UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-Graduação em Ciência da Computação

Larissa Gomes Malagoli

Dissemination of political content and the viral aspect on Twitter: topological analysis, characterization and classification

Belo Horizonte
2025

Larissa Gomes Malagoli

Dissemination of political content and the viral aspect on Twitter: topological analysis, characterization and classification

Final Version

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 Master in Computer Science.

Advisor: Ana Paula Couto da Silva Ph.D.
Co-Advisor: Carlos Henrique Gomes Ferreira Ph.D.

Belo Horizonte
2025

Malagoli, Larissa Gomes.

M237d Dissemination of political content and the viral aspect on Twitter: [recurso eletrônico] topological analysis, characterization and classification / Larissa Gomes Malagoli - 2025.
1 recurso online (89 f. il., color.) : pdf.

Orientadora: Ana Paula Couto da Silva.
Coorientador: Carlos Henrique Gomes Ferreira.

Dissertação (Mestrado) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Ciência da Computação.
Referências: f. 74-85.

1. Computação – Teses. 2. Redes de computadores – Teses. 3. Redes Sociais on-line – Teses. 4. Disseminação seletiva da informação – Teses. 5. Twitter - Teses. 6. Eleições – Brasil – 2022 – Teses. I. Silva, Ana Paula Couto da. II. Ferreira, Carlos Henrique Gomes. III. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Ciência da Computação. IV. Título.

CDU 519.6*22(043)

Ficha catalográfica elaborada pela bibliotecária Irénquer Vismeg Lucas Cruz CRB 6/819 - Universidade Federal de Minas Gerais - ICEx

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

FOLHA DE APROVAÇÃO

DISSEMINATION OF POLITICAL CONTENT AND THE VIRAL ASPECT ON TWITTER: TOPOLOGICAL ANALYSIS, CHARACTERIZATION AND CLASSIFICATION

LARISSA GOMES MALAGOLI

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores(a):

Profa. Ana Paula Couto da Silva - Orientadora
Departamento de Ciência da Computação - UFMG

Prof. Carlos Henrique Gomes Ferreira - Coorientador
Departamento de Computação e Sistemas - UFOP

Prof. Fabrício Murai Ferreira
Departamento de Ciência da Computação - UFMG

Profa. Carolina Ribeiro Xavier
Departamento de Ciência da Computação - UFSJ

Belo Horizonte, 28 de fevereiro de 2025.

Documento assinado eletronicamente por **Ana Paula Couto da Silva**, **Professora do**

Magistério Superior, em 14/04/2025, às 09:20, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Carolina Ribeiro Xavier**, **Usuária Externa**, em 18/04/2025, às 18:56, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Carlos Henrique Gomes Ferreira**, **Usuário Externo**, em 28/04/2025, às 09:57, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Fabrício Murai Ferreira**, **Usuário Externo**, em 23/05/2025, às 15:43, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4127933** e o código CRC **DC3FB9EC**.

Referência: Processo nº 23072.223124/2025-80 SEI nº 4127933

Acknowledgments

I would like to start by thanking my family, because without their support I would not be where I am today. My mother, Rosângela, that in addition to raising three children with a lot of attention and affection, also committed herself to contributing to the education of several young people through her career, changing the future of many of them and serving as an example of perseverance. You are not only an inspiration, but a constant reminder of the power of hard work, as well as courage and kindness. Along with her, my father, Marcelo, always encouraged me to believe in Science, shaping my interest in logic. Through it I also learned to strive when facing new challenges, being able to stay calm to find solutions. Likewise, I recognize the support of my brothers, Bruna and Marcelo, who were by my side when I needed advice and also made me smile in difficult times. You know how important you are to me and how you helped me through this journey

I would also like to thank my dear friends, who were my anchor during this experience. Gabriel Oliveira, Camila Carlos, Rafael Sobreira, Rômulo Barzotto, Gian Menossi, Vivianne Predebon, Taynara Cruz, Victor Ballagueiro and Michaela Machado, who closely followed the moments of success and also the most troubled periods. You contributed by giving me the strength I needed to balance my studies with moments of joy and you were always by my side, rooting for me. They are incredible people and their friendship is very precious.

I want to express my gratitude to my advisor, Professor Ana Paula Couto da Silva, for her teachings, opportunities and especially her patience throughout this journey. Furthermore, for always encouraging me to improve myself more and more, motivating me to achieve our goals and encouraging me along the way. I am also very grateful to my co-adivisor, Professor Carlos Henrique Gomes Ferreira, who contributed with his vast knowledge and with a lot of support. I consider myself lucky to have had the opportunity to learn and collaborate with you. I deeply thank you for being exceptional mentors.

I could not forget to mention my colleagues on the course and at the LOCUS Laboratory, with whom I spent time and learned a lot. The collaboration and collective spirit provided by them transformed the academic experience into something pleasant and stimulating.

Finally, none of this would be possible without the excellent academic environment provided by UFMG and DCC. I feel extremely proud to have established my career, both Bachelor's and Master's, at this institution where I was able to advance in terms of professional and personal development. Part of the studies contained in this thesis were also partially financed by CNPQ, CAPES and FAPEMIG, therefore the recognition of these important research bodies is also recorded. Finally, I thank everyone who has participated with me on this journey at some point. I express here my gratitude for making this moment in my life so meaningful.

Resumo

As redes sociais online desempenham um papel central na disseminação de informações e na formação da opinião pública. Durante as eleições presidenciais brasileiras de 2022, o Twitter foi amplamente utilizado para compartilhar conteúdos políticos, moldando o debate eleitoral. No entanto, compreender os padrões de disseminação da informação e os fatores que impulsionam a viralidade das mensagens ainda é um desafio. Embora diversos estudos tenham explorado o debate político nas redes sociais, poucos analisaram sistematicamente a interação entre a estrutura de disseminação e as características textuais para prever a viralidade do conteúdo. Neste contexto, este trabalho investiga a propagação do conteúdo político no Twitter durante as eleições brasileiras de 2022, combinando análises topológicas, caracterização textual e modelos de classificação. Primeiramente, modelamos as redes de difusão a partir de retweets e aplicamos técnicas de extração de backbone para identificar os usuários mais ativos na disseminação do debate eleitoral. Em seguida, realizamos uma caracterização linguística dos conteúdos compartilhados, avaliando aspectos como diversidade lexical, complexidade das mensagens e temas predominantes. Por fim, exploramos a adição de informações topológicas a classificação de viralidade, considerando se as mensagens foram ou não disseminadas por usuários nos backbones das redes de difusão construídas anteriormente, visando analisar se tal informação contribui para aprimorar o desempenho da classificação de mensagens virais ou não virais. Os principais resultados indicam que a estrutura topológica da rede é fundamental para entender a dinâmica de disseminação: usuários identificados no backbone possuem um papel crucial na manutenção do fluxo de informações, enquanto as mensagens virais tendem a apresentar maior complexidade linguística. No entanto, a inclusão de informações estruturais no modelo de classificação resultou em um aumento da sensibilidade na identificação de mensagens virais, mas reduziu a precisão geral da classificação. Estes resultados destacam a importância de considerar simultaneamente fatores textuais e estruturais para prever a viralidade do conteúdo e apontam direções para estudos futuros sobre a interação entre comportamento do usuário e disseminação da informação em redes sociais.

Palavras-chave: eleições brasileiras 2022; modelagem de redes; disseminação de informação; Twitter; processamento de linguagem natural; extração de backbone.

Abstract

Online social media platforms play a central role in the dissemination of information and the shaping of public opinion. During the 2022 Brazilian presidential elections, Twitter was widely used to share political content, influencing electoral discourse. However, understanding the patterns of information diffusion and the factors driving content virality remains a challenge. While several studies have examined political debates on social media, few have systematically integrated dissemination structures and textual features to predict viral content. In this context, this research investigates the spread of political content on Twitter during the 2022 Brazilian elections by combining topological analysis, textual characterization, and classification models. First, we model the diffusion networks based on retweets and apply backbone extraction techniques to identify the most active users in political discussions. Next, we conduct a linguistic characterization of the shared content, evaluating aspects such as lexical diversity, message complexity, and dominant topics. Finally, we explored the addition of topological information in virality classification, considering whether or not the messages were disseminated by users of the backbones of the previously constructed diffusion networks, aiming to analyze whether such information contributes to improve the classification of viral and non-viral content. Our findings indicate that the topological structure of the network is essential for understanding the dynamics of information dissemination: backbone users play a crucial role in maintaining the flow of information, while viral messages tend to exhibit greater linguistic complexity. However, incorporating structural dissemination information into the classification model increased recall in identifying viral messages but reduced overall classification precision. These results highlight the importance of jointly considering textual and structural factors to predict content virality and suggest future research directions on the interplay between user behavior and information diffusion on social media.

Keywords: 2022 brazilian elections; network modeling; information dissemination; Twitter; natural language processing; backbone extraction.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

The dissemination of information or opinions is considered a complex contagion process, as a person's connections in their social group as well as their beliefs will influence the final picture of the spread of information, survival, and the intensity of the level of discussion [12, 69, 43, 74]. Building on this, the process of information dissemination often manifests as a collective behavior, where groups of individuals actively seek to expand or reveal their views on specific topics [100]. Understanding how this important process unfolds in our society is not a new task and has always been a huge challenge for different research realms [29, 58, 97]. However, the pervasive presence of online social media platforms (OSMPs) in the daily life of today's society has added much more complexity to these studies. These online environments are seen as an extension of our social relationships, as they facilitate dialogue and connect people, providing channels of communication and enabling the expansion of discussions on a wide variety of topics.

Several factors may influence the sharing of information on OSMPs. One significant aspect is the algorithms used to populate users' timelines [8, 108]. However, it is also well-established that the way users organize themselves to share information, as well as the textual characteristics of the content, play a crucial role in this diffusion process [111, 54, 82]. The combined impact of these two factors in promoting content on OSMPs remains a challenging research problem. This issue is critical to analyze, especially since boosted content can sometimes go viral and potentially cause harm to society [41, 53, 112, 67]. This is particularly relevant given the role of OSMPs as key platforms for organizing and driving major social movements worldwide, ranging from health [65, 37] to politics [38, 99].

This scenario is no different in Brazil, where OSMPs are often used for political debates. Several studies in the literature [62, 19, 94, 56, 111, 99, 25, 84, 54, 44, 40, 89] characterized Brazilian political debates across various platforms. In this plethora of OSMPs, platforms such as X (former Twitter)[1] [93, 17, 88, 99, 84], WhatsApp [78, 79, 89] and Telegram [111, 82, 11], have been the focus of many works that aimed at analyzing how political information is disseminated on OSMPs.

In the light of this, the widespread use of Twitter as a key platform for political

[1] Twitter has been recently rebranded as X. However, we maintain the reference to the old platform's name as our study is based on features commonly associated with it.

Figure 1.1: Twitter was a platform where the 2022 election debate was greatly encouraged and had a lot of engagement.

discussion in Brazil makes it crucial to understand how information spreads and what factors contribute to content virality. While prior research has explored political discourse on social media, few studies have systematically integrated both textual and structural dissemination patterns to classify viral content. In this master's thesis, we address this gap by analyzing the 2022 Brazilian Presidential Election debate on Twitter. First, we model the diffusion network and apply backbone extraction methods to identify the core set of users who play a central role in spreading political content. Then, we examine the linguistic properties of the messages being disseminated. Finally, we assess whether incorporating network-based dissemination signals alongside textual features improves the classification of viral and non-viral content. By comparing a traditional transformer-based classifier trained solely on text features with an improved model that incorporates structural dissemination information, we provide an orthogonal view of the interplay between content and user behavior in the dissemination process. This study contributes to a more in-depth understanding of how structural factors influence political information flow and helps refine methodologies for content classification in dynamic online environments.

1.1 Research Goals

The main research goals of this master's thesis are threefold: (i) to analyze how users organize themselves to promote specific pieces of information; (ii) to examine the textual features of content shared by users with varying levels of engagement in Twitter discussions; and (iii) to evaluate how the combined impact of user engagement and textual features can improve the accuracy of classification models for distinguishing viral content. To address these research goals, this master's thesis is structured around three primary research questions (RQs):

- **RQ1: Is it possible to identify a group of users who recurrently spread similar content on Twitter during the two rounds of the Brazilian elections of 2022? Do they remain active and persistent in the analyzed time intervals?**

  We built media-centric networks that connect people who have shared similar messages based on the collected retweets. Backbone techniques were then used to eliminate peripheral and bridging connections based on the connected components of the network structure, and highlight edges that show consistent and repetitive activity between pairs of users to identify users whose sharing habits are not random. The individuals at the core of the networks extracted from the backbones were labeled as key-users and may have stimulated the dissemination process. Within this group, there is an even more select group of users who were also analyzed and who shared content during the two election rounds.

- **RQ2: Do the textual characteristics of the content change based on the user's level of engagement?**

  We categorized the disseminated messages as widespread dissemination, dissemination by the key individuals and dissemination by the persistent individuals. Our aim was to understand how the debate evolved from these different perspectives and how each contributed to the process. To that end, we employed different techniques such as LIWC [106] (psycholinguistics analysis), BERTopic [47] (topic extraction analysis) and XLM-RoBERTa [61] model (sentiment analysis) to distinguish textual characteristics from each of the user groups.

- **RQ3: Does incorporating information about user engagement, captured through the presence of messages in the backbone of the dissemination network, alongside textual features, improve the classification of viral and non-viral content during the 2022 Brazilian election?**

  To evaluate whether the combination of user engagement—represented by the presence of messages in the backbone of the dissemination network—and textual features enhances the distinction between viral and non-viral content, we employed classifiers based on fine-tuned transformers. We assessed performance by comparing two scenarios: one relying solely on textual features extracted from the messages, and another integrating information about the sharing behavior of groups of users identified within the backbone topology. This approach allows us to determine whether structural dissemination patterns contribute to a more accurate classification of viral content.

1.2 Contributions

This master's thesis focuses on expanding knowledge about the political discourse during the 2022 elections and analyzing the phenomenon of dissemination. The key contributions of this work are outlined below.

Identification and characterization of users who participated in the social media diffusion. We built dissemination networks using shared retweets and applied backbone extraction methods. Our initial findings revealed that these networks were sparse and noisy, with backbone analysis highlighting a core of connections that promoted content spread. We identified users who significantly influenced the discourse and exhibited less random behavior in the network, as well as individuals who consistently participated in the diffusion flow across both electoral rounds. We accomplished this by comparing the intrinsic characteristics, like followage and if they were present or not in the backbones, of each type of set of users. Our main contributions on this topic were published in [64].

Characterization of textual content in the electoral debate. We employed well-established techniques from the literature to analyze the content of diffuse messages across different dimensions, highlighting the potential influence of these characteristics on the broader debate. This included identifying the most prevalent emotions conveyed in the messages and representing the primary psycholinguistic attributes correlated with the disseminated content. Our primary findings indicated that content dissemination centered on topics related to religion, family, information and opinions about deaths during the COVID-19 pandemic, and Lula's electoral victory. Additionally, we identified key topics that appeared to correlate with real-world events occurring during the electoral period, providing valuable perspective into the interaction between online discourse and offline events. Our main contributions were published in [64].

Enhanced BERT-based model to distinguish viral content. We extended the BERTimbau classification model by incorporating structural dissemination features to assess their impact on predicting viral and non-viral content. Specifically, we investigated whether integrating information about a message's presence in the backbone of the dissemination network could enhance classification performance. By comparing this enhanced model with a traditional BERTimbau-based classifier that relied solely on textual features, we analyzed differences in precision, recall, and overall separability between viral and non-viral messages. Our results revealed that the inclusion of topological information improved recall, making the model more sensitive to identifying viral messages, but at the expense of precision and overall classification balance. suggests that the backbone may be conveying information about consistent sharing patterns within specific user groups rather than indicating overall global virality. Our findings highlight the potential trade-offs in leveraging network structure for virality prediction and opens avenues for future research on integrating more nuanced topological signals.

1.3 Master's Thesis Organization

This work is organized as follows. Chapter 2 presents the literature review, covering studies on social media analysis related to network modeling, political events, and natural language processing within our context. Chapter 3 details the methodological approach, including data collection on Twitter, dataset consolidation, network modeling, and backbone extraction. It also describes the structure and development of the virality classification models. Chapter 4 explores the topological differences between complete dissemination networks and backbones, demonstrating how backbone extraction helps identify patterns of users with less random interactions. Additionally, we examine the main topics propagated, the linguistic and sentiment characteristics of the messages, and the profiles of disseminating users. Finally, we analyze the attributes of viral and non-viral messages and present the classification model results, evaluating the impact of both textual content and topological features. Chapter 5 concludes the study, summarizing key findings and suggesting directions for future research.

