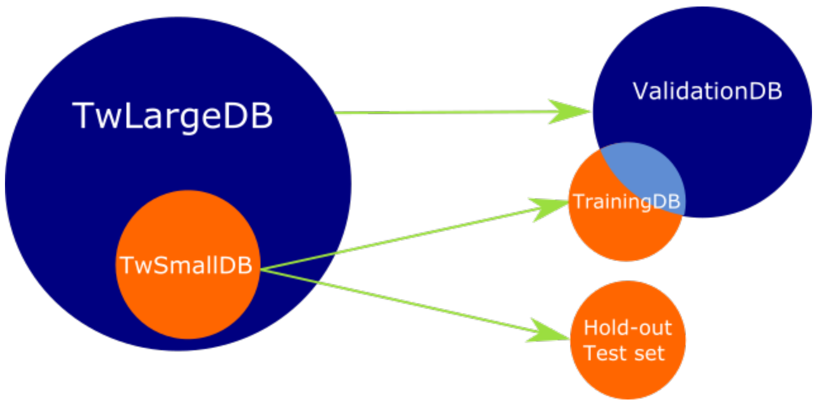


Figure 2.1: Overview of the public Twitter datasets.

The first dataset comprises the years of 2013, 2014, and 2015, and it was collected to train and validate the *political* message classifier and to characterize the communications of politicians in this period (see Chapter 3). From now on, we call this TWSMALLDB. The second, called TWLARGEDB, was collected later and comprises the TWSMALLDB. In addition, it expands by adding tweets posted by Brazilian deputies between 2013 and 2019. This dataset enabled us to perform a broader analysis of the communication behavior of deputies over the years.

From the TWSMALLDB, we extracted two datasets for training and validation of the proposed methodology (see Chapter 3). The first, called TRAININGDB, contains $2,000$ manually labeled tweets ($1,000$ *political* and $1,000$ *non-political*). The second, the hold-out test set, contains 814 labeled tweets with the unspecific distribution. Likewise, to validate the methodology over time (see Chapter 4), we created a dataset that comprises part of the TRAININGDB dataset and contains $3,116$ labeled tweets ($1,558$ *political* and $1,558$ *non-political*). From now on, we call this VALIDATIONDB.

Chapter 3

Characterizing Political Communications in Elections

In this chapter, we present a methodology to classify messages into two categories: *political* and *non-political*. From that, we characterize each politician according to the number of *political* and *non-political* messages they posted from October 2013 to October 2015. In addition, we consider for this analysis the three sets of politicians according to their position before and after the 2014 elections: *reelected* (RE), *loser* (LS) and *newcomer* (NC).

3.1 TwSmallDB dataset

To perform the politician communication characterization, 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`. We call this dataset TwSmallDB. 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* website[1]. 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.

[1]`https://dadosabertos.camara.leg.br`

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 3.1 summarizes the TwSmallDB. 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, on average, the most active congresspeople on 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.

Table 3.1: Dataset Summary

	# tweets	**# deputies**	**average**
Reelected	355,450	273	1,302
Newcomer	183,533	202	908
Loser	212,134	217	977
total	751,117	692	1,062

3.2 Identification of Political and Non-Political Content

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 are providing to the country. Congresspeople that post mostly *non-political* tweets are failing to account for the citizens who elected them.

Inspired by what Pal [2015] defined as a "banal" message and the arguments that compose a political communication identified by Bracciale and Martella [2017], we propose the following definition of a "political message":

Definition 1 *A* ***political message*** *is a message posted on social media whose content expresses subjects related to fundamental issues about the 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.*

Our intention is to provide an unequivocal definition of political message to minimize the subjective interpretations. Therefore, from Definition 1, we modeled the problem of identifying *political* and *non-political* messages as a supervised binary classification problem. Figure 3.1 shows the methodology, which can be divided into four steps. First, we sample from TwSmallDB a large set of tweets posted by deputies that are evenly distributed across deputies and across time. We call this dataset TrainingDB, which consists 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 [Mikolov et al., 2013a], 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 [Oliveira et al., 2018], 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 described next.

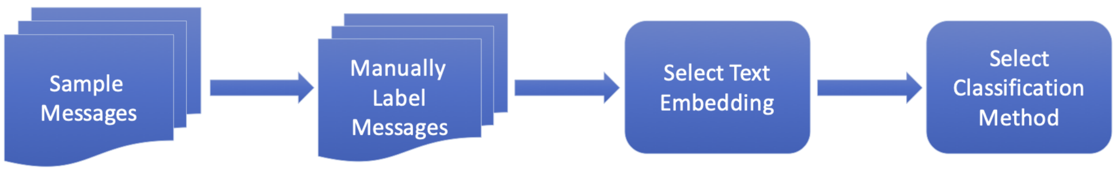


Figure 3.1: Overview of the methodology to identify *political* and *non-political* messages.

3.2.1 Meta Parameters

The process of identifying *political* messages using a classification approach involves several methodological decisions. These decisions can be thought of as the *meta parameters* of the methodology, which affects 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 3.2 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 the number of dimensions represents the vector size to which each word of the text is mapped; and the Convolutional Neural Network (CNN) architecture as the classification method.

Table 3.2: Meta Parameters

labeled tweets	period	embedding	embedding size	classification method
100	random	Word2Vec C-BoW	100	CNN
500	few months	Word2Vec Skip-Gram	300	LSTM
1000	few deputies	Glove		FastText
2000		Word2Vec C-BoW over our dataset		
		Word2Vec C-BoW hashtag over our dataset		

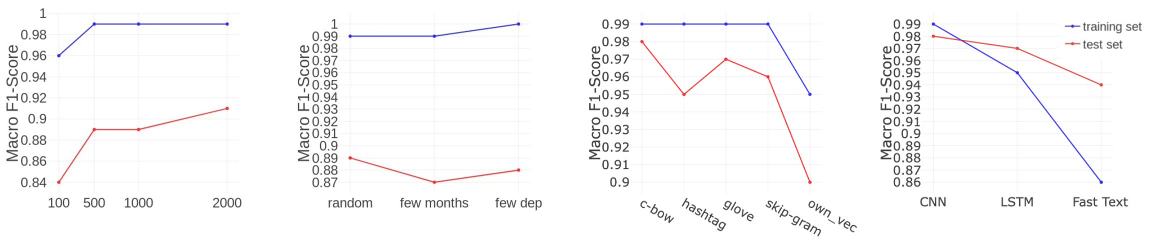
3.2.2 Sampling and labeling messages

The first challenge is to select the messages 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 messages are manually labeled as *political* and half as *non-political*. To do that, we adopt the following process. Given the whole collection of messages $\mathcal{U}$, we create a subset $\mathcal{P}$ of messages labeled as *political*, such that $\mathcal{P} \subseteq \mathcal{U}$. For each message $m \in \mathcal{U}$, m is labeled as *political* and assigned to the subset $\mathcal{P}$ if and only if there is a political position on m according to Definition 1. In case a political position is not clearly stated in the message, m is labeled as *non-political*. Note that this process is different from labeling a message 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 message is considered *political*.

The initial and most comprehensive sample of tweets to be labeled was randomly selected from the TWSMALLDB. 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 a size of $1,000$ tweets with this label and that their distribution over the months matches the original data. We call this dataset TRAININGDB. In order to check whether the distributions match, we calculated the KL divergence [Conover et al., 2011] between TrainingDB dataset of 1,000 *political* and *non-political* tweets and the TwSmallDB dataset. 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 TrainingDB is very similar to the original data. Then, the other samples of 100, 500, and 1,000 tweets were randomly sampled from TrainingDB of 2,000 labeled tweets. The other 814 tweets that were left out of this balanced set are used as a hold-out test set of arbitrary distribution.

Figure 3.2a 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 messages which the classifier labeled as *political* are in fact *political*. Having high recall means that from the total number of *political* messages, the classifier correctly labeled the majority as *political*. Moreover, the F1 score reaches its best value at 1 and the 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 an F1 score of 99% in the training set and 91% in the hold-out test set.



(a) Sample size (b) Dispersion (c) Embeddings (d) Classification method

Figure 3.2: Classification results. F1 scores for different configurations.

3.2.3 Dispersion of labeled messages

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 3.2b 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.

3.2.4 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 [Hartmann et al., 2017], which is able to produce an embedding matrix for a vocabulary of 1.3 trillion terms. These vectors were produced using the following methods: *Word2Vec C-BoW* [Mikolov et al., 2013a], *Word2Vec Skip-Gram* [Mikolov et al., 2013a] and *Glove* [Pennington et al., 2014]. Additionally, we trained the Word2Vec C-BoW model using the TWSMALLDB. We also used the pre-trained *Word2Vec C-BoW* [Hartmann et al., 2017] 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 3.2. Figure 3.2c exhibits that *Word2Vec C-BoW* and *Glove* have the same 99% of F1 score in the TRAININGDB. On the other hand, the result in the hold-out test set shows that *Word2Vec C-BoW* achieved a higher F1 score than *Glove*.

Also, it is important to note that the *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 *Word2Vec C-BoW* from [Hartmann et al., 2017]. Moreover, despite hashtags are good predictors of *political* tweets [Conover et al., 2011; Bovet et al., 2018; Pond and Lewis, 2019], 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 the Rio Grande, spreading the ideas of the well-prepared pre-candidate for governor in order to make our state strong again. #heinze #luiscarlosheinze #oriograndeforteoutravez #oriograndetemjeito.”

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

3.2.5 Classification method

The last decision is to choose which Neural Network architecture to use. More specifically, we compared three different architectures: Convolutional Neural Networks (CNN) [Kim, 2014], Long Short Term Memory Networks (LSTM) [Hochreiter and Schmidhuber, 1997] and FastText (the classifier only)[Joulin et al., 2016]. 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 with 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. Neural networks require to have inputs of the same size. For this purpose, we use the padding of size 28. Finally, 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 3.2d 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.

In fact, we tested many other architectures and classifiers, such as LSTM, LR, SVM, but neural network architectures obtained the better and most consistent results (Table 3.3).

Table 3.3: Political Message Classifiers

	Training set			**Test set**		
Classifier	Precision	Recall	F1-score	Precision	Recall	F1-score
CNN	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**	**0.98**
LSTM	0.95	0.95	0.95	0.97	0.97	0.97
FastText	0.88	0.87	0.86	0.94	0.94	0.94
SVM	0.87	0.87	0.87	0.81	0.81	0.81
Gradient Boosting	0.81	0.81	0.81	0.88	0.88	0.88
Logistic Regression	0.88	0.88	0.88	0.88	0.87	0.87

3.2.6 CNN for text classification

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 [Goldberg, 2017; Harris, 1954]. Such vectors, also called bag-of-words or bag-of-n-grams [Harris, 1954], are attractive due to their simplicity, efficiency, and often surprising accuracy. In this direction, recent work in learning dense vector representations of words [Mikolov et al., 2013a] using neural networks [Deriu et al., 2017; Bengio et al., 2003; Turian et al., 2010; Mikolov et al., 2013b; Pennington et al., 2014; Goldberg, 2017] were proposed. In all these studies, including Word2Vec [Mikolov et al., 2013a], the one we use in this thesis, 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 [Goldberg, 2016], whose resurgence greatly impacted text classification tasks [Yin et al., 2017]. In particular, Convolutional Neural Networks [Kim, 2014] 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 [Kim, 2014]. According to Yin et al. [2017], CNN performs well on tasks where feature detection in the text is more important. For example, searching for angry terms, sadness, abuses, named entities, etc. However, other types of Recurrent Neural Networks (RNN) such as Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM), perform better on tasks where the length of the text is important. These types of tasks include question-answering, translation, etc.

As CNN performed better among the tested classifiers, we inspire in the Camacho [2019] tutorial to explain how CNN works for text classification. First, before feeding a message into a neural network as an input, we have to convert each word into a numeric value, known as word embedding. Word embeddings are vectors of a specified length of n and each vector of n positions represents one word.

These embeddings are formed in an unsupervised manner by training a single-layer neural network—a Word2Vec model—on an input word and a few surrounding words in a sentence. Words that show up in similar contexts, such as "princess", "queen", and "woman" will tend to have similar vectors, given by cosine similarity, that point in roughly the same direction.

Figure 3.3 shows a simplified example of how a CNN classifies a message as *political* or *non-political*. In this figure, we can see some examples of words that were transformed into vectors. For example, the word "election" is represented by the vector $[0.8, 0.9, 0.1, 0.5, 0.1]$ of 5 positions. Therefore, we can represent a message as a list of word embeddings, which is used as input to the convolutional neural network.