Chapter 2

Related Work

The impact of OSMPs on the daily life of society has led a large number of researchers to propose different computational techniques and algorithms to study how offline social processes, such as the dissemination of information, are mimicked on these platforms [7, 39, 92, 111, 82, 84, 54, 80, 17, 57, 75] As the literature on this field of research is very extensive, in this chapter we limit ourselves to presenting some works that are better suited to our research topic, namely the use of OSMPs for the discussion of political events.

We begin the chapter by presenting some works that characterize the political worldwide debate on OSMPs (Section 2.1). We narrow our focus on presenting works that study the political debate in Brazil (Section 2.2). We conclude this chapter by pointing out how our work complements the previously studies in literature (Section 2.3).

2.1 Worldwide Political Debate on OSMPs

All over the world, political issues are always at the center of public interest, as they are of crucial importance for the sustainability of society. It is therefore not surprising that discussions that were previously only held offline are migrating to the online world. Next, we present a set of works that have analyzed how various political events were driven by the most popular OSMPs. We categorized OSMPs into four groups depending on the type of content they share most frequently: (i) Video Media OSMPs (Tiktok, Youtube); (ii) Mixed Content Media OSMPs (Facebook, Whatsapp, Telegram, Instagram); and (iii) Text Media OSMPs (Twitter, Reddit).

Video Media OSMPs. Cervi *et. al.* [21] reveal the political aspect of Tiktok content in Spain by analyzing how political parties used the tool as a means of communication using data collected manually by the authors. They categorized the posts in terms of genre, format, content, source and interaction and suggested that all parties used TikTok exclusively to publish political content without fulfilling the original purpose of the app, which is entertainment. Authors in [2] analyzed content from TikTok posts in Spain, Portugal, Brazil, and the United States, during the year 2020, a complex period as a consequence of COVID-19 and various global political episodes. The main paper findings revealed that TikTok's visual and viral nature facilitates the spread of disinformation while also offering opportunities for debunking hoaxes beyond the reach of traditional media.

Boyd *et. al.* [16] investigated the different roles that people who comment on YouTube videos can take on. The study analyzed comments from the video of Barack Obama's inaugural address on the platform and showed that pronouns involving the actor of the comment were the most popular lexical terms indicating involvement. When the comments were labeled as part of conversations, they were mostly labeled as constructive. The author also had findings that indicated that disruptive comments were associated with hateful or conspiratorial language, for example.

Mixed Content Media OSMPs. Instagram is one of the most popular OSMPs where users can express their opinions through posts. The authors in [3] investigated how political ideas were discussed during the Iranian parliamentary elections in 2016. They used a diffusion network model, the NetRate algorithm, to understand the process of hashtags spreading through followee-followers. They also identified communities in the networks using the Louvain algorithm. In [55], the authors analyzed the 2017 federal election in Germany. They proposed a framework in which the content of politicians' profiles as well as parties' accounts were used to characterize the strategies adopted in the election period. They focused mainly on the posted image or video and used additional information such as the hashtags or mentions in the posts. The results suggested that parties and candidates used complementary communication strategies instead of relying on a single tactic. Additionally, the analysis showed that parties in parts strategically violated Instagram's communication norms to convey hard-to-visualize policy messages while candidates preferred to post professional personalized posts.

The study by Ribeiro et al. [90] analyzes Facebook's advertising infrastructure to explain the impact of malicious ads allegedly placed by the Russians during the 2016 US election. The 2020 US election was analyzed by Guess et al. [49], where they investigated how Facebook and Instagram timeline algorithms could influence users' opinions and actions during the election period. Their suggestion was to change the way some users interacted with the platform so as not to expose them to the original algorithms, which would result in less social media use. According to the results, there was no significant change in polarization when addressing the research group's issues.

The research of [102] is based on the study of Telegram posts with political advertising in praise of the Russian Kremlin and how this affected the conflict between Ukraine and Russia from 2022 onwards. They collected information in five different languages and used natural language processing methods, including linguistic feature extraction and the application of Support Vector Machine (SVM) models, to identify linguistic and narrative features in the messages. The results indicate that the pro-Kremlin messages use language resources that are not very similar to natural language and modify some geopolitical terms.

The authors of [17] explored Latin American election scenarios by leveraging machine learning techniques to predict outcomes based on social media engagement and polling data. They gathered data from Facebook, Twitter, Instagram, and polling websites in Argentina, Brazil, Colombia, and Mexico during 2018 and 2019. The engagement metrics provided by the platforms were utilized as features to develop the Social Media Framework for Election Nowcasting (SoMEN). This framework integrates Multilayer Perceptrons with the Backpropagation strategy (MLP-BP) and General Regression Neural Networks (GRNN) to predict candidates' vote shares in elections. The framework was applied to publish predictions for the 2019 Argentine elections and the 2020 US elections. It demonstrated strong performance, achieving a Mean Absolute Error (MAE) lower than the average of comparable polls conducted during the same period.

Valenzuela et al. [107] examined the phenomenon of content sharing among Chilean users on WhatsApp during the 2017 elections in Chile. Through a survey conducted with a group of individuals, the study explored various aspects of platform usage. The findings highlighted how the subtle shift in context between personal matters and political issues is interconnected on the platform, with news consumption emerging as a significant use case in the region.

The significant influence of social media platforms on political issues is not inherently negative. The authors of [51] proposed a theoretical model to evaluate the effectiveness of political content sharing on platforms like Facebook and Twitter in persuading individuals to engage politically. This study, conducted in Chile, utilized data from a two-wave panel survey completed by adults in 2013. The findings revealed that Facebook is more effective for collective sharing of political content, while Twitter tends to to be suited for to individual-oriented political interactions.

Text Media OSMPs. Morini *et al.* [70] developed and validated a method for detecting echo chambers using data from Reddit during Donald Trump's presidency. Echo chambers, prevalent in social media, are environments where individuals with polarized beliefs have their opinions continuously reinforced by others who share similar views and motivations. The authors applied a Recurrent Neural Network (RNN) technique, Long Short-Term Memory (LSTM), and Bidirectional Encoder Representations from Transformers (BERT) models to measure users' ideologies and extracted network statistics to classify political stances as neutral, in favor, or against Trump. The findings revealed that users rarely maintain a neutral stance on all issues, often adopting polarized positions on controversial topics. Additionally, the study indicated a 30–40% likelihood that users might change their beliefs over time.

The study by [50] also examined a collection of Reddit posts from a U.S. political news forum that sparked negative reactions or were flagged as controversial. Their findings offered a systematic analysis of user actions, such as replies and votes, within discussion threads, revealing distinct patterns that refine the concept of controversy into three categories: disputes, disruptions, and discrepancies. Notably, the research linked user actions to a comprehensive feature space that included sentiment expressed in post content and topical variations across posts and replies.

Turning our attention to Twitter, Romero et al. [92] constructed a mention-based network using a dataset of tweets primarily focused on the United States. This dataset included various types of content, with a notable emphasis on politically themed messages, such as those referencing political figures, commentators, parties, or movements. The study aimed to explore the impact of hashtag usage on the platform using data from 2009 and 2010. To achieve this, the authors developed an extensive set of categories to annotate content topics, including political themes. They then analyzed subgraphs created from the initial users who mentioned a specific hashtag to understand how content spread from these early adopters throughout the network. Using simulations, they examined the cascading effects of hashtag propagation. A key finding of the study was that the repeated exposure of users to politically controversial hashtags significantly influenced their likelihood of adopting and propagating those hashtags.

Regarding 2016 US election, the authors of [20] presented a study on user engagement on Twitter during that election. The influence of homophily was demonstrated in different scenarios by applying sentiment analysis techniques to classify tweets and examining the number of connections between users. Their findings showed that users that interacted and connected more with individuals exclusively from their same class (i. e. had higher homophily levels) were classified in the Negative, Trump Supporters or Hillary Supporters classes.

Still in the North America, Dubois *et al.* [33] conducted a study using data from Canadian political communities to construct a follower network based on tweets from 2013 containing hashtags associated with one of the country's two major political parties. The study aimed to evaluate user influence within the network. They concluded that following others, interaction within the network, and an individual's structural position within their local neighborhood are critical factors in understanding influence on Twitter.

The 2013 Egyptian protest wave sparked significant polarization on Twitter [14]. Researchers employed data mining techniques to collect Arabic-language messages and user information from participants who engaged in the discussion more than 10 times.

Using this extensive dataset, tweets were classified as pro- or anti-military intervention with a Support Vector Machine model (SVM), revealing that the majority were anti-military. The study also examined shifts in user positions over time. To capture these dynamics, retweet networks were constructed, uncovering a transition where secularists initially formed a strong and committed group but experienced a decline in both strength and cohesion following the *Egypt's Coup d'Etat*.

Another study focusing on an African country analyzed the *Revolution Now* debate in Nigeria, a 2019 political activism event that also garnered significant attention on social media [81]. The research explored how linguistic features influence message dissemination on Twitter. Using a dataset of related keywords, the authors extracted prominent linguistic attributes with the Linguistic Inquiry and Word Count (LIWC) tool to examine their impact on engagement metrics, such as likes, replies, and retweets. The findings revealed that in political protests, tweets closely aligned with the main topic, containing quantifiers, quoted phrases, and proper punctuation, are more likely to generate high engagement.

2.2 Brazilian Debate on OSMPs

Our study is not the first to analyze the Brazilian political landscape on OSMPs. The content and activities of social media in Brazil have unique characteristics and can offer valuable insights into the country's political dynamics. In this section, we review several studies from the literature that address this topic. As before, we structure our discussion based on the categorization of social networks, acknowledging that different platforms often appeal to different audiences and purposes.

Video Media OSMPs. Brazilian politicians have recently started to address the population via livestream instead of relying on traditional media. Di Nubila *et al.* [31] investigated how this technopopulism via YouTube was used as a resource during the Bolsonaro government. Strategic decisions were made in the settings of these livestreams. The study provided a statistical analysis to show the probabilities of increasing manipulation of public opinion and the spread of government propaganda.

Electoral events are key moments for the diffusion of political content and can be shaped by various factors. The authors of [56] presented a method to examine gender bias through engagement and sentiment analysis of comments on YouTube videos that include the names of men or women in their titles or descriptions. The study covered the Brazilian elections of 2018 and 2022, focusing on the presidential, gubernatorial, and senatorial races. The quantitative analysis revealed that videos featuring men experience higher engagement, while videos featuring women show a greater prevalence of negative sentiments in the comments.

Regarding TikTok, Santana *et al.* [94] presented a study that analyzed the use and engagement of the TikTok profiles of the two leading presidential candidates Lula and Bolsonaro, during the 2022 election. The authors found out that Lula used the trend posts strategy in his profile and his supporters engaged more in a positive and neutral perspective, while Bolsonaro used collaboration with influencers to achieve higher metrics in the campaign videos, measured by the statistics of numbers of plays, likes, comments and shares. Bolsonaro's videos also had a more negative sentiment and agressive engagement from the commenters. Lima *et al.* [59] also examined the contributions of the candidates for Tiktok, focusing in a broader period of the same election, including the pre-electoral as well as the election event itself. Their results showed that politicians had similar usage compared to other OSMP platforms in the last election and that the share of engagement they received was consistent with the actual share of the vote.

Mixed Content Media OSMPs. Instagram is a popular communication channel for political discussions in Brazil. Analyzing the content posted by politics is an important research topic in this theme. By characterizing the posts of candidates for the Presidency of the Republic of Brazil in 2018, Sampaio *et al.* discovered that politicians preferred to emphasize and work on their public image as a main strategy, instead of actually providing proposals and discussing their government platforms. Costa *et al.* [24] took a similar research approach, examining how Instagram was utilized during the 2020 mayoral elections in a city in the South of Brazil. The results indicated that the platform was predominantly used for positive posts, with candidates employing this strategy to make politics more personal and accessible to voters.

The comprehension of how users organize themselves in order to engage in OSMPs is important, since these platforms promote public posts into users timelines and incentive users to interact. The study by Ferreira *et al.* [39] defined a probabilistic model for extracting interactions occurring on the backbone of Instagram constructed networks for Brazil and Italy, similar to the backbone analysis we performed in our work. The authors analyzed the formation of communities during the debate of presidential elections in 2018 in both countries, finding evidence of coordination. The network comparison provided insights on significant structural variations and unique dynamics that occurred in the communities observed.

Moving to Whatsapp, which is a highly used platform in Brazil, Nobre *et al.* [78, 79] analyzed messages exchanged during the Brazilian 2018 parliamentary elections in a network-focused research. The authors built a network based on users sharing the same content messages in one or more groups. They used backbone extraction techniques to search for highly connected user communities and tried to detect possible coordination actions when certain contents are shared.

In a related direction, the authors of [89] proposed a WhatsApp tool designed to enhance the fact-checking process by analyzing images and applying ranking-based strategies. The core technique involved calculating a score to estimate the likelihood that a given piece of content was fake news. Their tool aimed to support the news verification process and was validated using data from the 2018 Brazilian general elections. Similarly, Evangelista et al. [35] explored the dynamics of misinformation during the same election period on WhatsApp. Their study provided an in-depth analysis of misinformation examples and examined the role of user groups in fostering polarization. Furthermore, Machado et al. [62] investigated the dissemination of misinformation on WhatsApp during the 2018 elections. Their findings highlighted that viral content within WhatsApp groups predominantly consisted of hate speech and fake news.

Finally, the 2022 Brazilian elections marked a period of intense polarization, sparking widespread debate and capturing significant public attention. In this context, the study by [111] employed backbone extraction methodology to identify evidence of coordination in the dissemination of information on Telegram. This work underscored the growing impact of messaging apps on political mobilization and provided valuable insights into digital communication strategies in contemporary electoral processes. Similarly, the authors of [82] analyzed dissemination patterns on Twitter and Telegram using data from the 2022 elections. Their research focused on the timing of disseminated messages and its influence on dissemination patterns, offering potential indicators of user coordination in spreading information.

Text Media OSMPs. In the context of Twitter, political debates have long been encouraged by the platform's rapid response and trending structure. In [19], a characterization of pre-election advertising messages in Twitter was made for the 2016, 2018 and 2020 elections to compare the evolution over time. By applying psycholinguistic and sentiment analysis techniques, the results showed that initial advertising messages tend to be negative or neutral, with neutral sentiment increasing over time, and that there is a pattern in the use of hashtags and links, as well as mentions of different entities. The research showed that user sentiment when posting content can be an important factor when interpreting the possible relationship between social media users' opinions. The work in [25] in this analysis focused on understanding the final result of the 2018 elections in Brazil. Naive Bayes and SVM (Support Vector Machine) were used to classify the tweets. Their results indicate a correlation between the proportion of positive sentiment towards each candidate and their popularity in official election polls. In addition, Nobre *et al.* [80] also provided an overview of public opinion on political events on Twitter and showed how users use mixed feelings when discussing the topic.

Distrust in the electronic voting system and rumors of voter fraud are hot topics among voters nowadays. Recuero *et al.* [88] approached the topic with a survey of content created with the hashtag "#FraudenasUrnas"[1] in Twitter for the 2018 elections. The outcomes of the study suggested that the disinformation tweets with higher engagement used a discursive legitimization strategy, with story elements that "prove" the fraud and specific wording that gave authority to the text.

Next, we highlight studies that attempt to distinguish different perspectives on the political event that is the subject of our investigation in this master's dissertation: the Brazilian 2022 elections. Silva *et al.* [99] analyzed the opinions and sentiments expressed by users about the 2022 presidential candidates to verify whether the candidates' performance is related to their popularity on social media. The authors of [95] also relied on this technique for this political event, but focused on finding the machine learning algorithm with the best result for the topic. They used six different classifiers and concluded that SVM and Random Forest provided the best results in terms of TFIDF vectorization. The percentage of votes the presidential candidates received in the first round was compared to the sentiment score of the data collected from Twitter. Both authors believe that the candidates who received the most votes are also the ones who were discussed the most on social media.

In another direction, Paiva *et al* [84] took a different look at the 2022 elections in Twitter, focusing on the feminism debate. They analyzed more than 700 thousand tweets in Portuguese in the periods before, during and after the elections, using LDA topic extraction, sentiment analysis and psycholinguistic techniques to examine the data. The main findings were that negative sentiments prevailed in the messages and that they referred to sensitive topics. Pacheco *et al.* [83], in turn, investigated how bot accounts were used in the 2022 election debate on Twitter. The author applied a framework to determine the likelihood of coordinated accounts and the BotometerLite tool to search for bot users. Key findings were a significant correlation between the number of replies received on a given day and bot participation, as well as a summary of key profile characteristics of bot accounts.

In the context of polarization, Santos et al. [96] investigated the hostile content surrounding the political events of 2022. Using Twitter data, they applied stance detection to categorize users into polarized clusters, analyzing interactions between the two primary opposing sides in the presidential election. Their findings revealed that most of the antagonistic information consisted of insults, attacks on the identities of ministers, or derogatory remarks about the election process. The authors of [54] used topological structure analysis to investigate the contributions of specific groups to network polarization, also considering the 2022 elections. Their results revealed a significant degree of isolation among certain communities and showed that right-leaning users were particularly effective at coordinated communication within the network.

[1]Fraud in the voting machines

Table 2.1: Related work and our contributions.

Related Work	Work Domain	Netw. Model.	Backbone Extrac.	Content Charact.	Messag. Classific.	Our Work
[Ferreira et al., 2020] [39]	Use backbone extraction based on null models to understand how political and general content is shared across Instagram users.	✓	✓	✓	✗	Use backbone techniques (DF+NB) to characterize user engagement on Twitter and use this information to classify virality.
[Romero et al., 2011] [92]	Analyze the process of information dissemination on Twitter by comparing the dissemination of political and non-political content in a US-focused context using keywords/hashtags.	✓	✗	✗	✗	Examine the Brazilian political aspect and focuses on the user groups in the network, the textual characteristics and the viral content.
[Venâncio et al., 2024] [111]	Application of backbones and analysis of message topics in Telegram for the 2022 Brazilian elections and the 2023 Brazilian coup attempt, focusing on the investigation of possible user coordination.	✓	✓	✓	✗	Provide a view of the 2022 elections on Twitter where content tend to be highly diffused by the retweet feature, also providing network modeling, backbone extraction and content analysis. Addition of viral messages classification.
[Paiva et al., 2023] [84]	The textual analysis of content from the 2022 elections to understand the meaning of the messages offers a view that focuses on the issue of feminism in political debates.	✗	✗	✓	✗	Provide a general analysis focusing in virality as well as network features.
[Interian and Rodrigues, 2023] [54]	Built a graph network with Twitter data from the 2022 Brazilian election and attempted to contribute to knowledge about how this political event unfolded with a focus on polarization and coordinated communication.	✓	✗	✓	✗	Examine virality and dissemination, taking a deeper look in the networks backbone.
[Nobre et al., 2018] [80]	Analyze Brazilian political events using data from Twitter, characterize the content based on the sentiment facet of the messages.	✓	✗	✓	✗	Focus on the dissemination process, providing deeper analysis of the retweet diffusion network.
[Brito and Adeodato, 2023] [17]	Use of Machine Learning techniques to predict political content in social media in a Latin American context.	✗	✗	✗	✓	Focus on the network aspect of dissemination of content, as a basis to the prediction tasks.

2.3 Summary

The analysis conducted in this master's thesis builds upon and complements the findings of the aforementioned studies by offering a more comprehensive range of analyses using a large Twitter dataset. Our work employs backbone extraction methods to explore how political debates unfold across different levels of user engagement on the platform. Additionally, we present an initial investigation into how textual characteristics and topological features can enhance the understanding of how political content becomes viral.

Table 2.1 provides a comparative overview of studies closely aligned with ours, emphasizing how our findings contribute to a richer understanding of the evolution of online discussions on social media during this important event in the recent history of democracy in Brazil.

Chapter 3

Methodology

In this chapter, we present the methodology underlying the analysis of the Brazilian political debate on Twitter. We divide this chapter into six main sections. First, in Section 3.1, we present the methodological process used and then elaborate on each of the steps. We describe how Twitter data were collected (Section 3.2), followed by network formalization of the proposed dissemination modeling (Section 3.3), as well as the backbone extraction process definitions (Section 3.4). Afterwards, we detailed the content analysis performed using shared content (Section 3.5) and finally explain the classification models used to label a content as viral or non-viral (Section 3.6).

3.1 Overall Methodology

Figure 3.1 shows all phases of the methodology we used to study the process of information dissemination during the two rounds of the Brazilian elections. We started by collecting approximately 741K Portuguese-language tweets shared during the two rounds of the 2022 Brazilian general elections, which took place on October 2^{nd}, 2022 and October 30^{th}, 2022 (**Step 1: Data Collection**). Next, we formalize the media-centric network built for information diffusion, one network for each round (**Step 2: Network Modeling**).

We then applied backbone extraction techniques (**Step 3: Backbone Extraction**) to remove noisy or sporadic edges of the media-centric network, identifying the core of the information dissemination networks and addressing **RQ1**, which aims to identify the users who intensively participated in the debate. Subsequently, we characterized the content shared by the users in our dataset (**Step 4.a: Content Analysis**) and developed a classifier based on textual and topological features to predict whether a piece of content is viral or not (**Step 4.b: Virality Classification**), addressing **RQ2** and **RQ3**. The content analysis uncovers linguistic characteristics that offer insights into why certain topics generate higher engagement, while the virality classification model evaluates how combining textual features and user engagement can help distinguish viral from non-viral content. This approach also aids in identifying potentially unreliable or undesirable content that may spread widely across the network.

In the following sections, we go into the individual steps of our methodology and describe the techniques we have chosen to achieve the goals of this master's thesis. We believe that our methodology can be applied to analyze how real-world events in general are discussed on Twitter.

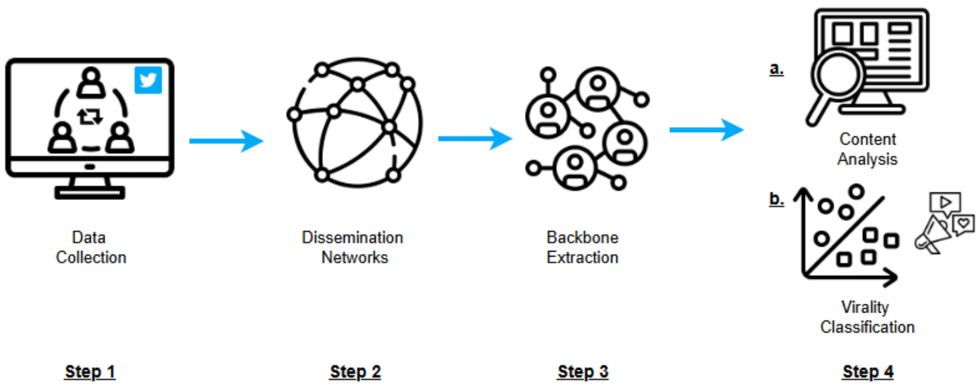


Figure 3.1: Methodology of the research to understand the dissemination process.

3.2 Data Collection

In this section, we describe the dataset used in our study. The dataset we analyzed consists of approximately 741K Portuguese-language tweets shared during the two rounds of the 2022 Brazilian general elections, which took place on October 2^{nd}, 2022 and October 30^{th}, 2022. Subsection 3.2.1 describes the crawling strategy we used and Subsection 3.2.2 gives an overview of our dataset.

3.2.1 Collected Dataset

Our dataset comes from data collected using the Twitter API Search[1] and the Tweepy tool.[2] We built a list of keywords related to the discussion topics that have ap-

[1] https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets

[2] https://docs.tweepy.org/en/latest/

Table 3.1: Collected keywords.

Context	**Keywords**
Candidate names	**Lula2022**, **Bolsonaro2022**, Ciro2022, **Eleições2022**, **eleições**, **Lula13**, **Bolsonaro17**, **Bolsonaro22**, **lula**, jair, **bolsonaro**, **cirogomes**, **Ciro**, **SimoneTebet**, Eleicao2022
Campaign words	**LulaPresidente**, Vote13, LulaNoPrimeiroTurno, **Lulinha**, **BrasilComBolsonaro22**, EuVotoBolsonaro22, LulaNoPrimeiroTurno13, LulaNoSegundoTurno, CapitaoPeloBemDaNacao, BolsonaroPresidente22, LulaEoPTTudoNoPrimeiroTurno, , BrasilVota22, BrasilDaEsperanca, luladay, BolsonaroNoPrimeiroTurno, BolsoNey, **FechadoComBolsonaro**, **VotoNuloSim**, LulaNo1ºTurno, QuemAnulaVotaLula, EsseÉMeuVoto, **Lulindo**
News media	JornalNacional, CPIDasPesquisas, Bonner, Datafolha, CNNnasEleicoes, BolsonaroMente, oLeDeLadrao, TiraGomes, PadreDeFestaJunina
Agressive or offensive	**BolsonaroMentiroso**, **Bolsolixo**, **Bonoro**, **Bozo**, ForaBolsonaroVagabundo, PadreDeFestaJunina, Bolsominion, LulaXenofobo,
Debates and events	SpaceDoLula, LulaNoFlow, LULANODLSHOW, DebateNoSBT, CiroNoFlow, SuperLiveBolsonaro22, LulaNoSBT, LulaNaGlobo, LulaNaCNN, BolsonaroNoJN, BolsonaroNaBand, CiroNaBand, LulaNoJN, LulaNaBand, LadraoNoJN, DebateNaBand, CiroNoJN, DebateNaGlobo, LulaNaGlobo, SimoneNoJN, CiroManifestoANacao, LulanoJornalNacional, 7DeSetembroFoiGigante, setedesetembro

peared as trending topics on Twitter.[3] A trending topic is the one that is heavily discussed on Twitter in a given time period. By looking at the trending topics, we highlight the most discussed topics on this platform. Moreover, we also selected a complementary set of election related keywords, such as the presidential candidates and the Twitter official hashtags used during the Brazilian election. Our collect strategy enables to gather messages containing popular debate terms, such as candidate- and campaign-related words, events in which the presidential runners-up participated, election-related news topics, and even aggressive terms disseminated by users whose stance was anti particular candidates. Table 3.1 presents the crawled keywords. We gathered 27,411,279 tweets and 90,211,448 retweets.

We then take a closer look at the temporal evolution of election-related discussions by showing the daily time series of the number of tweets and retweets from August 21,

[3] https://help.twitter.com/en/resources/glossary

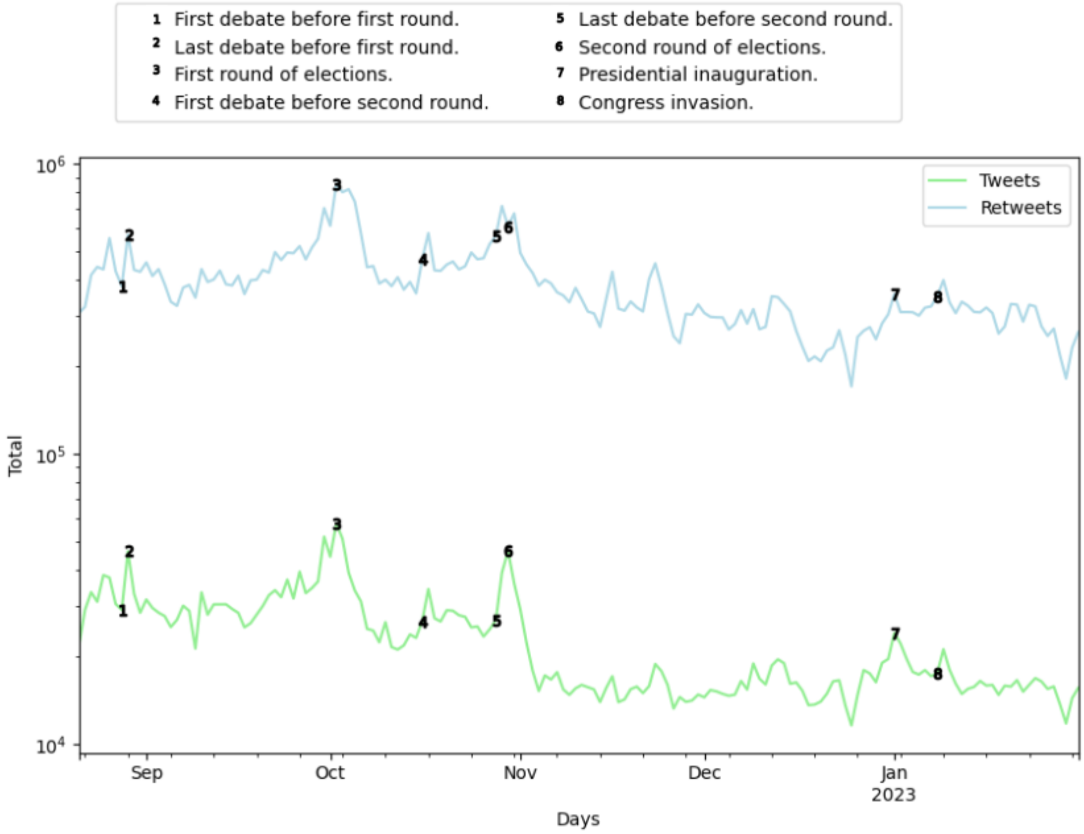


Figure 3.2: Daily numbers of tweets and retweets with annotations of dates with relevant election related events.

2022 to January 31, 2023 in Figure 3.2. To create both time series, we considered the bold keywords in Table 3.1, as they consistently appeared in the trending topics throughout the data collection period. The figure also highlights the days of some important real-life events related to 2022 Brazilian election including, presidential candidate debates (events #1,#2,#4 and #5), the first and second round elections (events #3 and #6) and the presidential inauguration (event #7) and the Brazilian congress invasion (event #8). Note that there are some significant spikes in the volume of tweets and retweets, which coincide with the first and second election rounds. Thus, several external events may have propelled the increase in the debate about 2022 Brazilian election on Twitter during the selected period.

Based on this preliminary analysis of the dataset, we focused on the Twitter discussions that took place during the first and second rounds of the election (October 2^{nd}, 2022 and October 30^{th}, 2022, respectively), i.e. the events that cause the spikes in the time series shown in Figure 3.2. To have a specific view of the political discussion, we built a list of keywords that includes terms such as the official election keyword used by Twitter and the most voted presidential candidates. Specifically, we consider the following list of keywords for the research experiments: *Eleições2022* [4], *Lula, Bolsonaro, Ciro,*

[4] Elections2022

Table 3.2: Data overview from the first and second election rounds.

Keywords	**First Round**			**Second Round**		
	# Unique Tweets	**# Retweets**	**# RT/ TW**	**# Unique Tweets**	**# Retweets**	**# RT/ TW**
Eleições2022	2,145	186,594	86.99	875	153,525	175.46
Bolsonaro	3,771	154,165	40.88	2,716	112,840	41.55
Lula	3,094	100,066	32.34	2,930	96,546	32.95
Ciro	1,705	44,185	25.91	445	11,995	26.96
SimoneTebet	4	64	16	4	197	49.25
Total	8,774	409,956	46.72	5,835	331,175	56.76

SimoneTebet.[5] [6]

3.2.2 Election Dataset Overview

Our final election dataset is formed by approximately 741K Portuguese-language tweets (retweets) shared during the two rounds of 2022 Brazilian general elections, occurred on October 2^{nd}, 2022 and October 30^{th}, 2022. Table 3.2[7] shows the total number of unique tweets and retweets for each keyword, after filtering tweets with at least two users in the dataset who retweeted that message. In addition, tweets with the same ID that contain multiple keywords were only counted once. As expected, in our dataset, the official Twitter keyword for the general elections (*Eleições2022*) is the most shared, followed by the keywords with the name of the two main presidential candidates.

To provide an overview of our data, we look into the contents of the collected retweets. We do so by showing in Figure 3.3 the word clouds with the top 100 most frequent words (in numbers of the retweets) during the first and second election rounds. In the first round (Figure 3.3.a) we note that elections related words are predominant, such as *vote* and *president*. Interestingly, the term electronic voting machine is also one of the most used terms, probably due to the suspicion about its credibility raised by the supporters of Jair Bolsonaro candidate.[8]

In the word cloud of the second round (Figure 3.3.b), we observe the presence of words celebrating the victory of the Luis Inácio Lula da Silva, such as *victory*, *lulapresident2022*, *democracy*. We can also highlight the term *electronic voting machine*, a topic that came up frequently in this election since some people started to question the

[5] These candidates accounted for over 98% of all votes.
[6] https://www.cnnbrasil.com.br/eleicoes/2022/apuracao/presidente/br/primeiro-turno/
[7] $RT/TW = Retweets/Tweets$
[8] https://www.nytimes.com/2022/09/29/world/americas/election-bolsonaro-brazil-fraud.html?smid=url-share

(a) First Round. (b) Second Round.

Figure 3.3: Retweets' word clouds.

credibility of the voting process in Brazil, prompting government entities to declare that the process is safe and reliable in order to combat the rumors and fake news that had emerged[9]. Finally, it is worth noting the presence of the word *Northeast* in both ballots, which is a Brazilian region where Lula has many supporters.[10]

3.3 Network Modeling

To investigate the 2022 election debate on Twitter, we used a network model known as *media-centric* network, which connects users who have shared similar content [78, 79, 26, 60]. By analyzing the characteristics of such media-centric networks, we can determine which actors contribute the most to content distribution.

We built two graphs, where each graph represents a media-centric network capturing the user sharing patterns during each election round. In each graph $G(V, E)$, a node $v \in V$ corresponds to a user who retweeted a tweet, and an undirected edge $e{=}(v_i,\, v_j)$ is included in E if the users corresponding to v_i and v_j shared the same tweet (exactly the same textual content) at least once. The weight of e is the number of tweets both users shared in common. The total amount of retweets shared across each network is the sum of all edge weights.