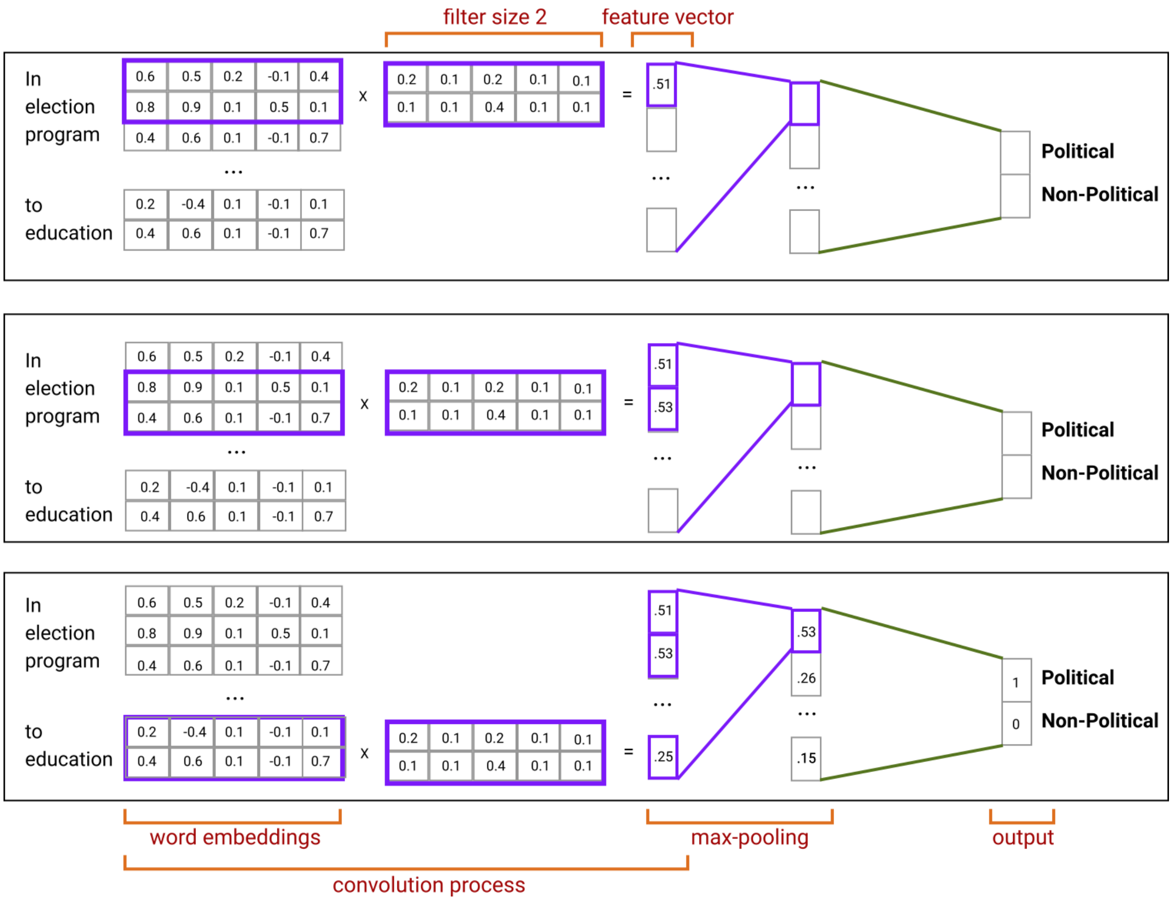


Figure 3.3: A simplified example of how CNN performs on text classification.

On text classification, a convolutional kernel is a sliding window, which is used to look at embeddings for multiple words. The height of the kernel is the number of embeddings it sees at once, similar to representing an n-gram in a bag-of-n-gram model. The width of the kernel should span the length of an entire word embedding. Figure 3.3 shows an example of a kernel of height 2 (a 2-gram) and width 5 (embedding size), which is represented by the purple rectangle.

We use a 2×5 convolutional kernel to look at two words in this example. The downwards-direction represents time, so, the word "election" comes right after "In" in this short sequence. The kernel filter weights and embedding values are multiplied in pairs and then summed to get a single output value of 0.51. Likewise, the next two vectors that represent the words "election" and "program" are multiplied by the filter of size 2 and result in an output value of 0.53.

To process an entire sequence of words, these kernels slide down a list of word embeddings, in sequence. This is called a 1D convolution because the kernel is moving in only one dimension: time. Sometimes a convolutional kernel does not perfectly overlay on the word embeddings and so some padding may need to be included to account for the height of the kernel.

Therefore, the convolution process can be viewed as window-based feature extraction, where the features are patterns in sequential word groupings that indicate traits of the grammatical function of different words. Recognizing these kinds of high-level features is useful in text classification tasks, which often rely on general groupings. For example, in *political* message classification, a model would benefit from being able to represent *political* and *non-political* groupings. A model could use those general features to classify entire texts.

In a typical convolutional neural network, one convolutional kernel is not enough to detect all the different kinds of features that will be useful for a classification task. In this thesis, we use 120 kernels in total; 40 kernels for each height: 3, 4, and 5. These heights effectively capture patterns in sequential groups of 3, 4, and 5 words. We chose a cutoff of 5 words because words that were farther away than that were generally less relevant or useful with respect to identifying patterns in a message. The stacked output feature vectors that arise from several of these convolutional operations are called convolutional layers.

When we are trying to classify *political* messages, and we see the phrase, "election program", we consider it as a good indicator of a *political* message. In order to indicate the presence of these high-level features, we need a way to identify them in a vector, regardless of the location within the larger input sequence. Therefore, the max-pooling operation forces the network to retain only the maximum value in a feature vector, which should be the most useful local feature.

In Figure 3.3, in the last CNN, we can observe that a vector of features is generated right after the convolution step. Then, in the max-pooling stage, only the max-value of this vector is passed forward. In this example, the max-value is 0.53. Finally, the max-values of each convolutional feature vector from different kernel filters, are concatenated and passed to a final fully-connected layer that produces the *political* or *non-political* output.

3.2.7 CNN architecture for political text classification

In the previous subsection, we explained the stages of a Convolutional Neural Network for text classification. In Figure 3.4, we show the architecture of the Convolutional Neural Network used in this thesis for *political* and *non-political* message classification, which is similar to [Kim, 2014] architecture, however, with the following adjustments and improvements in the network parameters.

In the input layer, we represented each word in a message as a dense vector retrieved from Word2Vec C-BoW [Mikolov et al., 2013a] with 300 vector positions. For each message, we use the padding of 28 to ensure that it has the same size. 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 a 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 message is *political* and 0 if *non-political*. Moreover, we optimized the neural network by means of cross-entropy loss function using RMSProp [Duchi et al., 2011] optimization algorithm.

3.2.8 Evaluation and Validation

Before running the classifier, we labeled 2,000 tweets evenly distributed across time and congresspeople and further validated them by other six independent researchers from the applied social sciences field, with research related to the political context investigated in this thesis. In this dataset that we named TrainingDB, we evenly

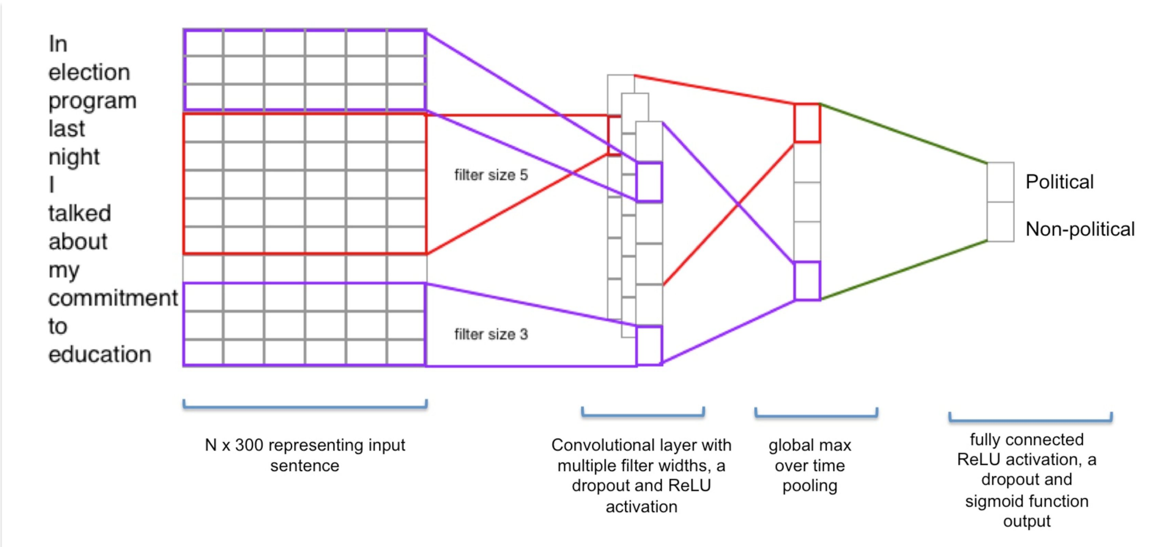


Figure 3.4: Convolutional Neural Network architecture for political message classification.

distributed 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 TRAININGDB by month and by deputy. After that, we classified the tweets in each group using our CNN and calculated the F1 score. Figure 3.5 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 TRAININGDB 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 TRAININGDB. The second contains 200 randomly selected tweets (100 *political* and 100 *non-political*) that were labeled by the classifier and not labeled previously. All tweets from this second dataset have been prepossessed to simulate how the classifier receives them as input. Then, we asked the 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

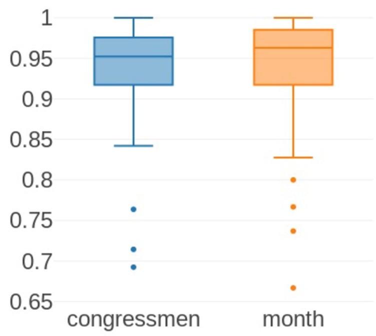


Figure 3.5: F1 score per deputy and per month.

agreement percentage and the Cohen's Kappa coefficient (κ) [Oren and Gilbert, 2011; Landis and Koch, 1977; Savage et al., 2015].

Cohen's kappa measures the agreement between two raters who each classify N items into C mutually exclusive categories. The definition of κ is:

$$\kappa \equiv \frac{p_o - p_e}{1 - p_e},$$

where p_o is the relative observed agreement among raters (identical to accuracy), and p_e is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category. If the raters are in complete agreement then $\kappa = 1$. If there is no agreement among the raters other than what would be expected by chance (as given by p_e), $\kappa = 0$. It is possible for the statistic to be negative, which implies that there is no effective agreement between the two raters or the agreement is worse than random [Wikipedia contributors, 2021]. According to Landis and Koch [1977], when κ 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 3.4 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 Landis and Koch [1977], 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 [Landis and Koch, 1977]. 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 3.4: 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*.

	Res 1	**Res 2**	**Res 3**	**Majority**	**Res 1 & 2**	**Res 1 & 3**	**Res 2 & 3**
Kappa coefficient κ	0.70	0.70	0.68	0.73	0.78	0.73	0.80
Agreement %	85%	85%	84%	86%	90%	88%	91%

The agreement results between the three researchers and the classifier in the *classified validation set* are described in Table 3.5. 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 $\kappa = 0.54$.

Observe that the majority vote has the same κ 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. Possibly, due to the prepossessing of the messages, however two of these values are still in the category of "moderate agreement" [Landis and Koch, 1977]. Moreover, in recent work and under a similar context, Resende et al. [2019] 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 Marcius Sidarta."* Probably the classifier took into account the number 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 3.5: Cohen's Kappa and Agreement percentage among our classifier and three independent researchers over the *classifier validation set*.

	Res 4	**Res 5**	**Res 6**	**Majority**	**Res 4 & 5**	**Res 5 & 6**	**Res 4 & 6**
Kappa coefficient κ	0.55	0.49	0.49	0.55	0.60	0.64	0.59
Agreement %	77%	74%	74%	77%	80%	82%	79%

To illustrate the performance of our classifier, we used the same methodology employed by Grant et al. [2010] to summarize the content of different sets of tweets. In Figure 3.6, we show the word clouds for the sets of *political* and *non-political* tweets during the elections period. In Figure 3.6a, we can identify the most popular words related to politics, which include "debate", "federal", "support", "government" and "policy". In contrast, Figure 3.6b 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"[2], "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.

(a) Political (b) Non-political

Figure 3.6: Word clouds of *political* and *non-political* messages in election term.

[2] In Portuguese, "I published" can be written using a single word: "publiquei".

3.3 Results

After the classification of all messages from TWSMALLDB, 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 Twitter and the propensity to post more about political issues. From now on, we call the decision to post a more or less political tweet as her/his *communication behavior*. In the following sections, we analyze and describe the *communication behavior* of Brazilian congresspeople.

3.3.1 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_i of politician i as:

$$PCI_i = \frac{|P_i|}{|NP_i| + |P_i| + 1}, \tag{3.1}$$

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[3]. 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 3.7 shows the scatter plot of the PCI and the number of tweets for all politicians of TWSMALLDB, 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

[3]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 equal 1.

20% share of *non-political* messages. Particularly, note that congresspeople who post more about politics are the majority (57%).

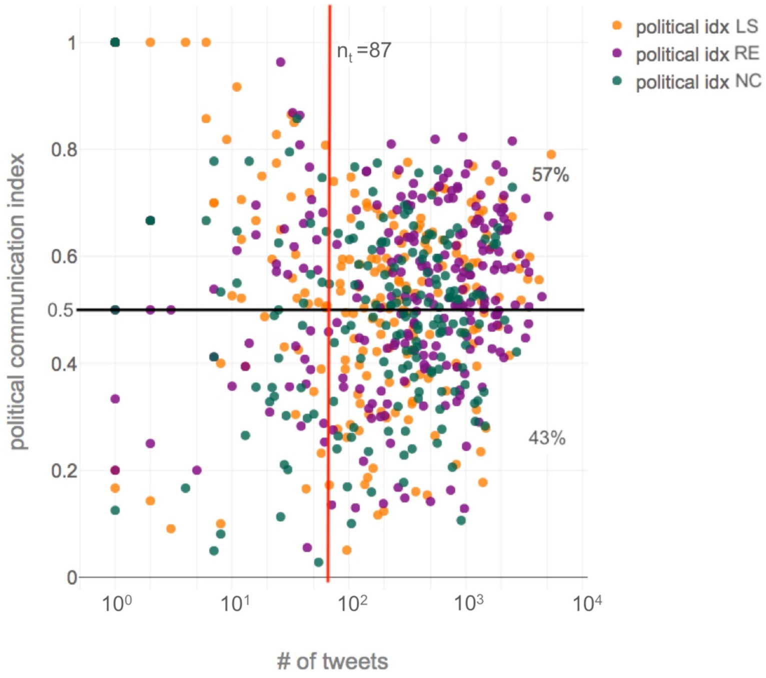


Figure 3.7: Scatter plot for the PCI and the number of tweets for Losers, Reelected and Newcomers

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 3.8, 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 [Graham et al., 2013; Gainous and Wagner, 2014] also pointed for at least 20% of *non-political* messages, to the best of our knowledge no previous work showed that this fraction is surprisingly invariant with respect to how *active* a politician is.