[9]https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=494958&ori=1

[10]https://www.cnnbrasil.com.br/politica/nordeste-e-a-unica-regiao-em-que-lula-obteve-mais-votos-que-bolsonaro-confira/, https://www.theguardian.com/world/2022/nov/01/brazil-election-how-lula-won-the-runoff-from-sao-paulo-to-the-north-east

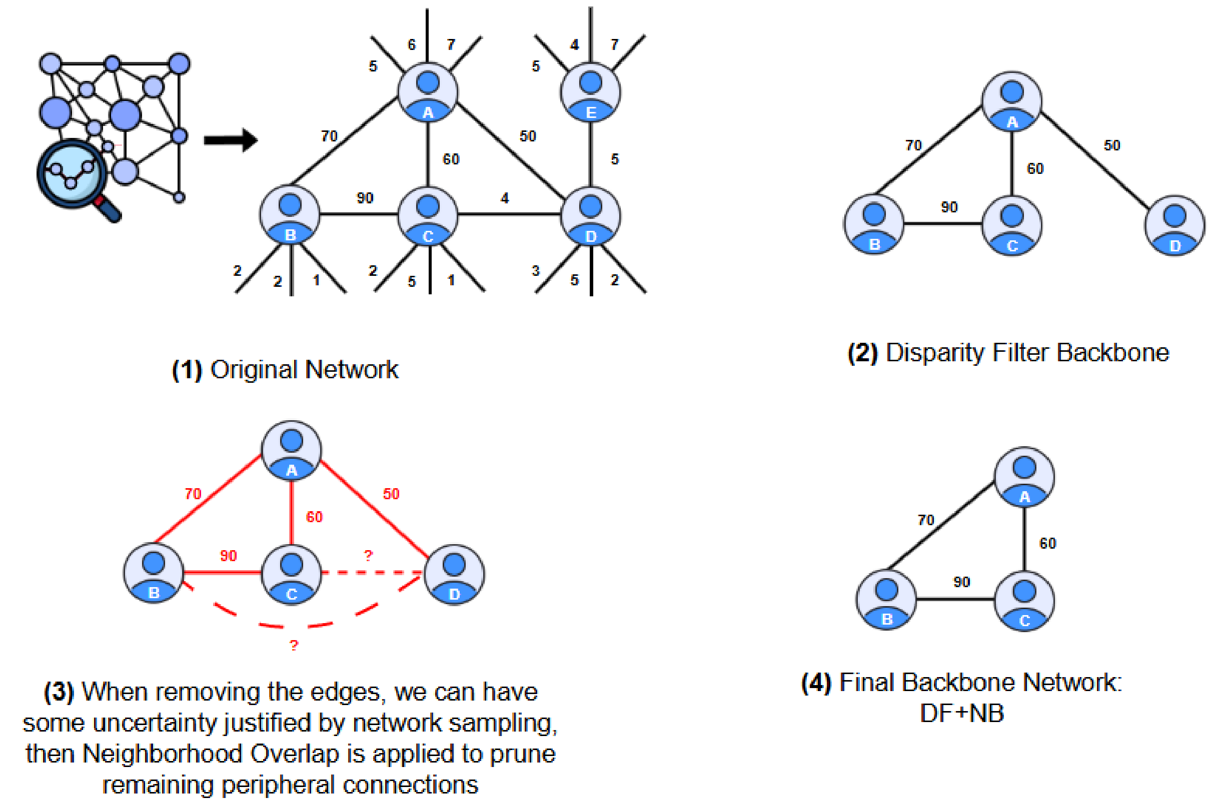


Figure 3.4: Illustration of Backbone Extraction Methods considering potential sample limitation scenarios.

3.4 Backbone Extraction

The primary objective of the media-centric network model is to provide a comprehensive overview of the distribution of content on Twitter and to enable the identification of groups of users who frequently share the same information. To refine this analysis, we focus on identifying pairs of users whose sharing patterns demonstrate strong, non-random associations. In other words, we filter out noisy or sporadic edges and retain only the pairs of users whose shared tweets deviate significantly from the expected number of shared content. These filtered users, referred to as **key users**, represent the core of the information dissemination network and are likely to be the primary drivers of the election debate on Twitter. By isolating this group, we aim to understand the dynamics of content diffusion and the role of influential users in shaping the political discourse during the election period.

To achieve that, we applied two methods to extract the network backbone to filter out weaker edges, leaving only stronger edges: Disparity Filter (DF) [98] and Disparity Filter with Neighborhood Overlap (DF+NB), which combines the Disparity Filter method with the concept of Neighborhood Overlap [60]. In a nutshell, Disparity Filter

(DF) assumes a reference model for users who share content independently of the others with a uniform distribution of edge weights at the corresponding vertices. Therefore, an edge (v_i, v_j) is left in the backbone if its weight (from a statistical point of view) strongly deviates from this reference model for both v_i and v_j. This method effectively highlights edges that exhibit consistent and repeated behavior between user pairs. With the Neighborhood Overlap (NB) technique, DF+NB goes further by removing peripheral and bridge connections and focusing on edges between users with common neighbors that also exhibit similar content dissemination patterns. After applying the two aforementioned methods, we limit our analysis to the DF+NB backbone extraction method, as DF+NB has proven to be more effective in high-noise scenarios, which is typical in Twitter data analysis [60].

Figure 3.4 illustrates an example of an original network and its backbones after applying the extraction methods. Starting with step (1) in the figure, we observe the original network where, from the perspective of a given node, there are both strong edges (with higher weights) and weaker ones. For instance, from the perspective of node D, edges have weights ranging from 2 to 50. In step (2), the Disparity Filter is applied. As previously explained, this method retains edges that significantly deviate from others in the local context of a node. Focusing on node D, only the edge with a weight of 50 is preserved because it clearly stands out compared to the others, which have weights between 2 and 5.

However, two key issues arise, motivating the use of the combined Disparity Filter and Neighborhood Overlap (DF+NB) approach proposed in [60]. First, considering node D again, the edge with a weight of 4 might actually have a higher weight (e.g., 30 [11]), which could lead to its retention in the backbone if data were complete. Second, the edge between nodes D and B could exist in the real network, indicating that D is more strongly connected, but it might not appear in the sampled data due to limitations imposed by the Twitter API. Since uncertainty is present, we adopted a more conservative approach, prioritizing the sampled network's provisioned data and removing weak connections in such cases. To address these potential effects of data sampling, step (3) shows the backbone of the Disparity Filter, where the Neighborhood Overlap method is applied to further refine the backbone. The idea is to remove edges from the Disparity Filter backbone that do not belong to strongly connected parts of the network. Specifically, it retains only edges that connect nodes with a sufficiently large number of common neighbors, based on the Disparity Filter backbone structure, to refine and prune any leftover peripheral connections. Finally, in step (4), the resulting backbone of DF+NB provides a conservative representation of the original network, highlighting only the strongest and most meaningful connections. This refined structure facilitates the identification of tightly connected,

[11] Due to data sampling limitations, some user actions and connections may not be captured, resulting in a lower weight assigned to a given link.

non-trivial topology within the sampled network, ensuring a robust depiction of the underlying topology despite the limitations of partial data. By doing so, the method focuses on distinguishing users with high interaction and significant connections from peripheral users, who exhibit weaker ties and less relevance to the core of the network.

To parameterize the Disparity Filter (DF) method, we set $\alpha = 0.05$, which serves as the significance level for testing the hypothesis that users are sharing the same content in a non-random manner [45]. This parameter represents the p-value used to test the null hypothesis that the sharing behavior of users is random and follows a uniform distribution. Deviations from this assumption indicate significant, non-random patterns of shared content. For the filter based on the neighborhood overlap metric, we use a threshold corresponding to the 95^{th} percentile of the neighborhood overlap distribution, ensuring that only the most strongly connected edges are retained.

After extracting the backbone of each graph, we applied the widely used Louvain community detection algorithm [13] to identify and analyze patterns of user groupings and their organization in each backbone. The goal of the Louvain algorithm is to maximize the modularity of communities.

This is a key metric that represents the density of connections within the communities compared to a hypothetical random network. The values for modularity range from -0.5 to +1, with higher values (above 0.3) indicating well-defined community structures [76]. In this study, we used the modularity metric of the communities as a proxy to assess the effectiveness of the backbone extraction technique in preserving the core of user interactions, revealing a more cohesive network topology.

3.5 Content Analysis

Besides identifying the key users who shared a high volume of similar information during the two election rounds, we are also interested in characterizing what they were talking about. For example, content analysis can help us understand how certain content attracts more attention from users, potentially contributing to its virality on a particular social network.

Here, we focus on retweets to analyze content dissemination and user behavior in the campaign context. We conducted the experiments using the original textual content of the messages, later scaling it to the total number of corresponding retweets in the network to optimize computational resource usage. To achieve this, we analyzed the text they shared from three different (but complementary) perspectives: (i) Topic Extraction; (ii) Sentiment Analysis and; (iii) Psycholinguistic Analysis.

Topic Extraction. During a debate, participants address various topics, with certain recurring themes capturing the interest of social media groups. These popular themes reflect the positions, opinions, and interests of the individuals involved, offering insights into the prevailing discourse.

To identify the discussed topics, we applied the BERTopic model [47], which has proven to be one of the most effective methods for analyzing short-text data [34]. The BERTopic process begins by converting a collection of retweet texts into vector representations using BERTimbau, a Portuguese pre-trained language model, as the foundation for the Transformer-based encoding [103, 42, 87, 68]. BERTimbau is derived from the original Bidirectional Encoder Representations from Transformers (BERT) model and has been fine-tuned to better capture semantic nuances in Portuguese text. Once vector representations are generated, Uniform Manifold Approximation and Projection (UMAP) is applied to reduce the dimensionality of these embeddings. This step is crucial for improving the efficiency of the subsequent clustering process while preserving the essential semantic relationships in the data. After dimensionality reduction, the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm is used to group similar low-dimensional vectors into clusters based on their semantic similarity.

To extract representative words for each identified topic, BERTopic employs class-based Term Frequency-Inverse Document Frequency (c-TF-IDF). Unlike the traditional TF-IDF, which measures word importance across an entire corpus [1], c-TF-IDF is adapted for topic modeling by computing term importance at the cluster level rather than at the document level. This approach enables better differentiation of topic-specific words by emphasizing terms that are characteristic of a particular cluster while downweighting those common across multiple topics. The c-TF-IDF score is calculated as:

$$W_{x,C} = \|\mathrm{tf}_{x,C}\| \times \log\left(1 + \frac{A}{f_x}\right) \tag{3.1}$$

where:

- $W_{x,C}$ is the c-TF-IDF score for word x in class (cluster) C,
- $\|\mathrm{tf}_{x,C}\|$ is the frequency of word x in class C, often normalized,
- f_x is the frequency of word x across all classes,
- A is the average number of words per class.

This formulation ensures that words that clearly stand for a certain topic are given a higher weighting, which enables a clearer interpretation of the extracted topics. After topic extraction, maximum marginal relevance (MMR) was applied to increase the diversity of extracted keywords. MMR ensures that the most relevant terms are selected for each topic while minimizing redundancy through a balance between relevance and diversity. This is done by comparing the word embeddings with the topic embeddings, which reduces the occurrence of synonyms and repetitive terms, making the topics more concise and easier to interpret.

For the parameterization, we followed the recommendations in the BERTopic documentation to find a balanced compromise between the number of topics and the size of the dataset.[12] As a result, we obtained the following parameterization: the number of neighbors and the component parameters required by UMAP were set to 10 and 5, respectively. The minimum topic size was set to 5, which controls the minimum number of unique retweets on a topic. The minimum number of words required to visualize topic content was set to 20 to inform broader content about the topics of the issues and their relationship to the elections. Finally, the MMR was set to 0.6 (on a scale of 0 to 1).

Psycholinguistic Analysis. We delve deeper into the content analysis by understanding the psycholinguistic properties of the shared text. We rely on the Linguistic Inquiry and Word Count (LIWC) lexicon [106] to categorize words in the text in linguistic style, affective and cognitive attributes. Using a predefined dictionary of words and linguistic categories, LIWC classifies words and terms in a given text into a variety of hierarchical attributes related to psychosocial dimensions, linguistic style and cognitive concepts (e.g., word categories, personal concerns, emotions). We compute the average frequency of the attributes over the retweets. In our data, we identify all 64 attributes, out of the available in LIWC's Portuguese dictionary. We rank the attributes according to their capacity to discriminate the retweets, estimated by the Gini Coefficient [113] and we use the top-10 to create heatmaps that can better highlight attributes associated with our dataset. The cells of the heatmap in a column show the relative deviation of the respective attribute for the respective keyword from the other keywords. In other words, each column (attribute) is normalized by the z-score, – i.e. $z = (x - mean)/std$. This means that each value is subtracted from the average of the column and then divided by the standard deviation of the column. The locations are color-coded red (resp. blue) if the attribute is present more (resp. less) than the average.

Sentiment Analysis. People tend to express themselves based on their particular feelings about a situation. Positive feelings typically express approval and agreement, while negative feelings can emphasize sadness and indignation. We then analyze whether the expressed opinion in the retweets is mostly positive, negative, or neutral. To this end, we used the XLM-RoBERTa[13] (Cross Lingual Language Model - Robustly Optimized BERT-Pretraining Approach) model available in the Hugging Face library. This model is

[12] https://maartengr.github.io/BERTopic/
[13] https://huggingface.co/docs/transformers/model_doc/xlm-roberta

a fine-tuned version of RoBERTa [61] trained on a Twitter database of 198 million multilingual tweets, with Portuguese being the second most frequently occurring language in these tweets. The XLM-RoBERTa model provides the probabilities of categorizing a given tweet as positive, negative or neutral. In our analysis, we classify the sentiment of a tweet as the class with the highest probability assigned by the model.

Part-of-Speech Tagging diversity. In order to explore the textual aspects of the messages with different virality types used in the classification process detailed in the next section, we decided to perform a Part-of-Speech Tagging (POS Tag) diversity analysis [66, 63]. This technique consists of an NLP method that gives each word in a sentence a grammatical category or tag, like noun, verb, adjective, etc. These identifications are used to define the textual syntactic functions and connections of words in their respective sentences. The first step was to calculate the ratio of distinct POS to the total number of tokens (words) in a text comment, to measure how diverse was our POS counting when compared to the text's length for the viral and non-viral classes. Next, estimated the 95% confidence intervals for the POS tag diversity values of two categories of virality.

Type-Token Ratio (TTR). The Type-Token Ratio (TTR) was used to assess the lexical diversity of the messages in the viral and non-viral sets, providing a view into the complexity and variability of the vocabulary across the two classes. A higher TTR indicates greater lexical diversity, suggesting that a class contains a wider range of unique words relative to its total word count, while a lower TTR suggests a more repetitive and constrained vocabulary. This measure is particularly relevant for comparing the semantic richness of viral and non-viral tweets, as variations in vocabulary complexity may reflect differences in writing style, informativeness, or audience engagement. The TTR is computed by dividing the total number of distinct words (types) by the total number of words in a text (tokens) [91].

3.6 Virality Classification Task

Content shared on OSMPs can quickly reach a vast audience, often going viral. Understanding the factors contributing to this phenomenon is a challenging task [15, 28]. In this work, we take the initial steps in this direction.