3.3.2 Consistent use of Twitter

As depicted in Figure 3.7, 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

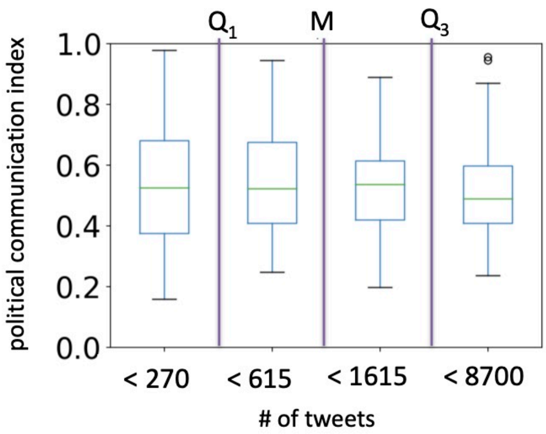


Figure 3.8: Distribution of PCI per number of tweets

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, on 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 3.7 denotes this threshold n_t. Observe that 30% of the deputies have not posted messages consistently on Twitter over the two years of our analysis. For the remainder of this thesis, we call such deputies *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.

3.3.3 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

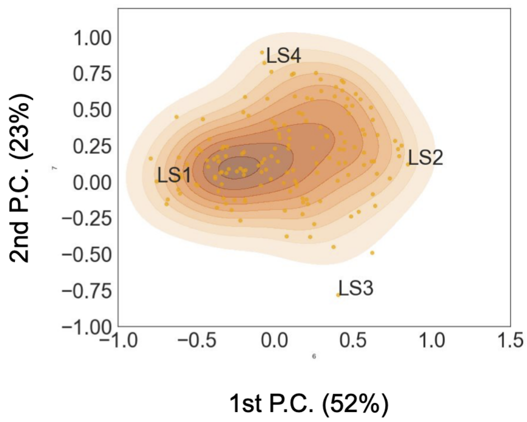
(P_1), Mar 2014 to Jun 2014 (P_2), Jul 2014 to Sep 2014 (P_3), Oct 2014 to Dec 2014 (P_4), Jan 2015 to Apr 2015 (P_5) and May 2015 to Sep 2015 (P_6). Then, for each deputy, we created a six-dimension numerical vector containing the PCI for each period[4]. Our conjecture is that there is typical deputy 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) [Jolliffe, 1986].

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.

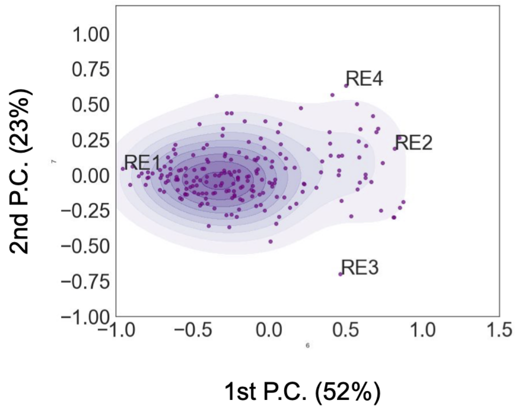
Figure 3.9 shows the scatter plot for the two principal components (PCs) of the PCA for all deputies and also the kernel density estimation [Silverman, 2018] for each class. First, note that the first two PCs were able to explain most of the variance in the data (75%). After analyzing the values of each 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 with 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 with 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.

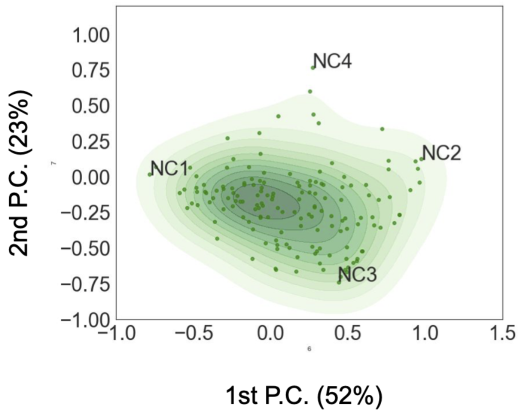
[4] Each deputy i is characterized by a vector $(PCI_i^{P_1}, ..., PCI_i^{P_6})$, where $PCI_i^{P_j}$ is calculated using only the tweets deputy i posted during P_j. This simple temporal vector characterizes how each deputy communicates in terms of *political* and *non-political* messages over time (see Figure 3.10 for some examples)



(a) Losers.



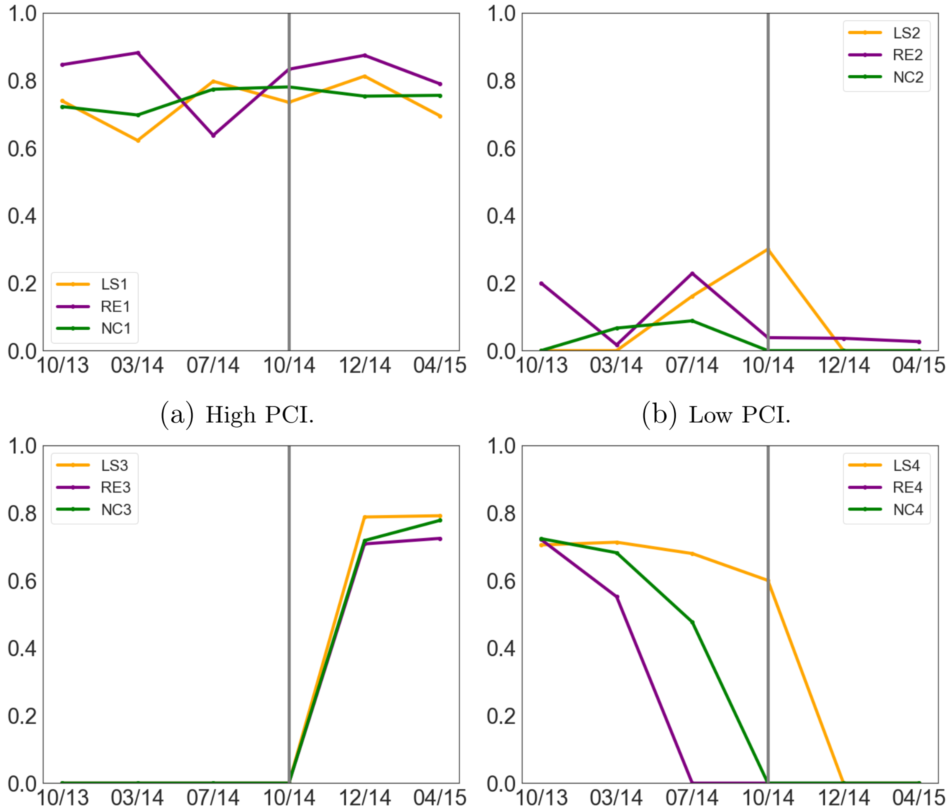
(b) Reelected.



(c) Newcomers.

Figure 3.9: Principal Component Analysis for PCI

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 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 3.10 we plot the vectors $(PCI_i^{P_1}, ..., PCI_i^{P_6})$ used to generate the PCA for all these twelve deputies. Observe that for all congresspeople classes in Figures 3.10a, 3.10b, 3.10c and 3.10d, 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.



(a) High PCI. (b) Low PCI.

(c) Increase PCI. (d) Decrease PCI.

Figure 3.10: The PCI over the time for nine selected deputies. The vertical line marks the Election Day.

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 3.11 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.

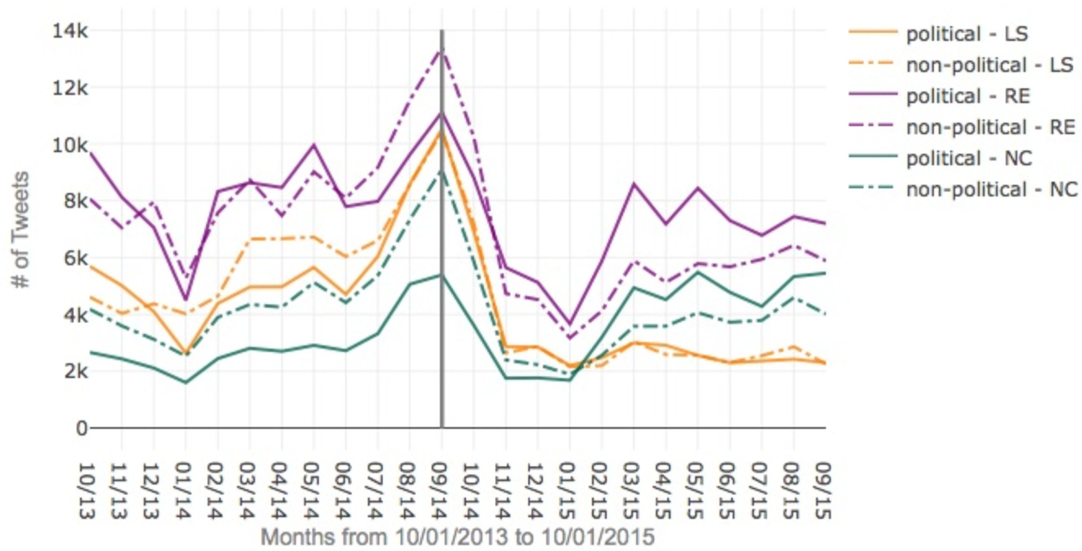


Figure 3.11: *political* and *non-political* tweets distributed over time.

3.3.4 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 the election term.

H0: Deputies do not change their posting behavior during the election term.

In order to verify these hypotheses, we performed the following task. We consider the pre-election period the four months from February/14 to June/14 and the election period the four months from July/14 to October/14. Recall that the 2014 elections occurred on 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 the 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 3.6 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 the 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 who 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 analyzes the politician's behavior change, before and after the election for such a large period.

Table 3.6: How deputies changed their posting frequency as the 2014 elections approached.

	Political				**Non-Political**			
	RE (%)	LS (%)	NC(%)	**total (%)**	RE (%)	LS (%)	NC (%)	**total (%)**
increased	37 (19%)	33(21%)	44 (32%)	114 (24%)	61 (32%)	40 (26%)	54 (39%)	155 (32%)
maintained	113 (59%)	96(63%)	84 (61%)	293 (60%)	110 (57%)	94 (61%)	75 (54%)	279 (58%)
decreased	42 (22%)	24(16%)	10 (7%)	76 (16%)	21 (11%)	19 (13%)	9 (7%)	49 (10%)

3.3.5 Public engagement on congresspeople political tweets

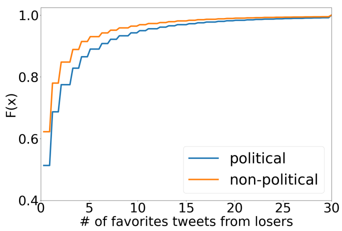
In the last sections, we saw that politicians usually post more *non-political* messages in the election term. From that, a question arises: does the public following of these politicians prefer tweets that are *non-political* or *political*? To answer this question, we collected the number of favorites[5] 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 [Zhao et al., 2015]. Moreover, given that each tweet is labeled as *political* or *non-political*, we can verify which class of tweet is more popular among users.

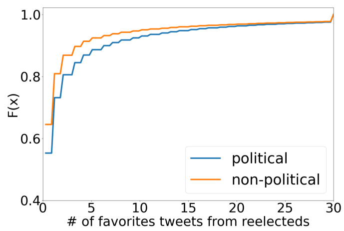
Figure 3.12 shows the cumulative distribution function (CDF) of the number of favorites (3.12a, 3.12b, 3.12c), and retweets (3.12d, 3.12e, 3.12f) received by the messages posted by each class of congresspeople. For better visualization, values greater than 30 were grouped together. Observe in all figures that *political* tweets tend to be favorited more than *non-political* ones. The number of tweets with at most one favorite

[5]Twitter swaps favorites for likes in November 2015. However, our database is from a period prior to this swap.

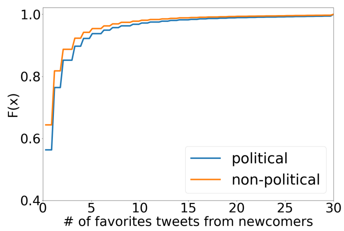
is between 68% and 73% for *political* and between 78% and 82% for *non-political* tweets. Similarly, for retweets, we can observe that *political* messages also have more retweets than *non-political* ones. The number of messages with at most one retweet is between 64% and 78% for *political* and between 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.



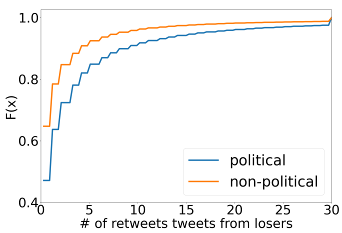
(a) Losers favorite CDF



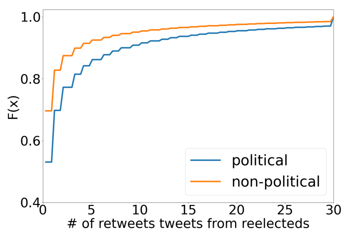
(b) Reelected favorite CDF



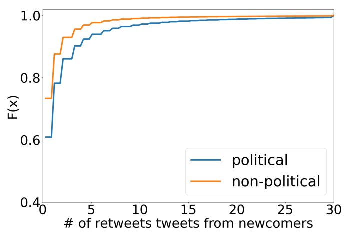
(c) Newcomers favorite CDF



(d) Losers retweet CDF



(e) Reelected rewteet CDF



(f) Newcomers retweet CDF

Figure 3.12: Cumulative Distribution Function of number of political and *non-political* favorited and retweeted tweets per class

Figure 3.13 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 terms 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 3.13 also suggests that *reelected* are the most popular politicians. Also, note that the popularity of *political* tweets is greater than *non-political* ones almost all the time. Similarly, the popularity of tweets of *newcomer* also increases during elections and keeps growing moderately afterward. Additionally, the popularity for both classes is almost the same over the entire period, with *political* tweets reaching their 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 the election term. Again, the popularity of *political* tweets is also always higher during the analyzed period.

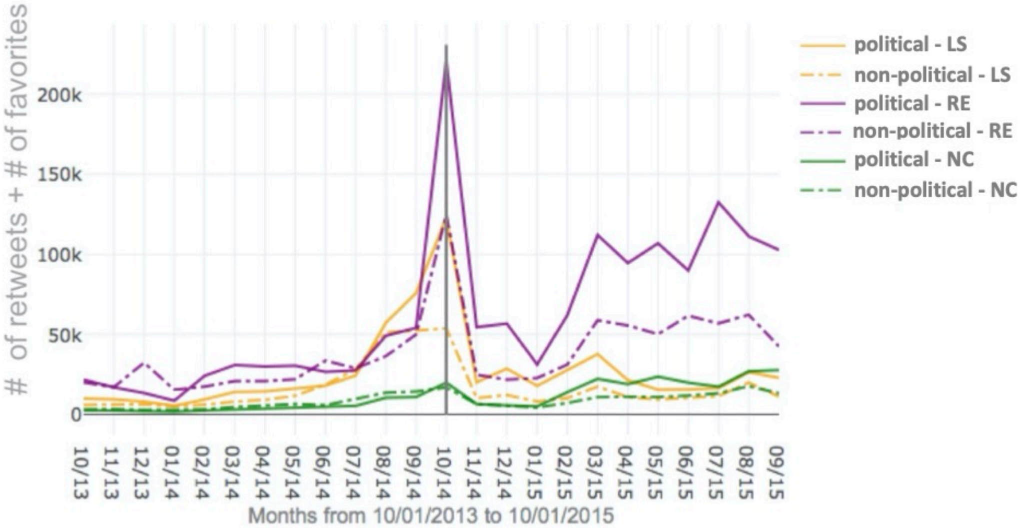


Figure 3.13: Popularity of tweets over time.

3.4 Concluding remarks

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 [Paul et al., 2017; Conover et al., 2011; Gao et al., 2017], 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, which corroborates with Graham et al. [2013], who showed that about 70% of UK politicians tweets were used for broadcasting *political* messages and, conversely, contrasts with other reports [Pal, 2015; Jungherr, 2016; Small, 2010], 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.

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.

Chapter 4

Long Term Characterization of Political Communications

As it is an incremental work throughout the development of the thesis, in our second case study, we carried out a broader and deeper analysis of the communication behavior of Brazilian politicians. We analyzed a set of politicians' tweets from October 2013 to October 2019, taking into account their ideological spectrum: *right*, *center*, and *left*.

4.1 TwLargeDB dataset

To perform the politician communication characterization over time, we used the Twitter API to collect $3,377,744$ public tweets from 914 Brazilian politicians with a valid account on Twitter and who had an office in Congress for at least one year from October 2013 to October 2019. This dataset is available at `https://doi.org/10.6084/m9.figshare.13297889.v2`. This dataset comprises the TwSmallDB and expands it by adding tweets posted by Brazilian deputies between 2013 and 2019. We call this TwLargeDB dataset.

The names of the active congresspeople during this period were retrieved and validated by a researcher in nine different moments between December 2013 and March 2020 using the *Chamber of Deputies Open Data*. The list of the Twitter accounts associated with the congresspeople was collected and validated from the personal profile pages of each congressperson. All politicians were labeled as *left*, *right* or *center*, according to the ideological orientation of their parties, which was collected from the online news web-page *Congresso em Foco* [Sardinha and Costa, 2019]. Each account was manually validated and the collection of their messages was performed in April

2015, April 2019, and April 2020. We prepared the text of the tweets for processing by removing duplicated tweets, punctuation, Portuguese stop words, URLs, and mentions.

We classify the tweets as *political* or *non-political* using our CNN classifier (see section 3.2), trained using the TrainingDB of 2,000 manually labeled tweets, 1,000 *political*, and 1,000 *non-political*. In addition, we also created a ValidationDB dataset containing 3,116 randomly sampled tweets to verify the performance of the classifier over the years. For each year comprising the ValidationDB, we sampled around 500 tweets. 1,104 tweets were randomly sampled from TrainingDB from October 01, 2013, to October 01, 2015. Then, from October 1, 2015, to October 1, 2019, we randomly sampled 2,012 tweets from TwLargeDB.

To qualitatively validate the ValidationDB, we asked two independent researchers to label the sampled 2,012 tweets as *political* and *non-political* according to the Definition 1 of political messages. A third researcher served as a judge in cases of disagreements, so the final label is the majority vote. The researchers are from our lab and they have research topics related to politics. The agreement among the annotators was calculated using the agreement percentage and Cohen's Kappa coefficient (κ) [Landis and Koch, 1977].

Table 4.1 shows the result of the agreement over the ValidationDB. Observe that Researchers 1 and 2 agreed in 92% of the labels, with $\kappa = 0.81$. Researcher 1 and the Majority vote agreed in 97% of the labels, with $\kappa = 0.93$, and Researcher 2 and the Majority vote agreed in 95%, with $\kappa = 0.87$. According to Landis and Koch [1977], these Kappa scores fall into the range of scores referred to as "almost perfect" agreement.

Table 4.1: Cohen's Kappa and Agreement percentage among the researchers.

	Res 1 & 2	**Res 1 & Majority**	**Res 2 & Majority**
Kappa coefficient κ	0.81	0.93	0.87
Agreement %	92%	97%	95%

4.2 Classification Over Time

Because of the natural change in language and terms used over time to designate a *political* message, it is necessary to update [He et al., 2018] or to train a new classifier [Krawczyk et al., 2017] when it loses its efficiency (see Figure 4.3). Li et al. [2018] cite some alternatives to mitigate this problem, however, they are costly or depend on external resources. In this chapter, we propose the LOCPOC (*LOng-term Classification* *of POlitical Communications*), a methodology based on several other techniques: our pre-trained CNN classifier, the DDAL [Costa et al., 2018] method to identify drifts and active learning strategy, and the transfer learning technique to retrain and update the classifier from He et al. [2018].

From the work of Costa et al. [2018], we use the drift identification method and the Active Learning strategy to explicitly detect drifts in an unsupervised way. Instances' true labels are required only after drift detection and new-labeled instances are used only to update the classifier, not to detect drifts [Costa et al., 2018]. Consequently, the computational cost involved in retraining the classifier is negligible and no extensive use of human effort is necessary to label new instances.

Figure 4.1 shows the LOCPOC methodology. First, we use our CNN classifier to label the unlabeled messages as *political* or *non-political*. To use the concept drift approach, messages are labeled by the classifier in fixed-size batches of 500 messages, the best parameter found by Costa et al. [2018]. The batch of labeled messages and their associated class probability (the classifier output) are passed as input to the drift detection algorithm. If no drift is identified, the algorithm returns to the previous phase and the next batch of messages is processed by the classifier. Otherwise, if drift is identified, we start the Active Learning phase, in which the batch of messages is sent to a human specialist who will label the instances whose associated probabilities are less than a λ threshold, which is set to 0.7.

In this thesis, a researcher acted as a domain specialist. In other words, the specialist labels messages from the batch where the classifier is less than 70% certain about their true class. To reduce the specialist's labeling effort, we (randomly) select at most 100 messages whose uncertainty is less than λ. Finally, this batch of (at most) 100 messages is used to retrain and update the classifier, which is done by the transfer learning approach of He et al. [2018] instead of training a new classifier, as suggested by Costa et al. [2018]. This approach uses a diachronic propagation mechanism to incorporate the historical impact into currently learned features. In this phase, the weights of the neural network are updated to learn the new language patterns.

The drifts are identified by an Active Learning strategy of virtual margin that is interpreted as the projection of hyperplanes equidistant to the separating hyperplane based on a user-defined uncertainty threshold (λ). Figure 4.2 shows virtual margins defined for $\lambda = 70\%$. Messages within the gray zone have associated uncertainty below the λ threshold and, if drift occurs, they are sent to be labeled by a human specialist. To identify a drift, the algorithm computes the density (δ) within the virtual margins (gray zone). Then, δ is compared to the historical maximum and minimum density values (δ_{min},δ_{max}) and it replaces the historical values when it is greater than δ_{max}

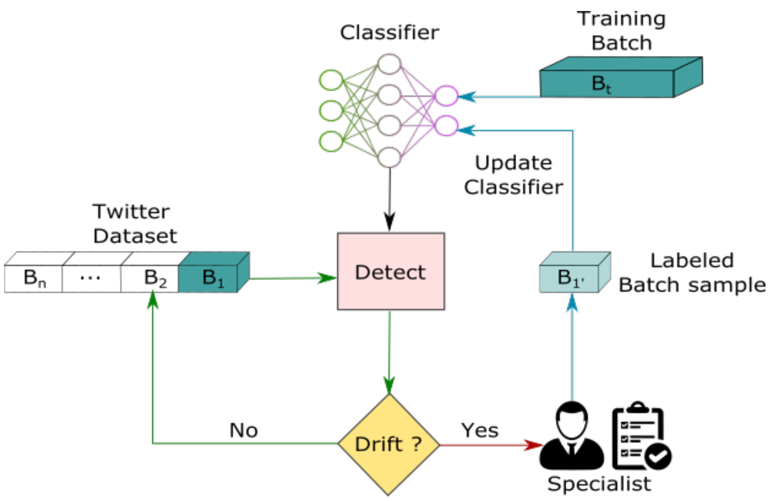


Figure 4.1: LOCPOC Methodology

or lower than δ_{min}. Lastly, if the difference between δ_{max} and δ_{min} is greater than a drift threshold (θ), then drift is signaled and the Active Learning phase is triggered. In this thesis, we use the threshold $\theta = 0.2$. Different values were tested, however, values greater than 0.2 did not identify drift in the data and smaller values are very restrictive in our dataset, which leads to the identification of false drifts in the data, with almost weekly frequency, making the solution infeasible and worse than other literature approaches.

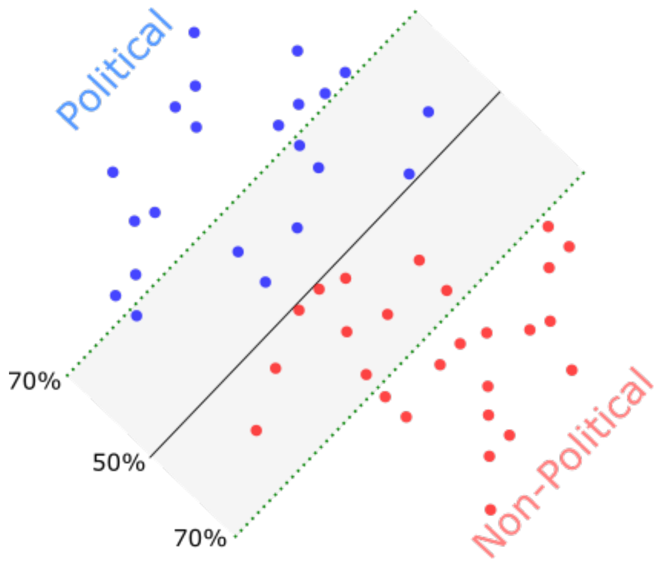


Figure 4.2: Separating hyperplane (solid line) and projection of virtual margins (dotted green lines) with uncertainty threshold set to 70%.

4.3 Results

From the LOCPOC methodology, we classified all tweets from the TwLargeDB over its span of 6 years. We evaluate the classification performance using the ValidationDB. Figure 4.3 shows the comparison between the LOCPOC (green line) and our CNN classifier without updates (red line). We also evaluate the classifier performance using a recent pre-trained Portuguese version of BERT [Souza et al., 2020] (purple line) in which we performed a fine-tuning using the TwSmallDB. We intend with this strategy to assess whether the LOCPOC outperforms a state-of-the-art classifier trained in a large volume of data over time that, theoretically, would not require retraining. The vertical blue lines indicate the dates where drifts were identified. Note that after 2015, i.e., the period in which it was initially trained, the F1 score decreases for all classifiers, suggesting a change in message patterns. More specifically, the F1 score of CNN drops drastically, reaching 0.76 by the end of 2016. Also note that in 2018, the year of national elections, CNN increases the F1 score again, probably because the terms related to elections are present in the messages again, making them similar to the messages of the TrainingDB.

On the other hand, note that LOCPOC has its F1 score almost constant over time, which varies from 0.89 to 0.94. Also, note that BERT performed better than the CNN and slightly worse than LOCPOC. Therefore, it is possible to observe that even though BERT is trained with a large volume of data from different periods, it still suffers from concept drift.

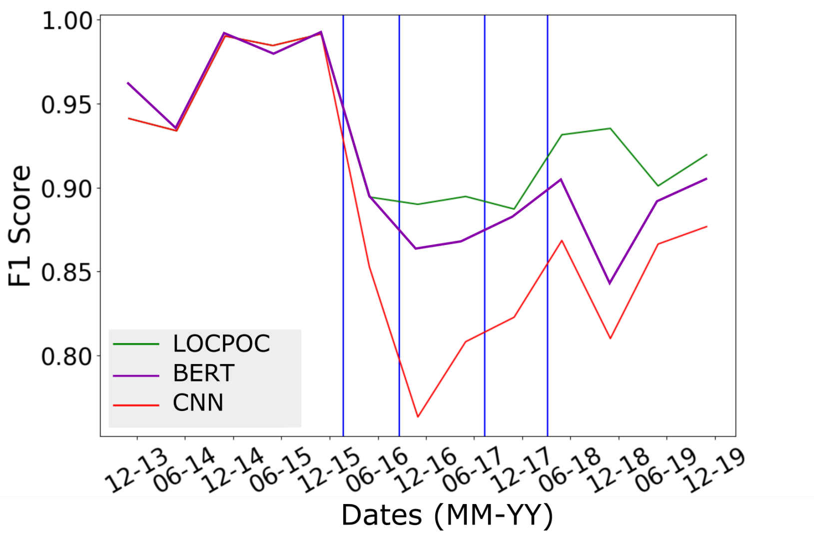


Figure 4.3: Comparison among classifiers over time

In total, four drifts were identified in the TwLargeDB dataset, which is related to events in the Brazilian political scenario, such as corruption scandals and the municipal election in 2016, and parliamentary activities such as the public agenda and campaign.