To that end, we propose the following classification task. We define the problem of distinguishing viral messages as a binary classification task that, given a message, classi-

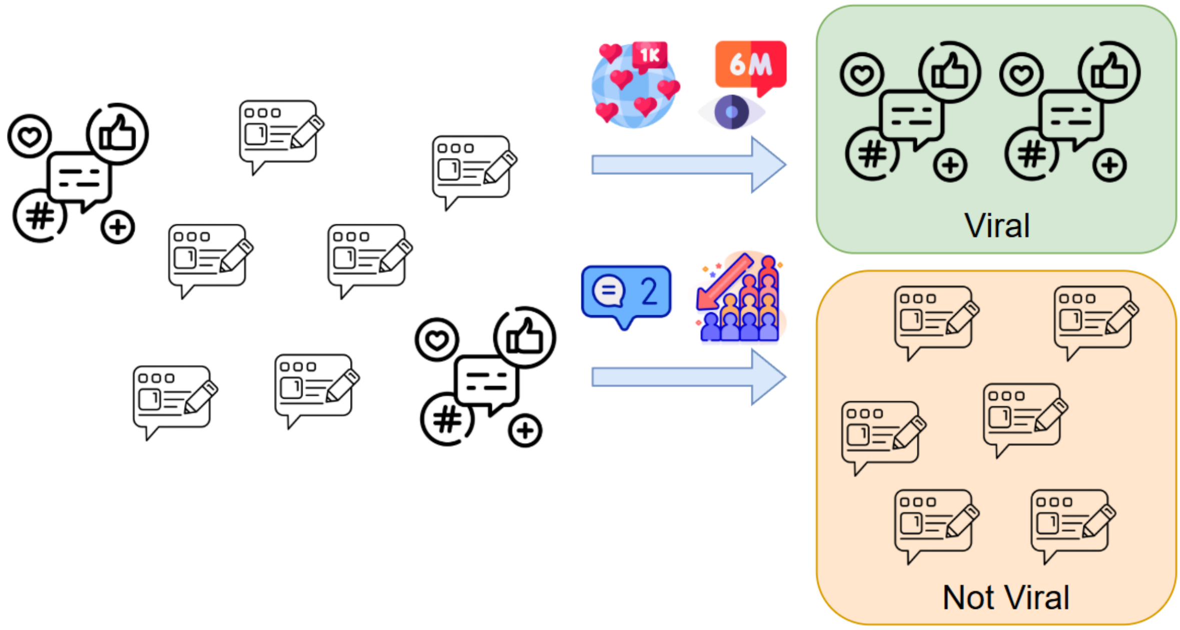


Figure 3.5: Illustration of the binary classification process proposed.

fies whether it is a viral or non-viral content. Formally, let $\mathcal{X}$ represent a set of messages and $\mathcal{Y} = \{0, 1\}$ the label space, where 1 corresponds to viral and 0 to non-viral. The goal of binary classification is to learn a function $f : \mathcal{X} \to \mathcal{Y}$ using a training set of m instances $\{(x_i, y_i)|1 \leq i \leq m\}$. Here, each instance $x_i \in \mathcal{X}$ represents a message described by its features, and $y_i \in \mathcal{Y}$ denotes the messages' corresponding target value (i.e., viral or non-viral). This process is exemplified by Figure 3.5, where messages can be gathered in a social media with different characteristics. Those which can reach high metrics of shares, comments, likes or visualization for example end up labeled as viral since they achieve a wider audience. Meanwhile, common messages usually don't perform that high, having less engagement and therefore being non-viral.

Such task requires defining what is a viral or non-viral content. Using the Cumulative Distribution Function (CDF) of the number of retweets per tweet, we identified the 10% and 90% cutoff points of the distribution. Tweets below the 10% threshold were classified as non-viral, while those above the 90% threshold were considered viral.

3.6.1 Classification Models

To classify a tweet as viral or non-viral, we designed two classification models based on BERTimbau, a language model specifically pre-trained for the Portuguese language

[103]. BERTimbau is built upon the BERT (Bidirectional Encoder Representations from Transformers) architecture [30], which introduced a revolutionary bidirectional approach to contextual word representation in natural language processing (NLP) tasks [110]. Thus, BERTimbau follows the Transformer-based architecture, in which multiple identical layers process text representations iteratively, refining contextualized embeddings at increasing levels of abstraction.

One of the main advantages of Transformer-based models like BERT and BERTimbau is their ability to generate contextualized word embeddings. Unlike traditional static embeddings, which assign a fixed vector to each word, BERTimbau dynamically adjusts word representations based on the surrounding context in a sentence. This contextual awareness significantly enhances language understanding, making it particularly effective for handling the complexities of Portuguese. BERTimbau was pre-trained on large-scale Portuguese-language datasets, allowing it to capture morphological, syntactic, and semantic patterns specific to the language. This pre-training process equips the model to handle linguistic phenomena such as rich inflectional morphology, verb conjugation complexity, and syntactic variations. By fine-tuning BERTimbau for our classification task, we adapt its pre-learned linguistic knowledge to the specific challenge of distinguishing between viral and non-viral tweets, enabling a more nuanced analysis of social media text. We then fine-tuned BERTimbau to perform our binary classification task, which involves determining whether a tweet is viral or non-viral. This task required adapting the model to recognize patterns in the text that correlate with virality, leveraging both the rich contextual embeddings generated by BERTimbau and additional features specific to our dataset. The two models differ in the features they use and the architecture they employ to incorporate these features into the classification process, as explained as follows.

The first model, referred to as the Baseline BERTimbau model, uses the standard pre-trained BERTimbau architecture fine-tuned for binary classification. The only input to this model is the text of the tweet, which is processed by BERTimbau to generate contextual embeddings. The representation of the tweet is extracted from the `[CLS]` token embedding, which is then passed through a fully connected transformation layer and activated using a Tanh function. The resulting vector is processed by an additional fully connected classification layer with a sigmoid activation function to output the probability of the tweet being viral. This model serves as a baseline to evaluate the effectiveness of incorporating additional features. The second model, which we named BERTimbau+B, extends the baseline model by incorporating an additional feature derived from the topological structure of the dissemination network. Specifically, this feature indicates whether a given tweet was shared by users identified in the backbone of the network, as extracted in earlier steps of our analysis. Recall that, the backbone represents the core structure of the dissemination network, capturing users and messages with the strongest connectivity and influence, which are likely to drive virality. The inclusion of this feature introduces a structural perspective to the classification task, as it carries information regarding non-random dissemination patterns and potential coordination between users [77, 111].

To integrate this additional feature, the architecture of BERTimbau was modified. After generating text embeddings using the standard BERTimbau model, a secondary input was introduced to accommodate the categorical backbone feature (*0* or *1*). This feature indicates whether the tweet was disseminated from users in the backbone (*value 1*) or from the broader network (*value 0*). We considered this information as a possible interesting input since content that were shared by the backbone users could have a high dissemination rate, which could eventually lead to higher engagement metrics and virality, exactly what we aimed to classify. Since BERTimbau produces high-dimensional embeddings, the backbone feature must be properly scaled to avoid dominance imbalance in the concatenation process. To achieve this, we apply Layer Normalization (L-Norm) [22, 6], which normalizes only the categorical feature using:

$$\text{L-Norm}(x) = \frac{x - \mu}{\sigma + \epsilon} \cdot \gamma + \beta \tag{3.2}$$

where x represents the input value (in this case, the backbone feature), μ and σ are the mean and standard deviation computed across the batch, and γ and β are learned parameters that allow the model to adjust the normalized values dynamically. The small constant ϵ prevents division by zero. After normalization, the embeddings from BERTimbau and the normalized backbone feature are concatenated into a single vector, which is then passed through additional fully connected layers before classification. Before computing the final logits, dropout is applied to prevent overfitting. The final classification is performed using Binary Cross-Entropy Loss (BCEWithLogitsLoss) [86], which combines a sigmoid activation function with a cross-entropy loss calculation. This avoids the need for explicitly applying a sigmoid function to the output layer, as the loss function itself incorporates the sigmoid transformation internally. Binary Cross-Entropy (BCE) is widely used in binary classification tasks, as it measures the difference between predicted probabilities and actual binary labels. The BCE loss function is defined as follows:

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i) \right] \tag{3.3}$$

where:

- $\mathcal{L}_{BCE}$ represents the Binary Cross-Entropy loss.
- N is the number of examples in the training batch.
- y_i is the ground-truth label for the i-th sample, where $y_i = 1$ for a viral tweet and $y_i = 0$ for a non-viral tweet.

- $\hat{y}_i$ is the predicted probability for the positive class (viral tweet), obtained from the model.
- $\log \hat{y}_i$ penalizes incorrect predictions when the true label is 1.
- $\log(1 - \hat{y}_i)$ penalizes incorrect predictions when the true label is 0.

In this formulation, higher penalties are assigned to confident yet incorrect predictions, encouraging the model to adjust its probability estimates accordingly. Since BCEWithLogitsLoss directly integrates the sigmoid activation function into the computation, it ensures numerical stability and simplifies the model's implementation by allowing the logits (raw outputs of the last layer) to be passed directly into the loss function [109].

In sum, while text embeddings capture linguistic and semantic properties, they may not fully reflect external factors influencing virality, such as network dynamics or patterns of coordinated dissemination from a set of key users. Thus, the backbone feature provides a complementary perspective by representing the structural importance of a tweet within the dissemination network. Since the backbone extraction method identifies users whose sharing behavior is *not random* but follows persistent, high-connectivity patterns, its inclusion in the model allows us to test whether such structural signals contribute to improving classification performance. The classification task remains binary, predicting whether a tweet is viral (*1*) or non-viral (*0*). By comparing the performance of the **Baseline BERTimbau** model with **BERTimbau+B**, we evaluate the contribution of the backbone feature to the overall classification accuracy. Thus, this dual-input architecture allows us to investigate the potential of combining linguistic and topological features in predicting virality in online social media platforms.

3.6.2 Evaluation Protocol

To evaluate the performance of our classifiers in distinguishing viral from non-viral tweets, we employed a rigid protocol. First, we divided the dataset into training (80%), validation (10%), and test (10%) sets. These splits were fixed and identical for both models, ensuring consistency across evaluations. The validation set was used for hyperparameter tuning and threshold optimization, while the test set was strictly reserved for final evaluation, preventing any data leakage and ensuring an unbiased assessment.

The models were fine-tuned using grid search, a systematic approach that evaluates all possible combinations of predefined hyperparameter values. The goal of this search was to determine the best configuration that minimizes the Binary Cross-Entropy (BCE) Loss explained in Equation 3.3, which ensures that the predicted probabilities align as closely as possible with the true labels. We selected three critical hyperparameters based on prior work [10, 4, 105, 104, 101]: the learning rate, dropout rate, and weight decay. The learning rate controls the step size during gradient updates, and an improper setting may cause either slow convergence or overshooting of the optimal solution. The dropout rate randomly deactivates a fraction of neurons during training, reducing overfitting and enhancing generalization. Weight decay adds a regularization penalty based on the magnitude of model weights, preventing the model from becoming overly reliant on specific features.

The total grid search space consisted of $3 \times 4 \times 3 = 36$ configurations. Each configuration was trained and evaluated on the validation set, and the one with the lowest validation loss was selected as the best. To prevent overfitting and unnecessary computational costs, we applied early stopping with patience set to 3 epochs [30, 85, 46, 32], meaning that training would halt if the validation loss did not improve for three consecutive epochs. The batch size was fixed at 4 due to memory constraints. A crucial step in the evaluation process was the optimization of the classification threshold. By default, binary classification models use a threshold of 0.5, but this may not always yield the best results, particularly in imbalanced datasets where false positives and false negatives have different costs. To ensure optimal classification, we adjusted the decision threshold using the validation set. This was done by analyzing the precision-recall curve, computing the F1-score at different thresholds, and selecting the threshold that maximized the F1-score [71].

The final hyperparameter configurations for each model were as follows. The Baseline BERTimbau model was optimized using a learning rate of $4e-5$, a dropout rate of 0.1, and a weight decay of 0.1. The BERTimbau+B model, which integrates topological information, achieved the best results with a learning rate of $2e-5$, a dropout rate of 0.4, and a weight decay of 0.1. These hyperparameters were then used to retrain the models on the combined training and validation sets (90% of the data) before final testing. For the final evaluation, we used several standard classification metrics [18, 73]. Accuracy measures the overall proportion of correctly classified instances:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{3.4}$$

where TP (true positives) represents the number of viral tweets correctly classified as viral, TN (true negatives) refers to non-viral tweets correctly classified as non-viral, FP (false positives) corresponds to non-viral tweets incorrectly classified as viral, and FN (false negatives) represents viral tweets that the model mistakenly classified as non-viral.

Precision evaluates how many of the tweets predicted as viral were actually viral:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.5}$$

Recall, also referred to as sensitivity, measures how well the model identifies actual viral tweets:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.6}$$

The F1-score is the harmonic mean of precision and recall, providing a balance between the two:

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.7}$$

Additionally, we considered the Area Under the Receiver Operating Characteristic Curve (AUC-ROC), a widely used metric in classification tasks that evaluates the model's ability to distinguish between viral and non-viral tweets across various classification thresholds. The ROC curve plots the true positive rate (recall) against the false positive rate at different decision thresholds, illustrating the trade-off between sensitivity and specificity. The AUC value, which ranges from 0 to 1, quantifies the overall performance of the model. A value of 0.5 indicates no discriminatory power, equivalent to random guessing, while a value closer to 1 suggests a model with strong classification capability. In our case, higher AUC values indicate that the model can effectively separate viral from non-viral tweets.

After determining the best models, the final evaluation was conducted exclusively on the test set, which had not been used during training or threshold selection. To ensure a robust statistical assessment, we applied bootstrapping on the test set, a resampling technique that estimates the variability of evaluation metrics by repeatedly sampling with replacement from the test set [72, 52]. A total of 300 bootstrap paired iterations were performed, enabling the computation of confidence intervals for accuracy, precision, recall, and F1-score. This process provided a more in-depth analysis of the models' stability and ensured that the reported results were not overly dependent on specific test set characteristics.

To compare the results, we rely on statistical tests, which must be chosen based on the characteristics of the data. Since we are comparing two models using the same set of observations, a paired statistical test is appropriate. In such cases, the choice between a parametric and a non-parametric test depends on whether the differences between paired samples follow a normal distribution. To verify this assumption, we conducted a normality test on the bootstrapped paired differences. Given the large sample size, it is essential to confirm whether the assumption of normality holds before applying parametric tests [23]. Specifically, we applied the Kolmogorov-Smirnov (KS) test to assess whether the distribution of the differences follows a normal distribution [27, 23]. If the p-value of the KS test surpasses 0.05, we fail to reject the null hypothesis, indicating that the differences are normally distributed, allowing us to apply a parametric test such as the paired t-

test. Otherwise, we reject normality and proceed with a non-parametric test, such as the Wilcoxon Signed-Rank Test [23].

Since the normality assumption does not hold in this case[14], we employ the Wilcoxon Signed-Rank Test for the zero-mean hypothesis. This test is a non-parametric alternative to the paired t-test, ranking the absolute values of the paired differences and assessing whether the median difference is statistically different from zero. Unlike the t-test, which assumes that differences are normally distributed, Wilcoxon considers only the signs and ranks of the differences, making it more robust to non-normal distributions. Then, for each bootstrap iteration, the difference between the evaluation metric values of Model A (BERTimbau) and Model B (BERTimbau+B) is computed as:

$$d_i = X_{A,i} - X_{B,i}, \quad \text{for } i = 1, \ldots, 300 \tag{3.8}$$

The Wilcoxon test assesses whether the median difference is zero:

$$H_0 : \text{Median}(d_i) = 0 \tag{3.9}$$

where d_i represents the differences between the paired observations. The test ranks the absolute differences, ignoring zeros, and computes a test statistic W, given by:

$$W = \sum \text{ranks of positive differences} - \sum \text{ranks of negative differences} \tag{3.10}$$

For large samples, the statistic W follows an approximately normal distribution and is standardized as:

$$Z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \tag{3.11}$$

where n is the number of nonzero differences. The computed Z-value is compared to the standard normal distribution to determine statistical significance. If the p-value is below the significance threshold (e.g., 0.05), we reject H_0 and conclude that there is a significant difference between the two models. Otherwise, we fail to reject H_0, indicating insufficient evidence to claim a difference in performance.

[14]The results of the KS test, which are listed in the Results section, indicate a significant deviation from normality. Therefore, the Wilcoxon Signed-Rank Test was performed instead of the paired t-test.

Chapter 4

Results

In this chapter, we present our findings to answer the proposed research questions. The first step was to describe the topological structure emerging from our data (Section 4.1), using network analysis techniques and metrics, as well as to present the networks derived from backbone extraction. Then, in Section 4.2, we characterized the most disseminated topics, highlighting the insights derived from the backbones of retweets for each election round. Finally, we discuss the classification models created and the key findings when considering the input of virality features (Section 4.3).

4.1 Topological Analysis

In this section, we aim to answer our ***RQ1: Is it possible to identify a group of users who recurrently spread similar content on Twitter during the two rounds of the Brazilian elections of 2022? Do they remain active and persistent in the analyzed time intervals ?*** To that end, we built media-centric networks in order to detect closely connected user groups that frequently share the same content.

Table 4.1 shows the analysis of these networks and reveals different structures between the complete networks and their respective backbones. DF+NB significantly reduces the weak links in both graphs, leading to a reduction in the number of nodes, called *key users*, and their edges.

Looking at the first round, the complete network consists of 71,585 nodes and 192,539,317 edges with an average degree of 5,379.32. By applying the DF+NB method, however, this network is reduced to 5,192 nodes and 137,165 edges. In the second round, on the other hand, the complete network consists of 60,288 nodes and 152,603,351 edges with an average degree of 5,062.48. Applying the DF+NB method reduces this network to 3,704 nodes and 33,480 edges. Both backbones highlight the users who are most active in sharing content: They retweeted approximately 16% and 10% of the tweets produced.

Backbones allow us to better understand the properties of the network core. An im-

Table 4.1: Characterization of the topology of the networks and backbones.