To exemplify the context drift that occurred in these periods, we selected and translated some messages whose λ thresholds were below 70% and were selected to be labeled by the human specialist. The first two drifts in data occurred during the impeachment of former president Dilma Rousseff, the 2016 Olympic games, and municipal elections in 2016. Most of the messages responsible for decreasing the classifier confidence (and accuracy) were related to these topics:

"PT (Labor Party) BROKEN OUR BRAZIL A series of negative records. This is how the PT is destroying our economy .... #URL"

"We just visited the Augusto Franco Market vendors and we are already going to Santa Maria #hashtag #URL"

"Gold again!!!!!!!! What an Olympic final!! Football yesterday and volleyball today!!! Brazil deserved the joy of an unforgettable Olympics! Bernadinho!!"

The third and fourth drifts occurred in the pre-election period of 2018 and are related to the bill that alters labor laws, the public agenda, and the condemnation of former President Lula for corruption:

"Moro's sentence: 9 and a half years in prison for Lula!"

"In Barra do Garcas visiting the works of the bridge over Araguaia and Garcas rivers. #URL"

"We must react to the end of Labor Law #URL"

We observed that the public agenda and political campaign activities are the most recurring topics in drift occurrences since these activities are confused with everyday activities not related to politics and, consequently, more difficult to classify. In addition to these topics, other less frequent ones, such as "Olympic games", "impeachment" and "corruption" also reduced the classifier's confidence.

To delve deeper into these messages, we evaluated the periods that are difficult to classify according to the LOCPOC, i.e, the periods in which the classifier's confidence decreased over time. For this task, we calculate the module of the difference between the message being *political* and *non-political* provided by the confidence of the classifier output. Then, in Figure 4.4, for each period, we compute the average of that value stratified by *political* and *non-political* messages. It is possible to observe that the confidence follows the F1 score behavior over time (Figure 4.3). There was a drop in the classifier confidence between 2016 and 2018, mainly for *non-political* messages.

When looking at these instances in which the classifier has less confidence, we noticed that, in fact, they are hard to label messages.

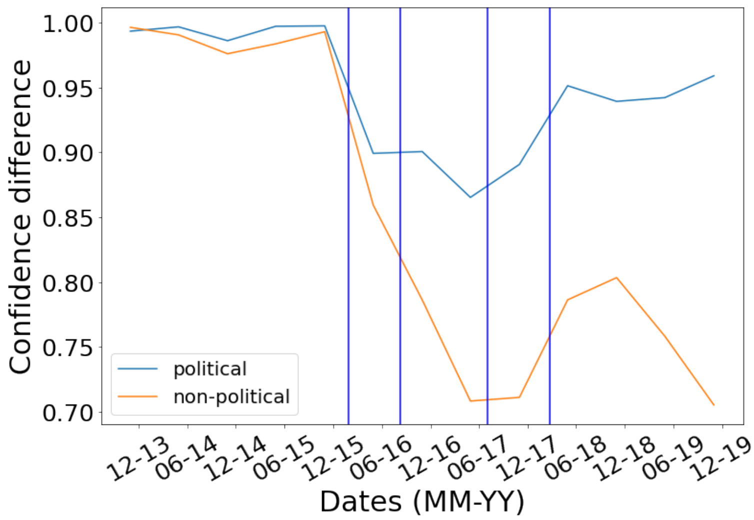


Figure 4.4: LOCPOC output confidence over time

To make it more evident, we identified the messages from ValidationDB in which the human labelers diverged among them and it was necessary for the judge to intervene. Then, we separate these messages into two disjoint sets. One containing messages with the agreement and the other with disagreement among the labelers. Finally, we compute the module of the difference of the classifier's confidence for each of these messages from the two sets. In Figure 4.5, we show the CDFs for the module of the difference between messages with agreement and disagreement. It is possible to notice that the confidence of the classifier is lower for the messages in which the labelers disagreed. For the messages that were in agreement, the classifier confidence was greater than 90% in more than 80% of these messages. Therefore, these results show that the classifier has less confidence in the labeling of messages in which there are divergences, even among humans, and which are, therefore, difficult to label.

Finally, to visualize how topics change over time, we created some word clouds from important periods of Brazilian politics. Figures 4.6a and 4.6b are from the impeachment period of former President Dilma Rousseff. Figures 4.6c and 4.6d are from the municipal election period in 2016. Finally, Figures 4.6e and 4.6f are from the federal election period in 2018.

In the *political* word clouds, it is possible to observe that some words appear in the different periods such as "government", "against" and "policy". However, it is also possible to note that some words characterize well each period, for example, the words "impeachment", "coup", "Dilma", "Mayor", "City", "congressperson" and "Bolsonaro". In

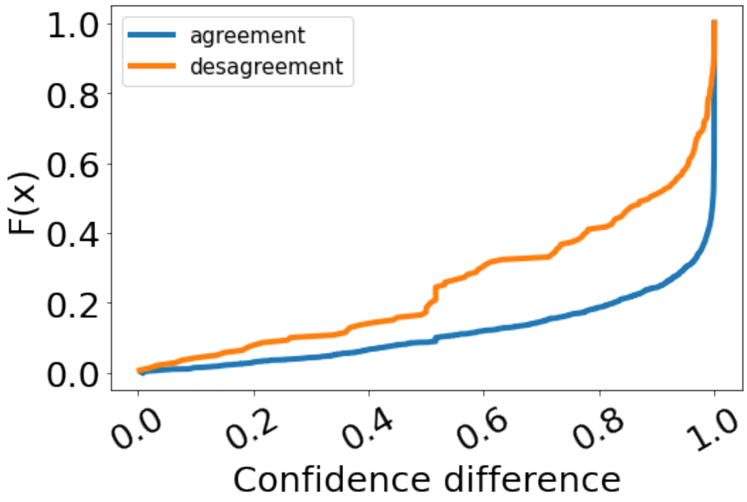


Figure 4.5: CDFs

turn, in the *non-political* word clouds, it is possible to note that words that reinforce the candidate's image and his relationship with voters are present in the three periods such as "God", "congratulations" and "friend". However, in Figures 4.6f and 4.6b, we can see that words related to politics such as "coup", "impeachment", "Bolsonaro", "support" and "Haddad" are present in the *non-political* discourse.

4.3.1 Political communication over time

Given that the LOCPOC obtained an F1 score around 90% in the VALIDATIONDB, the divergence of which is similar to human annotators (see section 3.2.8), we use it to label the TWLARGEDB. From this labeled dataset, we characterize the politicians' communication behavior over time stratified by their ideologies. We start our analyses by counting the total number of *political* and *non-political* messages they posted over time, which is shown in Figure 4.7a. The vertical lines denote important events in the Brazilian political scenario: gray lines mark the 2014 and 2018 national elections, and 2016 municipal elections; the red line marks the beginning of the impeachment process of the former president Dilma Rousseff.

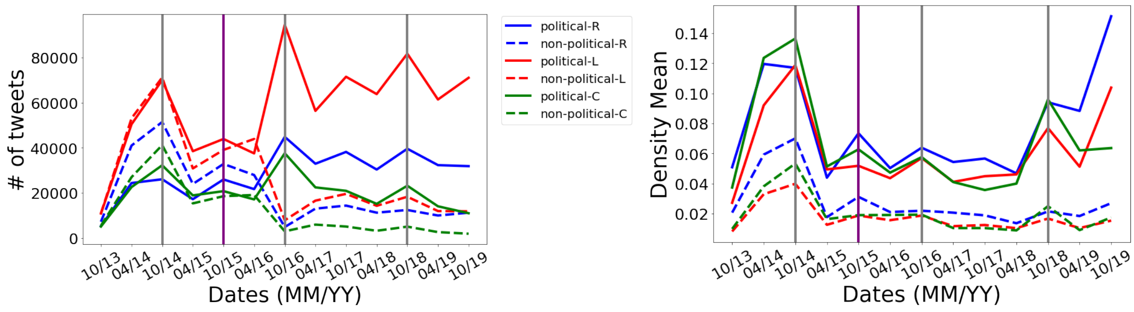
First, note that there is a peak in the number of tweets posted during the 2014 elections (as shown in section 3.3.3). However, *left* deputies have a similar number of *political* and *non-political* tweets, while the *right* and the *center* posted more *non-political* tweets. Second, observe that there are two important changes in the deputies' communication behavior during the period comprising the impeachment process and the 2016 municipal elections. First, there is a greater spike in the number of *political* posts, mainly by the *left*. Second, there is an inversion in the type of message disclosed,

(a) *political* tweets of impeachment (b) *non-political* tweets of impeachment
(c) *political* tweets of 2016 election (d) *non-political* tweets of 2016 election
(e) *political* tweets of 2018 election (f) *non-political* tweets of 2018 election

Figure 4.6: Word cloud of *political* and *non-political* tweets in different periods

i.e., deputies mostly from the *right* and *left* drastically reduced the number of *non-political* tweets and started to publish more about *political* subjects. Moreover, observe that the total number of the *center* tweets decreased over time, while the volume of the *left* and *right* messages remained practically stationary. Finally, note that the total number of the *left* messages is much higher than the *right* and *center*. This finding is in line with Amaral and Pinho [2017], who showed that *left* politicians are more active on Twitter.

However, we suspect that this significant difference in volume is mostly due to a small group of *left* politicians who post much more than the average. In Figure 4.8, we show the total number of tweets as we remove the most active politician of each spectrum. Observe how the *left* curve decays similarly to the *center* curve and much faster than the *right* one. This suggests that the activity distribution among *left* and *center* politicians is similar, with a few politicians being responsible for most of the



(a) Distribution of tweets by ideology over time. (b) Density mean distribution by ideology

Figure 4.7: Distribution of tweets over time.

activity, whereas, in the *right*, the activity is more evenly distributed.

We further investigate this by transforming the time series of each deputy into a density distribution over time, i.e., we divide the number of posts in each semester by the total number of posts of the politician. Figure 4.7b shows the mean of these transformed curves, which are very different from the ones shown in Figure 4.7a. First, note that all ideological spectra behave similarly: their posting frequencies increase significantly during executive elections, especially *center* politicians. Also, just before and after the 2018 elections, the difference between the *right* and the *left* curve increased significantly, as *right* politicians became more active and more numerous. These results are in agreement with Amaral and Pinho [2017], who showed that there are some "celebrity" politicians on Twitter, and with Recuero et al. [2020], who showed that the 2018 elections were marked by the massive use of social media by the *right* politicians and its supporters, which increased the number of right-wing deputies in the chamber.

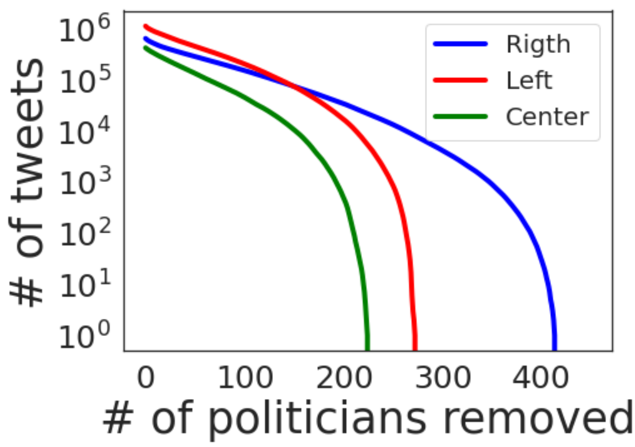
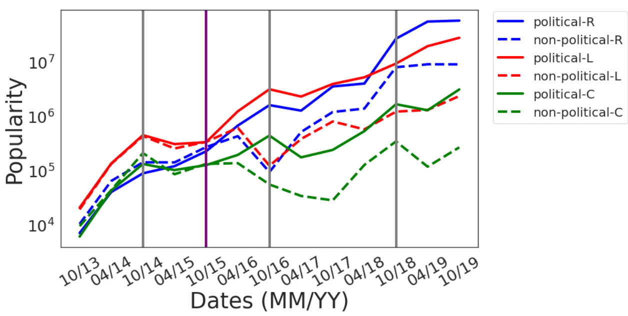


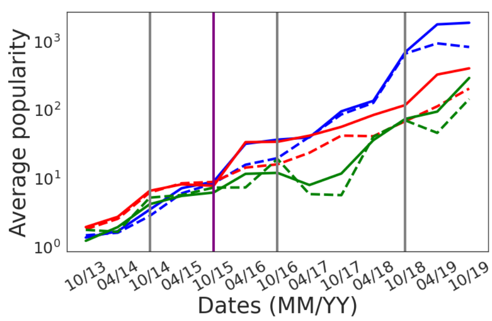
Figure 4.8: The number of tweets removing top-ranked politicians.

4.3.2 Popularity over time

Here we analyze the popularity of tweets by summing their received number of retweets and favorites. Figure 4.9a shows the popularity of the tweets over time, stratified by political ideology. First, note that the popularity of tweets posted by politicians has been steadily increasing since 2013 across all ideological spectra. Second, observe how *political* tweets are far more popular than *non-political*. Finally, note the popularity of posts from the *right* politicians, which grew more than from the other ideologies, both for *political* and *non-political* tweets, surpassing the popularity of the *left* just before the 2018 elections.



(a) Marginal popularity count.



(b) Average popularity.

Figure 4.9: Popularity over time, measured by the sum of retweets and favorites received by the tweets.

To estimate the expected popularity of a tweet, we divided the total number of favorites and retweets received by each ideology (e.g. *right*) by the total number of tweets they posted. Figure 4.9b shows this average popularity over time, and two striking findings can be taken from this figure. First, note that since 2017 the average popularity of *right* tweets is higher than the other ideologies, even for *non-political* ones. Second, and most surprisingly, the average popularity of *political* and *non-political* tweets for the *right* is roughly the same from April 2017 to October 2018, the month of the presidential election. This behavior is more evident in Figure 4.10 which shows the popularity of the complementary cumulative distribution function (CCDF) of tweets from April 2017 to October 2018. Note how the *right*'s popularity curves are different from other ideologies. Both *political* and *non-political* tweets distributions are practically the same, with average popularity much higher than the other ideologies. In other words, it seems that no matter the content of the tweets posted by the *right* during this period, their expected popularity is the same. This can be partially explained by reports of the massive use of robots by the *right*, as suggested by Silva and Silva [2019].

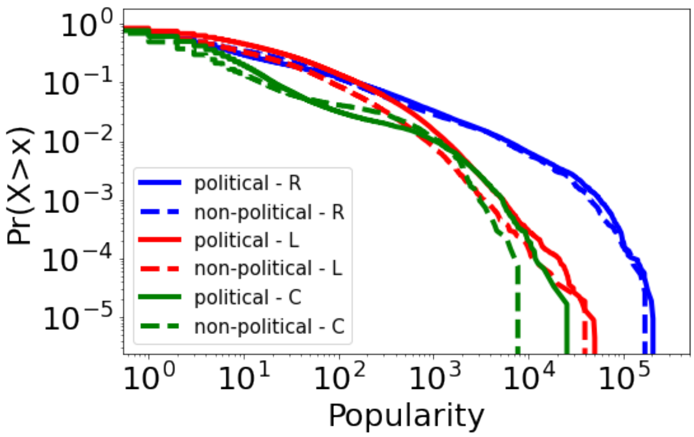


Figure 4.10: Popularity CCDF

4.3.3 Individual communication behavior

To quantify the individual communication behavior of politicians over time, we use the *Political Communication Index* (PCI) (see section 3.3.1), the ratio between their number of *political* tweets and their total number of tweets.

Figure 4.11 shows the scatter plot of the PCI and the number of tweets for all politicians grouped by year and stratified by the political ideologies. First, note that it is not possible to differentiate the politicians by their ideology. In general, most politicians posted between 10^2 and 10^4 tweets in this period. However, there are politicians who posted more than average, exceeding the number of 10^4 messages (marked by dotted gray line), mainly *left* politicians (fact noticed in Figure 4.7b). Moreover, the majority of politicians have a high PCI, greater than 50%, i.e., they decide to disclose their political views very often. This is in line with Graham et al. [2013], who showed that about 70% of UK politicians' tweets were used for broadcasting *political* messages and, contrasts with Pal [2015] who showed a predominance of *non-political* messages in politicians' communications.

4.3.4 Posting frequency over time

Although most politicians have a high PCI, the frequency with which the messages were posted over time is not evident. To analyze this behavior, we computed Shannon's entropy [MacKay and Mac Kay, 2003] for each deputy from the previously calculated densities. Let N_{sem} be the total of 12 semesters comprising the TWLARGEDB. Thus, for each deputy, if he/she uniformly posts the same number of messages each semester, then the entropy will be maximum $H_{max} = \log_2 N_{sem} = 3.58$. Conversely, if the politician posts all their messages in just one period, then the entropy will be equal to

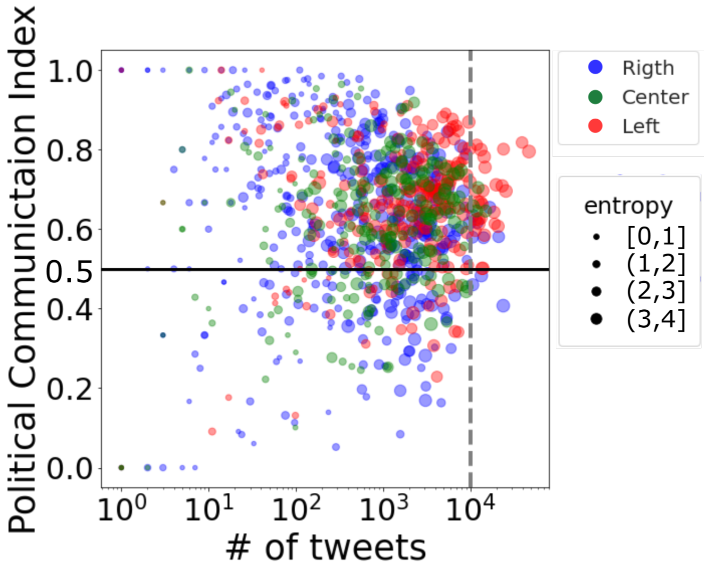
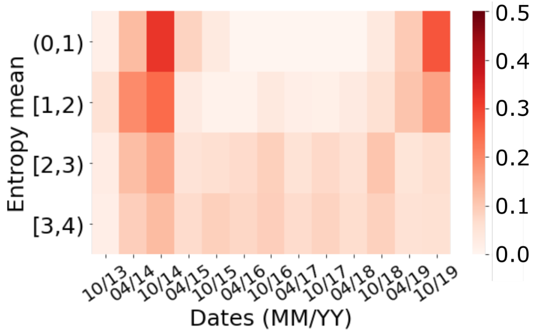


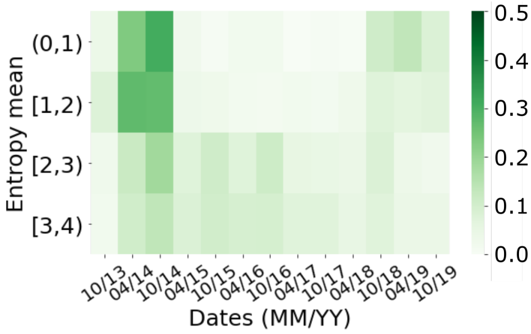
Figure 4.11: Scatter plot for the PCI and the number of tweets for *right*, *left* and *center*

$\log_2 1 = 0$. In other cases, deputies diversify their posting frequencies over time and the entropy varies between 0 and maximum entropy $H_{max} = 3.58$.

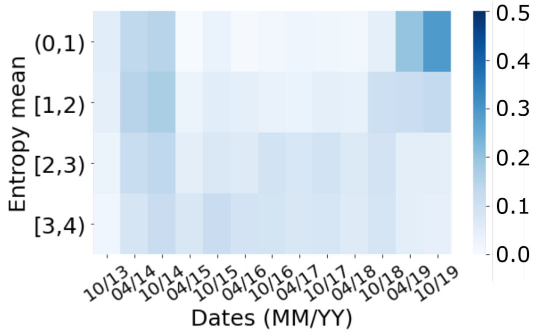
From the entropy calculated for each deputy, we divide it into ranges of values for better visualization. Deputies who chose a maximum of 2 semesters to post their messages, out of the 12 possible, have entropy values between [0,1]. Deputies who chose to post between 2 and 4 semesters have entropy with values between (1,2]. Similarly, deputies who posted between 4 and 8 semesters have entropy with values between (2,4]. Finally, deputies who posted regularly from 9 to 12 semesters have entropy values between (2,3.58]. In Figure 4.11, it is possible to observe that the majority of deputies have high entropy (larger dots), i.e., their messages tend to be uniformly distributed over time. However, less active deputies tend to have low entropy (smaller dots), i.e., they have the distribution of messages concentrated in a few periods. To visualize this behavior over time, we show in Figures 4.12a,4.12b,4.12c, for each entropy range, the average of the deputies' densities stratified by ideology, *left*, *center* and *right*, respectively. Naturally, observe that deputies with high entropy tend to post uniformly over time and the light colors in all semesters represent this uniformity. On the other hand, deputies with low entropy tend to concentrate their posts in particular semesters, which are represented by the darker colors in a few periods. For the *center*, communications of low entropy deputies were concentrated in the 2014 elections, which is also where some low entropy deputies of the *left* concentrate their communications. What is curious is the high concentration of low-entropy communications in 2019, for both the *left* and the *right* (and not for the *center*). Note that low entropy deputies from the *right* are mostly concentrated in this year. In addition, given that 2019 is not an election year, this suggests an increase in polarization and the rise of once silent deputies. In fact, for the 273 *left* politicians of our dataset, 24 have more than 50% of their communications in 2019. This is even more striking for the *right*: 85 out of 440 *right* politicians have more than 50% of their communications in 2019.



(a) Entropy Left



(b) Entropy Center



(c) Entropy Right

Figure 4.12: Entropy over time.

4.4 Concluding remarks

In conformity to prior research [Oliveira et al., 2018; Badawy et al., 2018; Caetano et al., 2018; Bagavathi et al., 2019] that analyzed communications of politicians on social media, in this section, we use the LOCPOC to characterize the communication of 914 Brazilian politicians over years in terms of the amount of *political* and *non-political* messages they post and that is robust to concept drifts. Using the proposed methodology in the TwLargeDB, which contains more than 3M tweets, we found that the data drifts occurred during important events of Brazilian politics, such as the impeachment of Dilma Rousseff and elections, and that the popularity of political communications increased steadily since 2013. Our analyses also revealed that, although left-wing politicians post, in general, more on social media, this number is distorted by a small group of very active accounts.

In particular, we showed how the *right* rose just before the 2018 elections, both in terms of the number of messages and public engagement. In the real world, the right also grew, while the left remained stable and the center decreased. In 2014 the *left* had 27% of the congress and in 2018 it remained. However, while the *right* increased from 46% in 2014 to 59% in 2018, the size of the *center* decreased from 27% to 14% in 2018. Regarding the total number of votes received, while the *left* held 20% of the vote in the two elections, the *right* increased from 51% in 2014 to 63% in 2018, and the *center* decreased from 29% to 17% in 2018. However, the increase in the number of right-wing politicians does not explain the activity burst in their accounts. Besides showing an increase of the *right*, the decrease of the *center* also points to an increase in polarization, which was also shown by Recuero et al. [2020]. Interestingly, this rise of the *right* was not exclusive to Brazil, other countries in South America such as Argentina, Colombia, and Chile also observed this same phenomenon [Ospina-Valencia, 2018].

Finally, our analyses also revealed a curious behavior regarding the popularity of right-wing tweets: between April 2017 and October 2018, *political* and *non-political* communications were equally popular among the public, which can be explained by the alleged use of robots to indiscriminately promote these tweets [Silva and Silva, 2019].

Chapter 5

Discussion

In the first two chapters, we characterize the communication behavior of the Brazilian congresspeople over time. In the first characterization, we analyzed the communication of politicians one year before and one year after the 2014 elections, stratifying the deputies for their electoral success. In the second characterization, we extended the period of analysis to 6 years and stratified each deputy by their ideological spectrum.

From these analyses, it was possible to note a slightly distinguishable behavior among and within the congresspeople classes, which responds to the research question **R1**. There is an increase in the number of tweets, both *political* and *non-political* over time, however, there was atypical behavior close to the election period. This behavior is partially corroborated by Lietz et al. [2014]; Wong et al. [2013]. During the 2014 elections, 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 [Jackson and Lilleker, 2011; Graham et al., 2013; Jungherr, 2016; Mainwaring, 2001].

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 [2017], 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 [Jackson and Lilleker, 2011], US Congress [Golbeck et al., 2010], candidates in Spain [Grimaldi, 2019], mayors in Turkey [Sobaci and Karkin, 2013] and party leaders in Canada [Small, 2010].

We also analyzed the popularity of the tweets through the number of likes (or favorites) and retweets they receive from 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 responds to the research question **R2** and it is antagonistic to Lee and Shin [2014], 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 [Amaral and Pinho, 2016; DiGrazia et al., 2013], and with the literature on the political behavior of Brazilian parliamentarians [Mainwaring, 2001; Samuels, 1999; Samuels and Zucco, 2014].

Regarding the use of social media by Brazilians, according to Machado et al. [2018], 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 [Recuero et al., 2020] and by the attempt to influence voters to change the outcome of the elections [Machado et al., 2018; Marques and MontÁlverne, 2016]. 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 [Cha et al., 2010]. Therefore, we show that during the 2014 elections congresspeople devote much of their communications to propagate *non-political* messages, which can be an erroneous practice. However, we observe an important change in the deputies' communication behavior right after the impeachment process started. There is an inversion in the type of message disclosed, i.e., deputies drastically reduced the number of *non-political* tweets and started to publish more about *political* subjects. Our analyses also revealed that people who follow the politicians of our data set are more likely to prefer *political* tweets. These results respond to the research questions **R0** and **R3**. Also, this is in line with the work of Hwang [2013], who showed that voters mostly expect that politicians actively share their candid opinions through the open public sphere of Twitter.

5.1 Limitations

Despite the importance of this work, we acknowledge that our thesis has several limitations. First, although the methodology can be replicated in other contexts, the results and conclusions found in this thesis are specific to the Brazilian scenario and are not necessarily valid for other countries and periods. We only tested machine learning-based models trained on messages in Portuguese, assessing the accuracy of such techniques on messages in other languages and testing how well these techniques are adapting to future elections remains an open question. Second, when labeling a message as *political* or not, we analyze only the content of the message and do not consider its subjective information. For instance, the following message apparently has no *political* content in its text: "Be proud of the love of your life!". However, it was posted by a deputy gay rights activist, kissing his partner as a form of protest. Therefore, there is a *political* intention in the post, but not in the message content. Third, we do not investigate an alternative explanation for the characterization 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.

5.2 Implications

Despite these limitations, our work was used in a different context than the one presented in this thesis. Silva et al. [2020] used the proposed CNN architecture to investigate the use of *political* Ads on Facebook during the 2018 election campaign.