Network	Date	# Nodes (%)	# Edges (%)	Tweets	Retweets	Avg. Degree	Density	Avg. Clustering	# Components	Size Giant Comp.	# Comm.	# Comm. With > 10 Users	Modul.
Complete	1st Round	71,585	192,539,317	8,774	409,956	5,379.32	0.075	0.6955	13	71,559	25	9	0.38
DF+NB	1st Round	5,192 (7.25%)	137,165 (0.0712%)	944	67,521	52.84	0.010	0.4261	433	3,658	457	18	0.42
Complete	2nd Round	60,288	152,603,351	5,835	331,175	5,062.48	0.084	0.6958	14	60,258	20	7	0.35
DF+NB	2nd Round	3,704 (6.14%)	33,480 (0.0219%)	556	33,637	18.08	0.005	0.4682	308	2,799	335	20	0.68

portant point is how users organize themselves into communities and how these structures influence the dissemination of information. To this end, we calculated some community-related metrics, as shown in Table 4.1. Our results show the increase in the modularity metric. This indicates highly connected and structured community networks and shows the potential of DF+NB in filtering out noise in the data. Furthermore, this increase in modularity shows the growing complexity of the interaction network across the two rounds of voting, which is further highlighted by the increase in the number of larger communities in the resulting topologies.

Takeaway. The identification of backbone networks reveals a group of users who engage through stronger interactions. This structured network can lead *key users* to act more coherently in promoting certain content, thus influencing the discourse.

4.2 Content Analysis

We now turn our attention to answer our ***RQ2: Do the textual characteristics of the content change based on the user's level of engagement ?*** To achieve this, we categorized our content analysis into three types of dissemination regarding user's level of engagement: *widespread dissemination*, which examines the entire media-centric networks across both election rounds; *key users' dissemination*, focusing on retweets shared exclusively by users within the backbone networks; and *persistent users' dissemination*, targeting users present in the backbones of both election rounds. While the first approach offers a comprehensive overview of the debate on the topic, the latter two delve into the core of the discussion, filtering out weaker interactions that could obscure the primary issues and concerns surrounding the Brazilian elections.

4.2.1 Widespread Dissemination

We first look at the topics shared by all Twitter users in our data set. This analysis was performed using the BERTopic model for the tweets that were disseminated during the two election rounds. Initially, 192 topics were identified through the application of BERTopic. However, to focus our analysis on the most influential discussions, we prioritize the 20 most popular topics discussed, measured by the number of retweets, which can provide insight into the prevailing topics of Twitter users. Table 4.2 provides an overview of the final topics, including the message metrics, the top 10 most discriminating topic words and a brief description of each topic.

Topics 1, 2, 3, 7, 11, 13, 14 and 18 mainly refer to Lula's victory. Topics 1 and 18 highlight the Brazilian regions where Lula was the winner of the election, as well as the fact that the then President of the United States, Joe Biden, was one of the first to internationally recognize and congratulate Lula's victory.[1] To illustrate, the content with the most retweets in topic 18 (4,419 retweets) is: *"Lula leads in all states of the Northeast"*. In addition, the discussions also celebrate *the victory of democracy*, reflecting disapproval of the former Brazilian president's governance.

Some topics reveal the main issues that attracted attention during the elections. For instance, Topic 15 focuses on a recurring theme echoed by far-right supporters: the possibility of electoral fraud. Topics 4 and 17 highlight the proliferation of discussions about the concept of the *third way*, an alternative proposed by certain individuals aimed at circumventing the polarization between the two leading candidates, Lula and Bolsonaro. This approach encourages voters to consider voting for candidates such as Simone Tebet and Ciro Gomes[2]. Topic 6 reflects the controversial decisions made by Jair Bolsonaro's government during the COVID-19 pandemic.[3]

In addition to the discussion about the presidential candidates, our data also shows retweets about the elections of state governments and members of parliament. Topics 8 and 16 focus on the discussion about the candidates of the state of Minas Gerais. Topic 16 in particular raised questions about the apparent contradictions in voting patterns, where state voters elected Lula as president while voting for predominantly far-right candidates for state government, deputies and senators. Furthermore, topics 9 and 20 are related to the increase in the number of female candidates elected in the 2022 elections.

Figure 4.1 shows the percentage of retweets per topic in each round. As expected, some topics were more represented in one round of voting than the other, due to the nature

[1] https://www.whitehouse.gov/briefing-room/statements-releases/2022/10/30/statement-by-president-joe-biden-congratulating-luiz-inacio-lula-da-silva-as-president-of-brazil/

[2] https://www.lemonde.fr/en/international/article/2022/09/28/brazil-election-third-way-candidates-gain-little-ground-against-lula-and-bolsonaro_5998463_4.html

[3] https://www.kcl.ac.uk/covid-19-in-brazil-how-jair-bolsonaro-created-a-calamity

Table 4.2: Top discussion topics found on Twitter.

ID	# Tweets	# Retweets	Most Discriminative Words	Description
1	587	48,978	inácio, silva, luiz, president, biden, elected, victory, new, brazil, luis	Discusses President Lula's victory in the election in the second round and the possibility of his upcoming victory during the first round. Cites Joe Biden, president of the United States, who was one of the first international figures to recognize Lula's election.
2	473	30,496	elections2022, turn, turned, lulaonFirstRound13, northeast, elections2022, lulapresident1, elections2022, turnaround, lulinha	Regarding the turnaround in votes that Lula had, when the votes from the northeast began to be counted.
3	76	28,892	supporters, celebrating, turnaround, victory, party, celebrate, streets, brasília, against, petista	Refers to Lula's victory and the voters' celebration.
4	167	24,892	way, third, fault, simone, chance, second, have, ciro, you, voted	Mentions the discourse of a third way of opposition to Lula and Bolsonaro, quoting candidates Ciro Gomes and Simone Tebet.
5	97	22,882	history, times, time, re-election, since, 1st, president, re-elect, term, succeeds	Topic that debates about the possibility of reelection of Bolsonaro and the fact of him being the first Brazilian president to not be re-elected.
6	145	18,158	mourning, thousand, pandemic, 700, covid, deaths, dead, people, lost, during	Issues and fatalities that occurred during Bolsonaro's administration in the COVID-19 pandemic period.
7	293	14,738	over, nightmare, goodbye, won, bye, lulapresident2022, end, well, above, finally	Electoral opponents of the Bolsonaro government celebrating the election results.
8	65	13,181	deputy, federal, paulo, ferreira, voted, mg, nikolas, elected, paraná, senator	Mentions the State's Elections for House of Representatives and Senate.
9	42	12,755	elected, federal, woman, first, paulo, historic, senate, all, damares, against	Comments on the electoral victory of women for the position of congresswomen.
10	25	11,790	stupid, general, voting, others, vote, for him, enough, regions, minas, right	Criticizes voters for their decision to vote on polemic candidates from far right, including Bolsonaro.
11	124	11,604	thank you, thank you, congratulations, god, good, democracy, country, all, sir, above	People celebrating and thanking the Brazilian democracy regime with Lula's election.
12	60	11,202	lost, neymar, fall, equal, lose, falling, stick, cup, this, in this	Mentions terms related to the World Cup, which took place close to the election period.
13	40	11,181	lo, lulapresident2022, let's go, turn, big, lulapresident1, victory, moment, luiz, listen	Talks about Lula's victory and his first speech.
14	53	11,140	urgent, missing, only, less, missing, thousand, victory, lulaonFirstRound13, elections2022, give	Refers to the first round when Lula led with 48.43% of the votes and almost was elected and the victory of Lula in second round.
15	51	10,758	was, fraud, winning, won, good, right, talking, up, the, turned	Debates about the turnaround, with some users using the discourse of electoral fraud.
16	168	10,030	zema, minas, nikolas, strange, general, something, mg, winning, wrong, vote	Discussion of the voting outcomes for the state of Minas Gerais, debating on how the senate and governor votes were for far right candidates, but the most voted for president in the region was Lula.
17	477	9,902	voted, simone, blank, null, asshole, voted, ciro, get, you, dick	Critics on null votes and about votes for the third and fourth place candidates of the presidential election.
18	53	9,704	states, leads, northeast, all, general, leading, minas, bahia, region, mato	Comments on the regions of Brazil that Lula was leading the dispute.
19	263	9,454	street, you, are, any, stay, what, someone, any, tweet, people	A topic with common used words in tweets in Portuguese, commenting the event.
20	41	9,367	first, woman, elected, federal, PT, paulo, new, something, support, sp	Discusses the first trans women elected for different Brazilian states as congresswomen.

of the discussions. For example, the elections to the Chamber of Deputies and Senators (topics 8, 10, 16 and 18) and Lula's victory (topics 1, 3, 11 and 13) were more present in certain rounds. However, some topics were discussed almost equally in both rounds (topics 2, 15, 19 and 20). Among these topics, it is worth highlighting Topic 15, which referred to possible electoral fraud, a topic that received a lot of attention, especially from

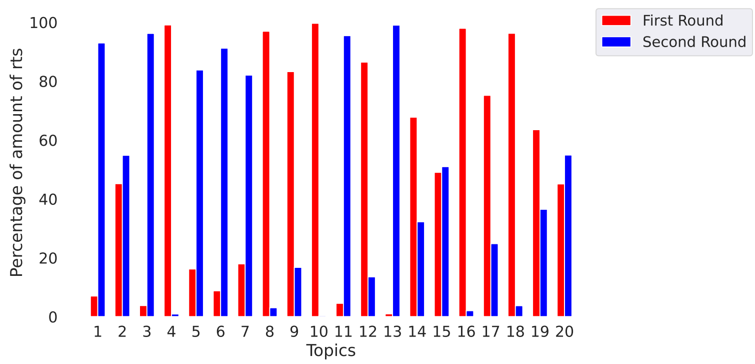


Figure 4.1: Percentage of normalized retweets per topic.

far-right voters due to the spread of fake news and misinformation about the matter.[4]

Let us now turn to the psycholinguistic analysis of the debate. Figure 4.2 shows the results. In the first round, we highlight the notable presence of words associated with *home, money, assent, see and sexual*. Retweets containing words associated with *home* often show people describing their voting experience (leaving their home to go vote)[5] or discussing family-related matters such as losses due to the COVID-19 pandemic. [6] Money related words are mainly related to economical concerns. In the second round, retweets frequently use words regarding *death, friend, religion* and *positive emotion*. Interestingly, religion (moral and religious concerns) was actually a theme highly emphasized by Bolsonaro's campaign.[7] Positive emotions were probably expressed by the Lula's supporters due to his victory. Death, instead, was closely related to the retweets regarding the COVID-19 pandemic and the way Bolsonaro's government deals with it.

Finally, we focus on the sentiment analysis. Table 4.3 shows the general sentiment distribution of the retweets. Negative sentiments predominate in both election rounds. However, the percentage of positive retweets increases 2.2-fold in the second round, which is confirmed by the increase in positive emotion-related words in the retweets (Figure 4.2).

We go further in our analysis by presenting the breakdown of sentiment by topic. Figure 4.3 summarizes the *contrastive score*, which is calculated as the difference between the proportion of positive and negative retweets. Overall, negative sentiment predominates in both rounds. Topics 3 and 14, which relate to Lula's victory, are exceptions to this trend. The analysis provides an important insight into the polarized nature of the political debate, especially on platforms such as Twitter [36, 48, 9].

[4] https://www.nytimes.com/interactive/2022/10/25/world/americas/brazil-bolsonaro-misinformation.html

[5] "They abused the public sector, lied, threatened believers and employees, attempted a coup, used the police to stop voters on their way to vote. It didn't help. "Good evening, President Lula! - popular resistance won." Read and enjoy @Maufalavigna #DomingoDetremuraSDV"

[6] "The route from home to the polling place passes through my work and the UBS where I took the 4th dose (the one in the photo). On the way, all I could think about was the 9 patients I lost. In my mother's desperation for me to get vaccinated... While Bolsonaro was riding a jet-ski. #Eleicao2022"

[7] https://edition.cnn.com/2022/10/29/americas/brazil-elections-gun-religion-intl-latam/index.html

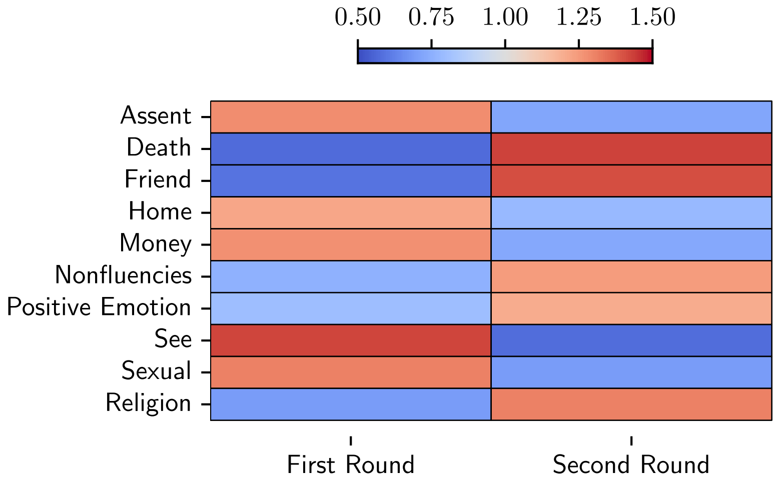


Figure 4.2: Top-10 LIWC attributes (complete networks).

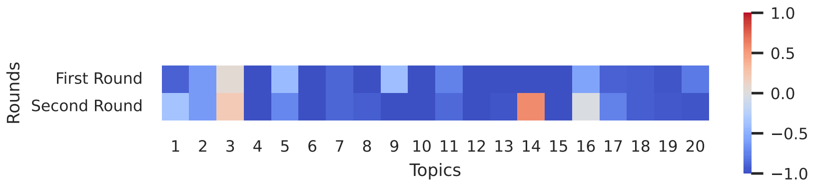


Figure 4.3: Contrasting sentiment score (complete networks).

Table 4.3: Sentiment distribution (complete networks).

Date	**Negative (%)**	**Positive (%)**	**Neutral (%)**
1st Round	243,098 (60.1%)	17,791 (4.4%)	143,333 (35.5%)
2nd Round	171,128 (53.3%)	31,863 (9.9%)	118,325 (36.8%)

4.2.2 Key Users' Dissemination

We now turn our attention to the content diffused by the *key users*, whose belong to the DF+NB backbones. Their interactions extend beyond random occurrences, being in the core of the discussion across the analyzed Twitter networks.

Table 4.4 lists the top 20 topics retweeted by these users. Twelve of these topics are the same as those shared by all users in the complete networks, even if they appear in a different order. When analyzing the complementary topics, by manually verifying the content of the most popular messages in each one, we find that these topics have more to do with Lula's performance in the election, as well as some topics usually raised by the opposition to Bolsonaro, such as concerns about previous government actions in the areas of education, health and the COVID-19 crisis. We also note that the majority of retweets in the top 20 for the backbones have to do with discussing or celebrating Lula's victory.

Our data reveals interesting changes in the psycholinguistic characteristics of the

Table 4.4: Top 20 discussion topics found on Twitter for backbones.