In an attempt to minimize the external influence, disinformation campaigns, and the misuse of social media during the elections, on May 24, 2018, Facebook changed its ToS policy to allow the launching of *political* ads only by advertisers that reside in the same country with the people targeted (this requirement does not apply to *non-political* ads) [Leathern, 2018]. Also, Brazilian authorities demanded that political figures that are advertising on Facebook *political* content along the electoral period, an established period near the elections, need to give information about their national identification numbers, namely CPF, for individuals, and CNPJ, for companies. Facebook responded by creating an interface that allows advertisers of *political* content to include disclosure information in their ads related to elections and also their CNPJs or CPFs.

Concerned by the eminent high potential misuse of Facebook ads and imminent risks to Brazilian electoral laws, Silva et al. [2020] developed the monitor of *political* ads on Facebook installed by more than 2000 users. From that application, they collected the AdCollector dataset containing 239k ads from 40k advertisers along the period of March 14, 2018, to October 28, 2018. From this dataset, they were able to exploit the ads self-declared as *political* from compliant advertisers to build a machine learning-based model that detected other similar ads coming from advertisers that do not comply with Facebook's Term of Service or electoral laws. They found that a small fraction of advertisers in their dataset have the right disclaimer stipulated in the Brazilian election law but they did not declare their ads as *political* to Facebook, hence, they do not appear in the available official Facebook Ad Library.

Therefore, while the measures taken by Facebook to diminish the elections threats were welcomed, many people including researchers, journalists, and organizations pointed out that they are not sufficient. *First*, advertisers have to declare themselves, on a voluntary basis, whether they are sending *political* ads. This is problematic because dishonest political parties and presidential candidates can avoid scrutiny of their ad messages by not declaring them as *political*. *Second*, beyond public opinion manipulation and spread of fake news, the Facebook ads platform can also be used for *slush funds* [Campos et al., 2018]. Brazilian electoral law states that companies are prohibited to make donations to any political party or candidate during the election period. Currently, dishonest companies can spend an unlimited amount of undeclared money in favor of a political agenda through the Facebook ads platform [TSE - Electoral Court, 2017].

5.3 Future directions

In the previous section, we show how our CNN architecture was used in a different context to identify Facebook Ads in disagreement with Brazilian electoral law and Facebook policies. However, this thesis opens up several possibilities for future research. For example, there is difficult for big tech companies like Facebook, Twitter, and Google to adapt to the laws of each country. Therefore, a methodology like ours can assist in organizing information and facilitate the application of the laws of each country.

Another possible extension of our work is to deepen into the textual analysis to identify what the deputies intend when posting a message. In the example shown in Section 5.1, the content of the message is *non-political*, but the deputy's intention when posting the message was of a *political* nature, as evidenced by the image attached to the message. Therefore, an extension of our textual classifier could be developed to incorporate other media formats, such as images, videos, and voice. Thus, the analysis could be enriched and new insights brought about the real intention of the message's author. Also, these other types of media formats are still little explored in political analysis literature.

Still, with regard to textual content, our proposed methodology can be extended to the analysis of topics [Shi et al., 2018]. Since our methodology classifies each message as *political* or *non-political*, one could use these two disjoint sets and analyze the topics discussed by deputies over time. Therefore, several analyzes could be carried out, for example, to analyze the recurring topics, whether the distribution of topics changes over time, whether there are deputies who prefer some specific topics over others, and whether there are clusters of deputies who post about of the same subjects.

Regarding the political message classifier, a possible unfolding of this thesis could be an improvement of the concept drift and active learning techniques and the adoption of more recent NLP classification models, such as T5 and GPT-3, that use Transformers and Attention Mechanisms. Also, our methodology could be extended to other languages and to other contexts, for example, to evaluate the tweets posted by politicians in the U.S. election campaign, especially the ones posted by Donald Trump, which were very popular and controversial. Additionally, one could deepen into the results obtained in this thesis during the post-election period of 2018 that suggests a polarization between *right* and *left* discuss and the possible use of robots to boost the popularity of tweets.

5.4 Conclusion

The classification of short texts is not a trivial task due to the amount of noise in the communication and the limited number of characters that make it difficult to infer the context of the message. There are several attempts in the literature to classify this type of communication, however, they are not effective or cannot be applied on a large scale. Therefore, our study is an important step in social computing literature, providing a supervised machine learning classifier that labels all textual messages from different social media platforms as *political* or *non-political*. Moreover, the classifier is robust to concept drifts, where topic changes are identified over the years by an unsupervised drift detection. From that, we parsimoniously characterized the social media communications posted by Brazilian parliamentarians over time. In other words, the *political* message classifier enabled us to address two case studies with different natures and characteristics.

More specifically, it enabled us to analyze how politicians present themselves in the digital environment and how the public reacts to them. We investigated 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 over time, especially during important political events. Moreover, we showed that *political* and *non-political* tweets become significantly more popular over time, especially the *political* ones, which are much more preferred than the *non-political*. Finally, we show that the *right* strongly increased its participation over time, which was reflected in the votes and the number of elected politicians.

Our thesis opens a wide range of possibilities for new applications and research. Twitter recently announced that it will no longer allow *political* ads on its website [Feiner, 2019]. We hope our findings and all the real-world experience of deploying a real system along the 2018 Brazilian elections will inform debates around public policies that regulate political advertising on the Internet. If a system like ours is implemented on a widespread scale, political campaigns might adopt adversarial strategies that change their marketing strategies in order to exploit our false-negative rate. To assist in this task, we developed a tool that uses our classifier and enables the automatic labeling of *political* and *non-political* messages. The tool receives as input a message or a file with several messages and returns the label and the confidence of the classifier for each message. Although there are some efforts to make the elections in Brazil transparent [voz ativa, 2019; Eleições Sem Fake, 2019], having multiple independent systems would make the monitoring of *political* messages more robust to attackers. We hope our effort will inspire other initiatives around the world. Our work not only highlights the importance of independent auditing platforms for *political* ads but also provides all necessary framework to make it feasible as our code is open source[1].

[1] https://github.com/lucasant10/

Bibliography

Amaral, M. S. and Pinho, J. A. G. d. (2016). Tuitando por votos: Congressistas brasileiros e o uso do twitter nas eleições de 2014. In *Proceedings of the XL Encontro da Anpad*, pages 1--19. Anpad.

Amaral, M. S. and Pinho, J. A. G. d. (2017). Ideologias partidárias em 140 caracteres: uso do twitter pelos parlamentares brasileiros. *Revista de Administração Pública*, 51(6):1041--1057.

Badawy, A., Ferrara, E., and Lerman, K. (2018). Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign. *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, pages 258--265.

Bagavathi, A., Bashiri, P., Reid, S., Phillips, M., and Krishnan, S. (2019). Examining untempered social media: analyzing cascades of polarized conversations. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 625--632.

Barberá, P. and Zeitzoff, T. (2018). The new public address system: Why do world leaders adopt social media? *International Studies Quarterly*, 62(1):121--130. ISSN 14682478.

Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. (2003). A Neural Probabilistic Language Model. *J. Mach. Learn. Res.*, 3:1137--1155. ISSN 1532-4435.

Bovet, A., Morone, F., and Makse, H. A. (2018). Validation of Twitter opinion trends with national polling aggregates: Hillary Clinton vs Donald Trump. *Scientific Reports*, 8(1):1--16. ISSN 20452322.

Bracciale, R. and Martella, A. (2017). Define the populist political communication style: the case of italian political leaders on twitter. *Information, Communication & Society*, 20(9):1310--1329.

Caetano, J. A., Almeida, J., and Marques-Neto, H. T. (2018). Characterizing politically engaged users' behavior during the 2016 us presidential campaign. *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, pages 523--530.

Camacho, C. (2019). Cnn for text classification.

Campos, R. R., Maranhão, J., and Benevenuto, F. (2018). Fake news and the chronicle of slush fund announced.

Cappella, J. N. (2017). Vectors into the Future of Mass and Interpersonal Communication Research: Big Data, Social Media, and Computational Social Science. *Human Communication Research*, 43(4):545--558. ISSN 14682958.

Carey, J. M. (2007). Competing Principals, Political Institutions, and Party Unity in Legislative Voting. *American Journal of Political Science*, 51(1):92--107. ISSN 0092-5853.

Cha, M., Haddadi, H., Benevenuto, F., and Gummadi, P. K. (2010). Measuring User Influence in Twitter: The Million Follower Fallacy. *ICWSM*, 10(10-17):30.

Chang, E. C. C. and Golden, M. A. (2007). Electoral Systems, District Magnitude and Corruption. *British Journal of Political Science*, 37(1):115--137. ISSN 0007-1234.

Conover, M., Ratkiewicz, J., and Francisco, M. (2011). Political polarization on twitter. *Icwsm*, 133(26):89--96. ISSN 15205126.

Costa, A. F. J., Albuquerque, R. A. S., and Santos, E. M. D. (2018). A Drift Detection Method Based on Active Learning. *Proceedings of the International Joint Conference on Neural Networks*, 2018-July.

Council on Foreign Relations (2018). WhatsApp's Influence in the Brazilian Election and How It Helped Jair Bolsonaro Win.

Deriu, J., Luca, V. D., Müller, S., Hofmann, T., and Jaggi, M. (2017). Leveraging Large Amounts of Weakly Supervised Data for Multi-Language Sentiment Classification. pages 1045--1052.

DiGrazia, J., McKelvey, K., Bollen, J., and Rojas, F. (2013). More tweets, more votes: Social media as a quantitative indicator of political behavior. *PLoS ONE*, 8(11):1--5. ISSN 19326203.

Dubois, E. and Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information Communication and Society*, 21(5):729--745. ISSN 14684462.

Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research*, 12:2121--2159. ISSN 15324435.

Eleições Sem Fake (2019). Eleições sem fake. `http://www.eleicoessemfake.dcc.ufmg.br`. Accessed: 2019-12-19.

Enli, G. (2017). Twitter as arena for the authentic outsider: exploring the social media campaigns of trump and clinton in the 2016 us presidential election. *European Journal of Communication*, 32(1):50–61.

Etudo, U., Yoon, V. Y., and Yaraghi, N. (2019). From facebook to the streets: Russian troll ads and black lives matter protests. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Feiner, L. (2019). Twitter Bans poltical ads – CNBC.

Forelle, M. C., Howard, P. N., Monroy-Hernandez, A., and Savage, S. (2015). Political Bots and the Manipulation of Public Opinion in Venezuela. *SSRN Electronic Journal*, pages 1--8. ISSN 1556-5068.

Gainous, J. and Wagner, K. M. (2014). *Tweeting to power: The social media revolution in American politics*. Oxford University Press.

Gao, L., Kuppersmith, A., and Huang, R. (2017). Recognizing Explicit and Implicit Hate Speech Using a Weakly Supervised Two-path Bootstrapping Approach.

Garimella, K., Morales, G. D. F., Gionis, A., and Mathioudakis, M. (2018). Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 2:913--922. ISSN 16130073.

Glassman, M. E., Straus, J. R., and Shogan, C. J. (2010). Social networking and constituent communications: member use of Twitter during a two-month period in the 111th Congress. *Congressional Research Service*, 41066.

Golbeck, J., Grimes, J. M., and Rogers, A. (2010). Twitter use by the US Congress. *Journal of the Association for Information Science and Technology*, 61(8):1612--1621.

Goldberg, Y. (2016). A Primer on Neural Network Models for Natural Language Processing. *J. Artif. Int. Res.*, 57(1):345--420. ISSN 1076-9757.

Goldberg, Y. (2017). Neural Network Methods for Natural Language Processing. *Synthesis Lectures on Human Language Technologies*, 10(1):1--309. ISSN 1947-4040.

Gonawela, A., Pal, J., Thawani, U., van der Vlugt, E., Out, W., and Chandra, P. (2018). *Speaking their Mind: Populist Style and Antagonistic Messaging in the Tweets of Donald Trump, Narendra Modi, Nigel Farage, and Geert Wilders*, volume 27. Computer Supported Cooperative Work (CSCW). ISBN 1060601893.

Graham, T., Broersma, M., and Hazelhoff, K. (2013). Closing the gap? Twitter as an instrument for connected representation. *Routledge Research in Political Communication.*

Grant, W. J., Moon, B., and Grant, J. B. (2010). Digital dialogue? australian politicians' use of the social network tool twitter. *Australian Journal of Political Science*, 45(4):579--604. ISSN 10361146.

Grimaldi, D. (2019). Can we analyse political discourse using twitter? evidence from spanish 2019 presidential election. *Social Network Analysis and Mining*, 9(1):49.

Hampton, K. N., Shin, I., and Lu, W. (2017). Social media and political discussion: when online presence silences offline conversation. *Information Communication and Society*, 20(7):1090--1107. ISSN 14684462.

Harris, Z. S. (1954). Distributional structure. *Word*, 10:146--162.

Hartmann, N., Fonseca, E., Shulby, C., Treviso, M., Rodrigues, J., and Aluisio, S. (2017). Portuguese Word Embeddings: Evaluating on Word Analogies and Natural Language Tasks. Number Section 3.

He, Y., Li, J., Song, Y., He, M., and Peng, H. (2018). Time-evolving text classification with deep neural networks. *IJCAI International Joint Conference on Artificial Intelligence*, 2018-July:2241--2247. ISSN 10450823.

Heller, W. B. and Mershon, C. (2005). Party Switching in the Italian Chamber of Deputies, 1996–2001. *The Journal of Politics*, 67(2):536--559. ISSN 0022-3816.

Hemphill, L., Otterbacher, J., and Shapiro, M. (2013). What's Congress Doing on Twitter? In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 877--886, New York, NY, USA. ACM.

Hemphill, L. and Roback, A. J. (2014). Tweet acts: How constituents lobby congress via twitter. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, pages 1200--1210, New York, NY, USA. ACM.

Hicken, A. and Simmons, J. W. (2008). The Personal Vote and the Efficacy of Education Spending. *American Journal of Political Science*, 52(1):109--124. ISSN 00925853.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. volume 9, pages 1735--1780, Cambridge, MA, USA. MIT Press. ISSN 0899-7667.

Howard, P. N., Savage, S., Saviaga, C. F., Toxtli, C., and Monroy-Hemández, A. (2016). Social Media, Civic Engagement, and the Slacktivism Hypothesis: Lessons From Mexico'S "El Bronco". *Journal of International Affairs*, 70(1):55--73. ISSN 0022197X.

Hughes, S. G.-F., Allbright-Hannah, K., Goodstein, S., Grove, S., Zuckerberg, R., Sladden, C., and Bohnet, B. (2010). Obama and the power of social media and technology. *The European Business Review (May-June 2010)*, pages 16--21.

Hwang, S. (2013). The Effect of Twitter Use on Politicians' Credibility and Attitudes toward Politicians. *Journal of Public Relations Research*, 25(3):246--258. ISSN 1062726X.

Jackson, N. and Lilleker, D. (2011). Microblogging, constituency service and impression management: UK MPs and the use of Twitter. *The Journal of Legislative Studies*, 17(1):86--105.

Jolliffe, I. T. (1986). Principal component analysis and factor analysis. In *Principal component analysis*, pages 115--128. Springer.

Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. ISSN 10450823.

Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 13(1):72--91. ISSN 1933-1681.

Karlsen, R. and Enjolras, B. (2016). Styles of social media campaigning and influence in a hybrid political communication system: Linking candidate survey data with twitter data. *The International Journal of Press/Politics*, 21(3):338--357.

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. pages 1746--1751. ISSN 10709908.

Kim, Y. M., Hsu, J., Neiman, D., Kou, C., Bankston, L., Kim, S. Y., Heinrich, R., Baragwanath, R., and Raskutti, G. (2018). The stealth media? groups and targets behind divisive issue campaigns on facebook. *Political Communication*, 35(4):515--541.

Klinger, U. (2013). MASTERING THE ART OF SOCIAL MEDIA: Swiss parties, the 2011 national election and digital challenges. *Information Communication and Society*, 16(5):717--736. ISSN 1369118X.

Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., and Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion*, 37:132--156. ISSN 15662535.

Landis, J. R. and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159. ISSN 0006341X.

Leathern, R. (2018). Shining a Light on Ads With Political Content.

Lee, E.-J. and Shin, S. Y. (2012). Are They Talking to Me? Cognitive and Affective Effects of Interactivity in Politicians' Twitter Communication. *Cyberpsychology, Behavior, and Social Networking*, 15(10):515--520. ISSN 2152-2715.

Lee, E. J. and Shin, S. Y. (2014). When the Medium Is the Message: How Transportability Moderates the Effects of Politicians' Twitter Communication. *Communication Research*, 41(8):1088--1110. ISSN 15523810.

Li, P., He, L., Wang, H., Hu, X., Zhang, Y., Li, L., and Wu, X. (2018). Learning from Short Text Streams with Topic Drifts. *IEEE Transactions on Cybernetics*, 48(9):2697--2711. ISSN 21682267.

Library, C. (2019). Jair Bolsonaro Fast Facts – CNN.

Lietz, H., Wagner, C., Bleier, A., and Strohmaier, M. (2014). When Politicians Talk: Assessing Online Conversational Practices of Political Parties on Twitter. *Eighth International AAAI Conference on Weblogs and Social Media*, pages 285--294.

Machado, C., Kira, B., Hirsch, G., Marchal, N., Kollanyi, B., Howard, P. N., Lederer, T., and Barash, V. (2018). News and political information consumption in brazil: Mapping the first round of the 2018 brazilian presidential election on twitter. *Computational Propaganda Project*.

MacKay, D. J. and Mac Kay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Mainwaring, S. (2001). *Sistemas partidários em novas democracias: o caso do Brasil*. Mercado Aberto.

Marques, F. P. J. and MontÁlverne, C. (2016). How Important is Twitter to Local Elections in Brazil? A Case Study of Fortaleza City Council. *Brazilian Political Science Review*, 10. ISSN 1981-3821.

Maruyama, M., Robertson, S. P., Douglas, S., Semaan, B., Faucett, H., and Program, I. S. (2014). Hybrid Media Consumption : How Tweeting During a Televised Political Debate Influences the Vote. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*, pages 1422--1432.

Mascaro, C. M. and Goggins, S. P. (2011). Brewing up citizen engagement: The Coffee Party on Facebook. *Proceedings of the 5th International Conference on Communities and Technologies - C&T '11*, (July):11. ISSN 9781450308243.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient Estimation of Word Representations in Vector Space. *Arxiv*, pages 1--12. ISSN 15324435.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013b). Distributed Representations of Words and Phrases and Their Compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, NIPS'13, pages 3111--3119, USA. Curran Associates Inc.

Moisés, J. Á. (2011). O desempenho do congresso nacional no presidencialismo de coalizão (1995-2006). In *O papel do Congresso Nacional no presidencialismo de coalizão*.

Oliveira, L. S., de Melo, P. V., Amaral, M., and Pinho, J. A. (2018). When politicians talk about politics: Identifying political tweets of brazilian congressmen. *International AAAI Conference on Web and Social Media*.

Oliveira, L. S. D., Vaz-de Melo, P. O. S., Amaral, M. S., and Pinho, J. A. G. (2020). Do politicians talk about politics? assessing online communication patterns of brazilian politicians. *ACM Transactions on Social Computing*, 3(4):1--28. ISSN 2469-7818.

Oliveira, L. S. D. and Vaz Melo, P. O. (2017). How to Find the Relevant Words Politicians Use in Twitter? In *Proceedings of the 23rd Brazillian Symposium on*

Multimedia and the Web - WebMedia '17, pages 465--468, New York, New York, USA. ACM Press.

Oren, M. A. and Gilbert, S. B. (2011). Framework for measuring social affinity for CSCW software. *CHI EA '11: CHI '11 Extended Abstracts on Human Factors in Computing Systems*, pages 1387--1392.

Ospina-Valencia, J. (2018). Is there a right-wing surge in south america?

Pain, P. and Masullo Chen, G. (2019). The President Is in: Public Opinion and the Presidential Use of Twitter. *Social Media + Society*, 5(2):205630511985514. ISSN 2056-3051.

Pal, J. (2015). Banalities turned viral: Narendra modi and the political tweet. *Television and New Media*, 16(4):378--387. ISSN 15528316.

Pal, J., Thawani, U., Van Der Vlugt, E., Out, W., Chandra, P., et al. (2018). Speaking their mind: Populist style and antagonistic messaging in the tweets of donald trump, narendra modi, nigel farage, and geert wilders. *Computer Supported Cooperative Work (CSCW)*, 27(3-6):293--326.

Paul, D., Li, F., Teja, M. K., Yu, X., and Frost, R. (2017). Compass: Spatio Temporal Sentiment Analysis of US Election. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '17*, pages 1585--1594, New York, New York, USA. ACM Press.

Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532--1543.

Pereira, N. B. (2011). Sob o piado do twitter: o novo tom das campanhas eleitorais no brasil com a difusão da internet. In *Congresso Luso Afro Brasileiro de Ciências Sociais*, volume 11, pages 1--23.

Persson, T., Tabellini, G., and Trebbi, F. (2001). Electoral Rules and Corruption. Technical report, National Bureau of Economic Research, Cambridge, MA.

Pond, P. and Lewis, J. (2019). Riots and Twitter: connective politics, social media and framing discourses in the digital public sphere. *Information Communication and Society*, 22(2):213--231. ISSN 14684462.

Recuero, R., Soares, F. B., and Gruzd, A. (2020). Hyperpartisanship, Disinformation and Political Conversations on Twitter: The Brazilian Presidential Election of 2018. *Ted Rogers School of Management, 3 Social Media Lab*, 1(2):569--578.

Resende, G., Messias, J., Melo, P., Vasconcelos, M., Benevenuto, F., Sousa, H., and Almeida, J. M. (2019). (Mis)information dissemination in WhatsApp: Gathering, analyzing and countermeasures. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 2:818--828.

Ribeiro, F. N., Saha, K., Babaei, M., Henrique, L., Messias, J., Benevenuto, F., Goga, O., Gummadi, K. P., and Redmiles, E. M. (2019). On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on facebook. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, FAT*'19, Atlanta, USA.

Rori, L. and Richards, B. (2017). *Understanding Online Political Networks: The case of the far right and far left in Greece Pantelis*, volume 10540 of *Lecture Notes in Computer Science*. Springer International Publishing, Cham. ISBN 978-3-319-67255-7.

Saad-Filho, A. (2013). Mass protests under 'left neoliberalism': Brazil, june-july 2013. *Critical Sociology*, 39(5):657--669.

Saldivar, J., Daniel, F., Cernuzzi, L., and Casati, F. (2019). Online Idea Management for Civic Engagement: A Study on the Benefits of Integration with Social Networking. *ACM Transactions on Social Computing*, 2(1):1--29. ISSN 24697818.

Samuels, D. J. (1999). Incentives to cultivate a party vote in candidate-centric electoral systems: Evidence from brazil. *Comparative Political Studies*, 32(4):487--518.

Samuels, D. J. and Zucco, C. (2014). Lulismo, petismo, and the future of brazilian politics.

Sardinha, E. and Costa, S. (2019). Direita cresce e engole o centro no Congresso mais fragmentado da história.

Savage, S., Monroy-Hernandez, A., and Acm (2015). Participatory Militias: An Analysis of an Armed Movement's Online Audience. *Proceedings of the 2015 Acm International Conference on Computer-Supported Cooperative Work and Social Computing (Cscw'15)*, pages 724--733.

Sethi, T. S. and Kantardzic, M. (2017). On the reliable detection of concept drift from streaming unlabeled data. *Expert Systems with Applications*, 82:77--99. ISSN 09574174.

Shah, D. V., Cappella Ramesh, J. N., and Neuman, W. R. (2015). Big Data, Digital Media, and Computational Social Science: Possibilities and Perils. *Annals of the American Academy of Political and Social Science*, 659(1):6--13. ISSN 15523349.

Shapiro, M. A. and Hemphill, L. (2017). Politicians and the Policy Agenda: Does Use of Twitter by the U.S. Congress Direct <i>New York Times</i> Content? *Policy & Internet*, 9(1):109--132. ISSN 19442866.

Shi, T., Kang, K., Choo, J., and Reddy, C. K. (2018). Short-Text Topic Modeling via Non-negative Matrix Factorization Enriched with Local Word-Context Correlations. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, pages 1105--1114, New York, New York, USA. ACM Press.

Shin, J. H. (2017). The choice of candidate-centered electoral systems in new democracies. *Party Politics*, 23(2):160--171. ISSN 1354-0688.

Silva, M., Santos de Oliveira, L., Andreou, A., Vaz de Melo, P. O., Goga, O., and Benevenuto, F. (2020). Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. In *Proceedings of The Web Conference 2020*, number April 2018, pages 224--234, New York, NY, USA. ACM.

Silva, V. d. N. and Silva, R. H. A. (2019). Are algorithms affecting the democracy in brazil? In *INTERNATIONAL SYMPOSIUM ON ETHICAL ALGORITHMS*.

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

Small, T. A. (2010). Canadian politics in 140 characters: Party politics in the Twitterverse. *Canadian parliamentary review*, 33(3):39--45.

Sobaci, M. Z. and Karkin, N. (2013). The use of twitter by mayors in Turkey: Tweets for better public services? *Government Information Quarterly*, 30(4):417--425. ISSN 0740624X.

Souza, F., Nogueira, R., and Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In *9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear)*.

Tan, C., Peng, H., and Smith, N. A. (2018). "You are no Jack Kennedy": On Media Selection of Highlights from Presidential Debates Chenhao. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 2:945--954.

The Economist (2019). Jair Bolsonaro, Brazil's president, is a master of social media.

TSE - Electoral Court (2017). Brazilian law nº 13.488, octuber 6, 2017. `http://www.justicaeleitoral.jus.br/arquivos/propaganda-eleitoral-na-internet`. Accessed: 2019-10-14.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *ICWSM*, 10:178--185.

Turian, J., Ratinov, L., and Bengio, Y. (2010). Word Representations: A Simple and General Method for Semi-supervised Learning. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10.

voz ativa (2019). Voz ativa. `http://www.vozativa.org`. Accessed: 2019-12-19.

Watts, J. (2016). Dilma Rousseff impeachment: what you need to know – the Guardian briefing.

Weber, I., Garimella, V. R. K., and Batayneh, A. (2013). Secular vs. islamist polarization in egypt on twitter. In *Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining*, pages 290--297.

Wikipedia contributors (2021). Cohen's kappa. [Online; accessed 22-fev-2021].

Wong, F. M. F., Tan, C. W., Sen, S., and Chiang, M. (2013). Quantifying political leaning from tweets and retweets. In *Seventh International AAAI Conference on Weblogs and Social Media*.

Yin, W., Kann, K., Yu, M., and Schütze, H. (2017). Comparative Study of CNN and RNN for Natural Language Processing.

Yoon, H. Y. and Park, H. W. (2014). Strategies affecting Twitter-based networking pattern of South Korean politicians: Social network analysis and exponential random graph model. volume 48, pages 409--423. ISSN 00335177.

Zhao, Q., Erdogdu, M. A., He, H. Y., Rajaraman, A., and Leskovec, J. (2015). SEISMIC: A Self-Exciting Point Process Model for Predicting Tweet Popularity.

Zliobaite, I., Bifet, A., Pfahringer, B., and Holmes, G. (2014). Active learning with drifting streaming data. *IEEE Transactions on Neural Networks and Learning Systems*, 25(1):27--39. ISSN 2162237X.