ID	# Tweets	# Retweets	Most Discriminative Words	Description (New Topics Only)
1	587	8,005	inácio, silva, luiz, president, biden, elected, victory, new, brazil, luis	
2	473	5,865	elections2022, turn, turned, lulaonFirstRound13, northeast, elections2022, lulapresident1, elections2022, turnaround, lulinha	
3	76	3,893	supporters, celebrating, turnaround, victory, party, celebrate, streets, brasília, against, petista	
4	79	3,684	difference, falls, million, less, fell, thousand, 46, only, elections2022, votes	Highlights the difference between Lula's and Bolsonaro's votes
5	53	3,112	urgent, missing, only, less, missing, thousand, victory, lulaonFirstRound13, elections2022, give	
6	97	2,987	history, times, time, re-election, since, 1st, president, re-elect, term, succeeds	
7	42	2,732	elected, federal, woman, first, paulo, historic, senate, all, damares, against	
8	56	2,418	must, minutes, prf, 19, 10, next, in this, night, globo, campaign	Discuss the projections by major news agencies, which estimate that Lula would surpass Bolsonaro in votes
9	41	2,179	first, woman, elected, federal, PT, paulo, new, something, support, sp	
10	177	2,126	2nd, datafolha, 1st, round, second, presidential, elections2022, governor, will, need	DataFolha survey indicating a high likelihood of second-round runoffs for the presidential race
11	53	1,976	states, leads, northeast, all, general, leading, minas, bahia, region, mato	
12	167	1,874	way, third, fault, simone, chance, second, have, ciro, you, voted	
13	65	1,838	deputy, federal, paulo, ferreira, voted, mg, nikolas, elected, paraná, senator	
14	45	1,776	advantage, over, continues, determined, 47, 90, million, ballots, leadership, almost	After the majority of voting machine results were cleared, Lula was leading the race, sparking widespread discussion among voters
15	83	1,667	northeast, arriving, always, north, pará, elections2022, bahia, region	Tweets celebrating the Northeast region votes were being counted, which significantly impacted the voting results in favor of Lula
16	69	1,652	health, want, education, people, good, freedom, life, live, govern, because	Concerns about education and health issues
17	40	1,565	lo, lulapresident2022, let's go, turn, big, lulapresident1, victory, moment, luiz, listen	
18	64	1,526	turn, delicious, lulinha, lulapresident, lulaonFirstRound13, calm, elections2022, god, turned, northeast	Tweets with the use of "He who laughs last, laughs best" to comment on Lula's victory in the election results.
19	293	1,454	over, nightmare, goodbye, won, bye, lulapresident2022, end, well, above, finally	
20	37	1,443	amazonas, pandemic, during, vote, for him, shame, seems, many, leading, leads	Controversial outcomes arose from the presidential election in the state of Amazonas, due to Bolsonaro's actions during the COVID-19 crisis[8]

content shared by users in the extracted backbones. Figure 4.4 shows these results. Words related to *family* predominate in the content shared by these users, mainly in the first round. To better understand what users shared on this topic, we analyzed our data manually. These messages mainly mentioned Bolsonaro's family, which is heavily involved in politics and was at the center of several controversial situations reported by the

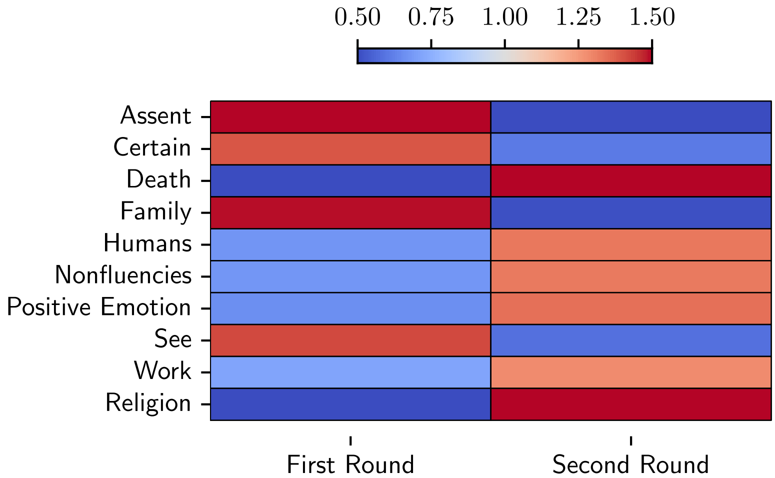


Figure 4.4: Top-10 LIWC attributes (key users).

Table 4.5: Sentiment distribution (key users).

Date	**Negative (%)**	**Positive (%)**	**Neutral (%)**
1st Round	29,815 (44.16%)	2,238 (3.31%)	35,468 (52.53%)
2nd Round	15,545 (46.21%)	4,399 (13.08%)	13,693 (40.71%)

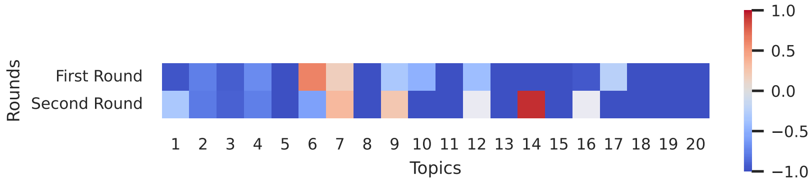


Figure 4.5: Contrasting sentiment score (key users).

media, as well as family conflicts due to the political polarization that was a remarkable characteristic of Brazil's 2022 elections. Regarding the second election round, retweets with *religion* related words attracted more attention of these users.

Finally, the results of the sentiment analysis are shown in Figure 4.5 and Table 4.5. Key users tend to balance the shared content in both rounds with negative and neutral sentiments. In contrast to users from the complete networks, key users shared more tweets with neutral sentiments. In addition, retweets shared in the second round tended to be more positive than those shared by total users. In particular, we highlight Topic 14, which refers to Lula's leadership in the second round. These retweets are very positive, suggesting that these users strongly support the possibility of Lula's victory.

Takeaway. The topics shared by users in the backbones (core) are different from those shared across the network (including with peripheral users). Interestingly, the most shared topics in the backbones, which are not found in the overall networks, are more focused on supporting Lula and celebrating his victory. The sentiment towards the content is more positive in the second round, especially in relation to the topic discussing Lula's victory (topic 14). The percentage of negative retweets is lower in the two rounds.

4.2.3 Persistent Users' Dissemination

In this section, we analyze the content disseminated by users classified as *persistent users*, who belong to the backbone networks of both election rounds.

In our data, a total of 624 users, representing 7% of users on the DF+NB backbones, shared information across the two rounds. This small percentage of users demonstrated significant activity, retweeting almost 22% of the messages shared across the events of interest for the backbones. Figure 4.6 shows the cumulative distribution of followers when we consider the persistent group and the users in the first and second round networks. All persistent users are unverified users [9] and they tend to have more followers than the users in the complete network.

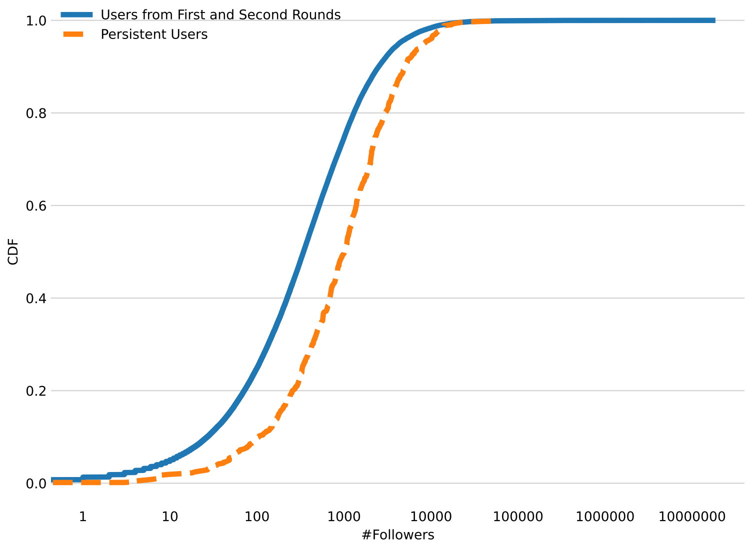


Figure 4.6: Followers Distribution - All users versus Persistent users.

In our final discussion on characterization, we focus on the main topics shared by persistent users on Twitter networks and the textual aspects of their dissemination. The most shared topics by persistent users are almost the same as those of the main users, with one exception, as can be seen in Table 4.6. This divergent topic is about Lula's lead in the second round, with a significant margin of votes in the state of Bahia. Some of them appeared in different positions because the total number of retweets in the network takes into account the only persistent group, but overall it is a very similar scenario to the main users' spread.

We now focus on the textual aspects of the content disseminated by this set of users. First, we present the LIWC analysis in Figure 4.8. Most of the content attributes align closely with those disseminated by the key users. Notably, the first round of content

[9] our data was collected before the introduction of the blue tick purchase plans

Table 4.6: Top 20 discussion topics found on Twitter for persistent users.

ID	# Tweets	# Retweets	Most Discriminative Words	Description (New Topics Only)
1	587	1,828	inácio, silva, luiz, president, biden, elected, victory, new, brazil, luis	
2	473	1,322	elections2022, turn, turned, lulaonFirstRound13, northeast, elections2022, lulapresident1, elections2022, turnaround, lulinha	
3	76	886	supporters, celebrating, turnaround, victory, party, celebrate, streets, brasília, against, petista	
4	79	786	difference, falls, million, less, fell, thousand, 46, only, elections2022, votes	
5	97	615	history, times, time, re-election, since, 1st, president, re-elect, term, succeeds	
6	53	582	urgent, missing, only, less, missing, thousand, victory, lulaonFirstRound13, elections2022, give	
7	42	576	elected, federal, woman, first, paulo, historic, senate, all, damares, against	
8	56	557	must, minutes, prf, 19, 10, next, in this, night, globo, campaign	
9	177	541	2nd, datafolha, 1st, round, second, presidential, elections2022, governor, will, need	
10	41	470	first, woman, elected, federal, PT, paulo, new, something, support, sp	
11	45	439	advantage, over, continues, determined, 47, 90, million, ballots, leadership, almost	
12	69	424	health, want, education, people, good, freedom, life, live, govern, because	
13	53	399	states, leads, northeast, all, general, leading, minas, bahia, region, mato	
14	65	398	deputy, federal, paulo, ferreira, voted, mg, nikolas, elected, paraná, senator	
15	49	347	bahia, 30, million, counted, still, people, day, voting, missing, voted	Talks about Lula having the majority of votes in the state of Bahia, a strong electoral college in the Northeast region, during the counting
16	40	343	lo, lulapresident2022, let's go, turn, big, lulapresident1, victory, moment, luiz, listen	
17	83	314	northeast, arriving, always, north, pará, elections2022, bahia, region	
18	37	307	amazonas, pandemic, during, vote, for him, shame, seems, many, leading, leads	
19	64	300	turn, delicious, lulinha, lulapresident, lulaonFirstRound13, calm, elections2022, god, turned, northeast	
20	167	251	way, third, fault, simone, chance, second, have, ciro, you, voted	

highlights a significant presence of family-related topics, while the second round shifts focus towards themes like death, religion, work, and positive emotion.

Regarding sentiment analysis, Figure 4.7 underscores that the group predominantly disseminated negative sentiment among the top 20 topics. Table 4.7 presents the sentiment distribution for the persistent users. In the first round, neutral messages were more frequent, while in the second round, negative messages became more widespread.

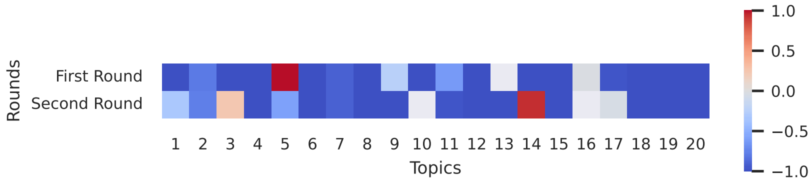


Figure 4.7: Contrasting sentiment score (persistent users).

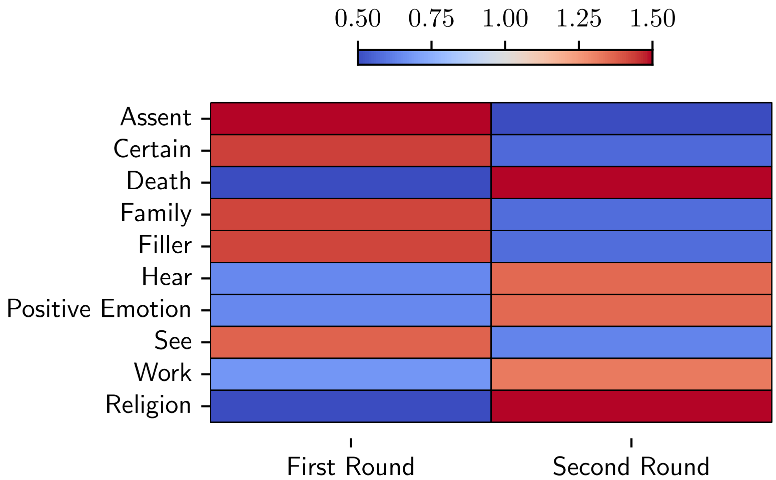


Figure 4.8: Top-10 LIWC attributes (persistent users).

Table 4.7: Sentiment distribution (persistent users).

Date	Negative (%)	Positive (%)	Neutral (%)
1st Round	5,704 (42.85%)	483 (3.63%)	7,123 (53.52%)
2nd Round	3,606 (45.36%)	1,165 (14.66%)	3,178 (39.98%)

Takeaway. Although the group of persistent users is small (7%), our results suggest that they play an important role in the dissemination of content by the main users, accounting for 22% of retweets. The topics they share are almost identical to those shared by the main users and their potential to reach a larger audience is higher than that of users in the complete networks, as they tend to have more followers.

4.3 Virality Classification

Natural Language Processing (NLP) techniques are widely used to analyze diverse contexts, particularly in tasks involving behavioral predictions based on historical data. Text classification models play a crucial role in this process by systematically categorizing information and detecting patterns [30, 105]. In this section, we aim to answer the research question: ***RQ3: Does incorporating information about user engagement,***

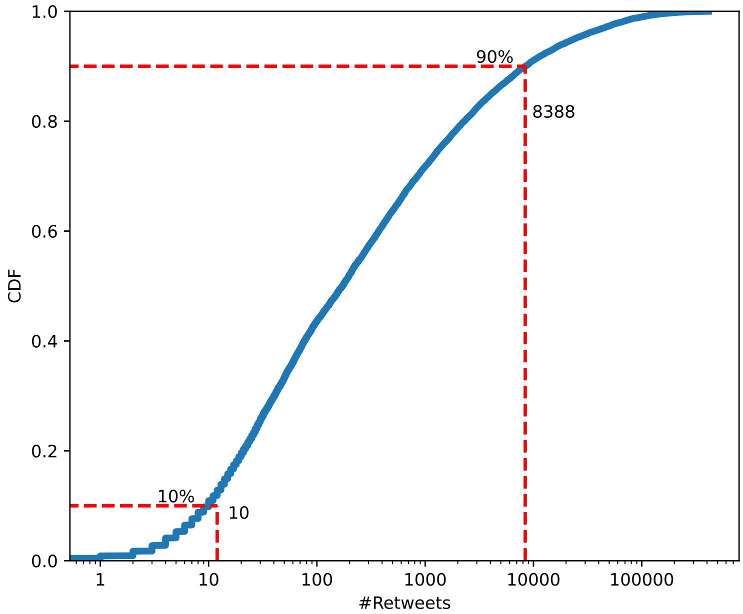


Figure 4.9: Retweet CDF with cutoff points for classifier analysis.

captured through the presence of messages in the backbone of the dissemination network, alongside textual features, improve the classification of viral and non-viral content during the 2022 Brazilian election?. To address this, we explored two classification approaches: one relying solely on textual features and another incorporating topological information from the dissemination network. Our analysis was grounded in the previously constructed dissemination network, where we examined the original messages that were shared and propagated. The goal was to classify messages based on their virality, distinguishing widely disseminated content from ordinary messages with limited reach. Thus, the first step in this process was defining the threshold that separates these two categories.

Figure 4.9 presents the Cumulative Distribution Function (CDF) of the retweet count for the selected messages. To establish a clear distinction between viral and non-viral content, we considered the two extremes of the distribution, selecting the first 10% and the last 10% of the CDF. This approach allowed us to focus on messages with either very low or very high dissemination rates, ensuring a meaningful comparison between both groups. As a result, the classification threshold was set at 10 retweets for non-viral content and 8,388 retweets for viral content. Following this approach, the final dataset for the classification task comprised 2,933 messages, of which 1,402 were labeled as viral and 1,531 as non-viral. This dataset was then used in the training and evaluation of our classification models, allowing us to assess the impact of incorporating both textual and topological features in predicting content virality.

4.3.1 Linguistic Characteristics of Viral Classes

Before proceeding with the classification models, we first examined the linguistic characteristics of viral and non-viral messages. This preliminary step aimed to identify textual differences between the two classes that could influence content dissemination. To achieve this, we analyzed both syntactic and semantic aspects of the messages using two complementary approaches: confidence interval estimation and statistical hypothesis testing.

First, we computed 95% confidence intervals for key linguistic features to observe the range of values that likely characterize each class. This provided an estimate of the expected variation in linguistic properties within viral and non-viral messages, allowing for a descriptive comparison between the two groups. Second, to formally test whether these linguistic distributions differed significantly, we applied the Kolmogorov-Smirnov (KS) test. The KS test measures the maximum difference between the empirical cumulative distribution functions of two samples, assessing whether they are drawn from the same underlying distribution. A statistically significant result ($p < 0.05$) indicates that the two distributions are unlikely to be identical, meaning that the linguistic characteristics of viral and non-viral messages exhibit meaningful differences. By combining these two approaches, we aimed to both describe and statistically validate differences in textual features that could play a role in content virality.

Figure 4.10: POS TAG diversity Confidence Intervals.

To analyze the lexical and grammatical diversity in viral and non-viral messages, we examined the diversity of POS (Part-of-Speech) tagging, as described in Chapter 3. This measure quantifies the variety of syntactic structures present in the messages, providing insights into whether viral content exhibits more or less linguistic complexity compared to non-viral content.

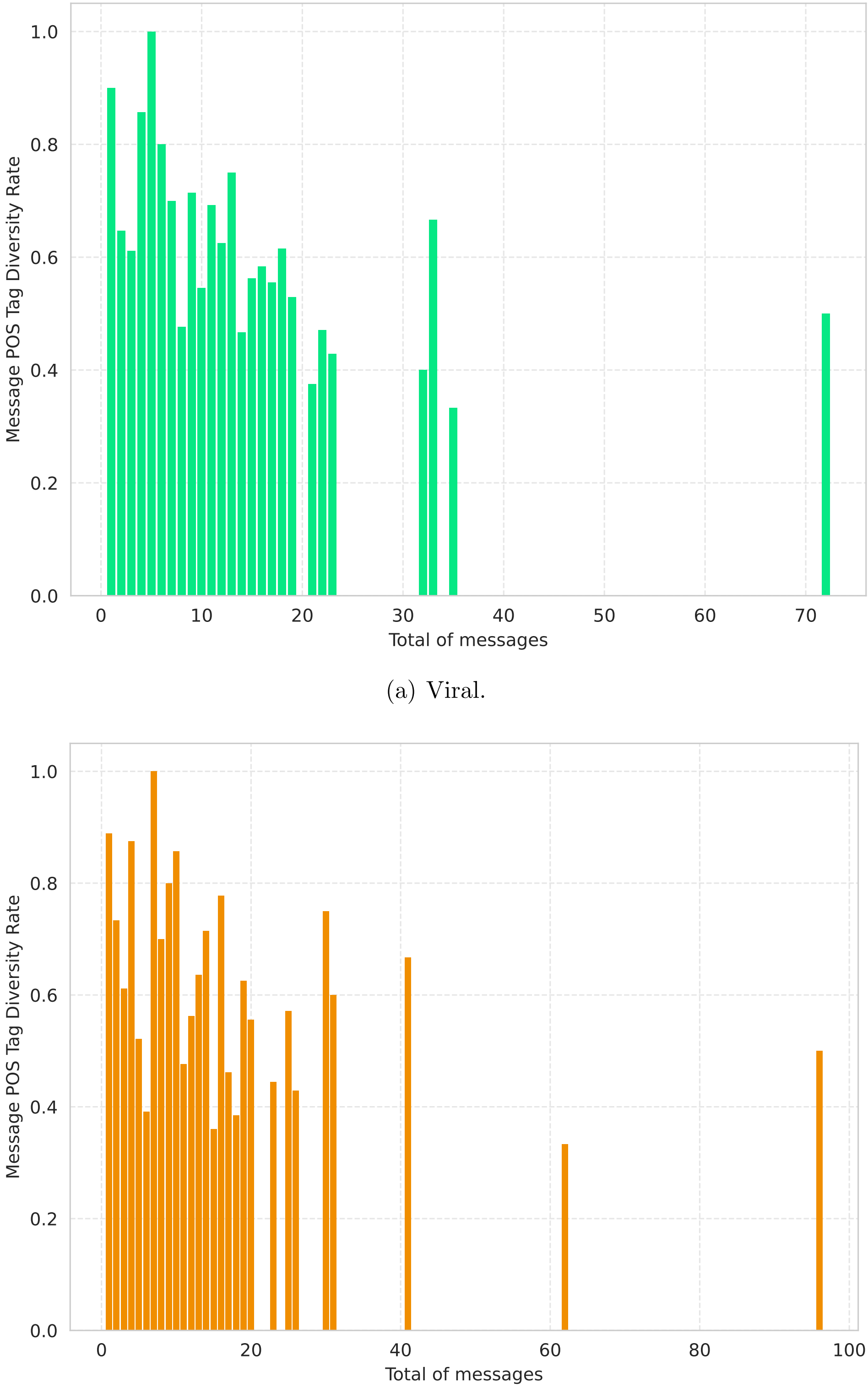
Figure 4.10 presents the confidence intervals for POS tag diversity in both classes. The viral messages had a confidence interval of $(0.40, 0.42)$, whereas the non-viral messages had a slightly higher range of $(0.42, 0.44)$. The difference between these distributions was found to be statistically significant, indicating that non-viral messages tend to exhibit slightly greater syntactic diversity than viral messages. To further illustrate this finding, Figure 4.11 shows the distribution of POS tag diversity across the dataset. We observe that, for both viral and non-viral messages, the majority of texts cluster around a POS tagging diversity score of approximately 0.50. However, a small subset of messages exhibits significantly higher diversity, suggesting that while some tweets employ more varied grammatical structures, the general trend in both classes leans toward linguistic uniformity.

This result aligns with common patterns of social media communication, where users tend to adopt concise and familiar language rather than highly diverse syntactic structures. The slightly higher POS tag diversity in non-viral messages could indicate that more complex or varied sentence structures might not necessarily favor engagement, reinforcing the idea that simpler, more standardized language is more likely to be widely shared.

The complexity of sentences in the classified content was also examined by calculating the Type-Token Ratio (TTR) for both viral and non-viral messages. This measure quantifies lexical diversity by assessing the proportion of unique words (types) relative to the total number of words (tokens) in a text. The goal was to determine whether one class of messages exhibited a richer vocabulary or tended to rely on a more limited set of words, irrespective of text length.

The results indicated a confidence interval of (0.89, 0.90) for viral messages and (0.87, 0.89) for non-viral messages, with a statistically significant difference between the two, as shown in Figure 4.12. While both classes contained messages with high lexical diversity, which is expected in a politically charged debate with multiple perspectives in a large country like Brazil, viral messages exhibited slightly greater linguistic complexity. This suggests that widely shared content might incorporate a broader range of expressions or combine different terms in a more engaging and creative manner, potentially contributing to audience retention and interaction.

Table 4.8 presents examples of messages from each class along with their corresponding TTR values. Messages with lower TTR tend to rely on repetitive patterns or characters, resulting in simpler texts that convey less information to readers. In contrast, messages with higher TTR employ a more formal language, exhibiting a more diverse lexical structure. Notably, the non-viral message with high TTR features a more aggres-



(a) Viral.

(b) Non-Viral.

Figure 4.11: POS TAG diversity distribution.

sive tone, whereas its viral counterpart reports on an incident classified as a crime in Brazil—voters taking pictures of their choices on the electronic voting machine [10].

[10]https://www.tre-sp.jus.br/comunicacao/noticias/2022/Outubro/eleitores-que-tiraram-fotos-ou-

Table 4.8: Example of messages with their corresponding TTR score.

Message Class	TTR Score	Text
Non-viral	1.00	“It scares me to see the strength of this demon Bolsonaro in 2022, how many stupid people deny having access to information”
Non-viral	0.07	“BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL BOLSONARO WILL FALL”
Viral	1.00	”Voter takes cell phone to the electronic voting machine and posts on social media, which is a crime. #Elections2022”
Viral	0.34	”Luiz Inácio Lula da Silva you will sign my diploma!!!!!!!!!!!!!!!!!!!!!!!”

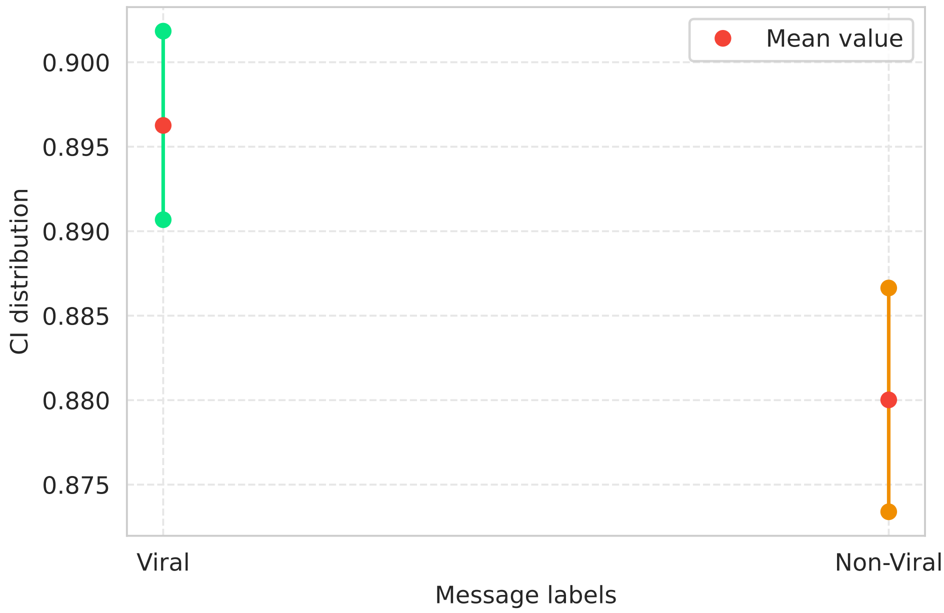


Figure 4.12: Complexity of messages Confidence Intervals.

The final analysis examined message length, as shown in Figure 4.13. Social media users often prioritize instant, concise content that delivers small, digestible pieces of information [5]. At the time of data collection, Twitter enforced a 280-character limit per post, though total text length could slightly exceed this when considering mentions, links, and media attachments. We sought to investigate whether viral messages followed the commonly assumed trend of being shorter or if, in this specific context, lengthier content was more likely to gain traction.

The results indicated that viral messages were significantly longer on average compared to non-viral ones, suggesting that, at least within the political discourse analyzed,

videos-do-voto-podem-ser-processados

succinct messages were less effective in driving engagement. Given the political nature of the dataset, users appeared to favor more contextual or informative content, which aligns with the tendency to share detailed updates, such as election results or political analyses. The confidence intervals for message length were (125.66, 133.18) for viral tweets and (110.35, 117.06) for non-viral tweets, with a statistically significant difference confirmed by the KS test (Figure 4.14).

Additionally, we conducted a word cloud analysis to uncover the most frequently used words in viral and non-viral messages. To highlight the key differences between the two classes, we removed common words appearing in both sets, ensuring that only distinctive terms remained in each category. Figure 4.15 displays the 300 most frequently used words in both viral and non-viral messages. In the viral class, electoral-related terms were prevalent, including words such as voting booth, congresswoman, poll worker, and mayor. Additionally, the names of several countries—Japan, Korea, Australia, and New Zealand—appeared frequently, likely because Brazilian citizens in these countries voted earlier in the election process.

Political figures also had a strong presence in the viral category, particularly Alexandre de Moraes and Carla Zambelli, both of whom were widely discussed in the media due to their involvement in controversial political actions at the time[11,12].

Beyond electoral terms and political figures, highly polarized and ideologically charged words such as dictatorship and fascism appeared prominently. These words were commonly used by opponents of Brazil's far-right movement to critique the government's policies[13]. Another noteworthy term was Marcola, the alias of Marcos Willians Herbas Camacho, leader of the PCC criminal faction. During the election, fake news spread alleging that he endorsed candidate Luiz Inácio Lula da Silva (PT), prompting the Superior Electoral Court (TSE) to intervene and demand the removal of such misinformation[14]. Finally, the word police surfaced frequently in the viral dataset due to controversies surrounding police roadblocks during the second round of voting. Reports suggested that the Federal Highway Police (PRF) disproportionately conducted operations in the Northeast, where Lula was leading in the polls, raising concerns about potential voter suppression[15].

By contrast, the non-viral class exhibited a distinct linguistic pattern. A key difference was the prevalence of offensive language and political slang. Several terms in this set were insulting or aggressive, including politically charged slurs. The initials JB (Jair Bol-

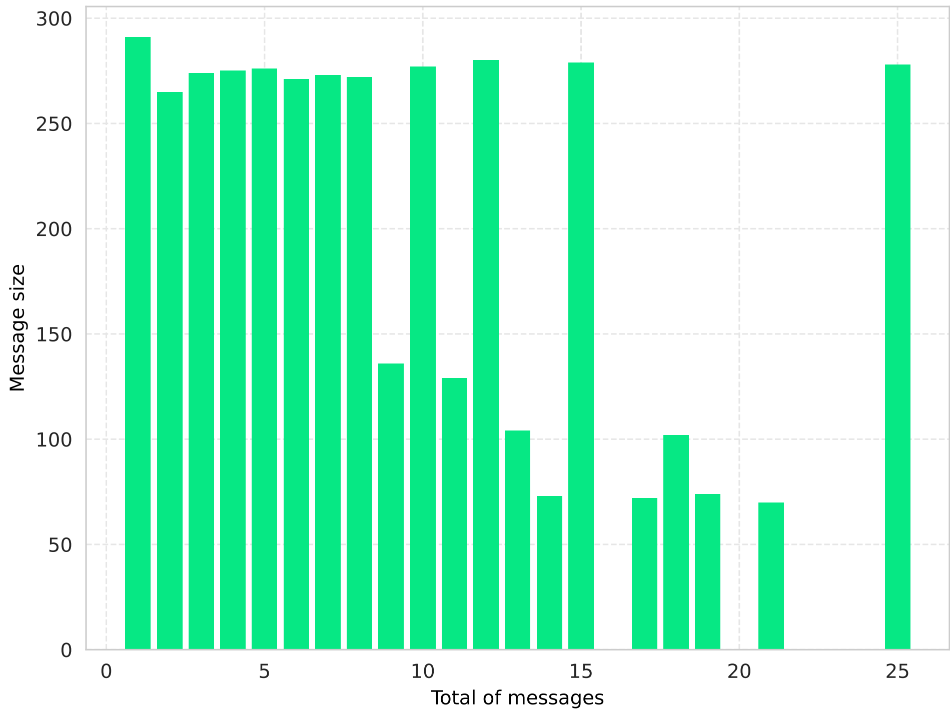
[11] https://g1.globo.com/sp/sao-paulo/noticia/2025/01/30/carla-zambelli-tem-mandato-de-deputada-cassado-pelo-tre-sp-e-fica-inelegivel-por-8-anos-apos-divulgar-fake-news-sobre-processo-eleitoral.ghtml

[12] https://oglobo.globo.com/politica/eleicoes-2022/noticia/2022/10/na-vespera-da-eleicao-alexandre-de-moraes-diz-que-democracia-e-construcao-coletiva-dos-que-acreditam-na-paz.ghtml

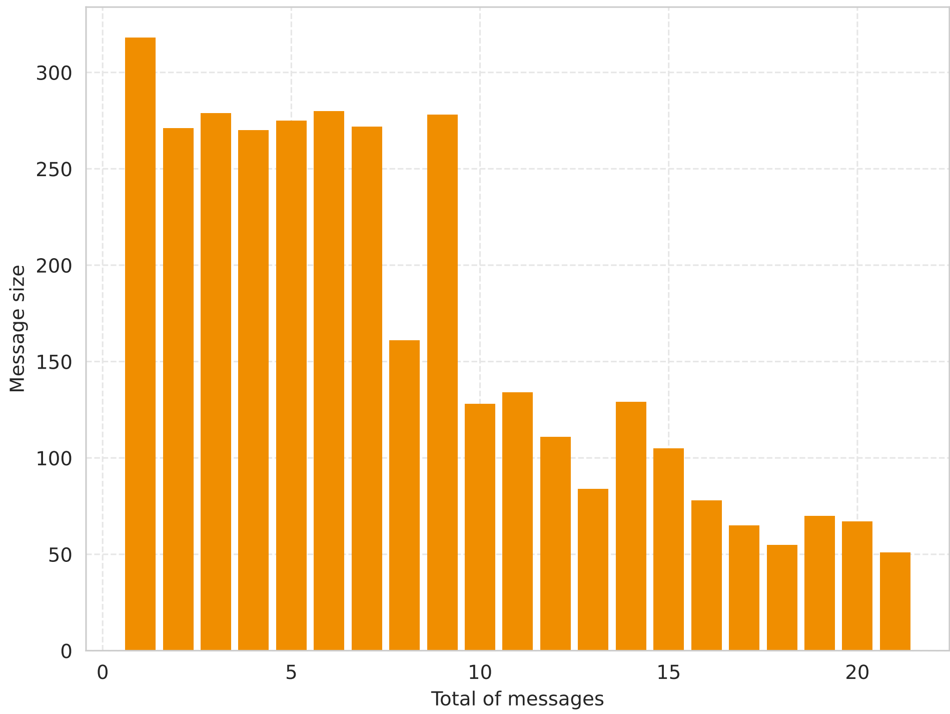
[13] https://oglobo.globo.com/blogs/portugal-giro/post/2022/08/bolsonaro-repete-lema-da-ditadura-de-salazar-diante-do-coracao-de-d-pedro-e-surpreende-portugueses.ghtml

[14] https://www.cnnbrasil.com.br/politica/moraes-manda-bolsonaro-e-sites-removerem-postagem-sobre-suposto-apoio-de-marcola-a-lula/

[15] https://www.cnnbrasil.com.br/blogs/debora-bergamasco/eleicoes/investigacao-da-pf-conclui-que-blitze-da-prf-impactaram-eleicoes-em-2022/



(a) Viral.



(b) Non-Viral.

Figure 4.13: Message size distribution.

sonaro) and cirista (a term for supporters of Ciro Gomes) appeared frequently, reflecting factional discourse. Mentions of political parties, such as PP, MDB, and Republicanos, also stood out, indicating a focus on party affiliation rather than specific election events.

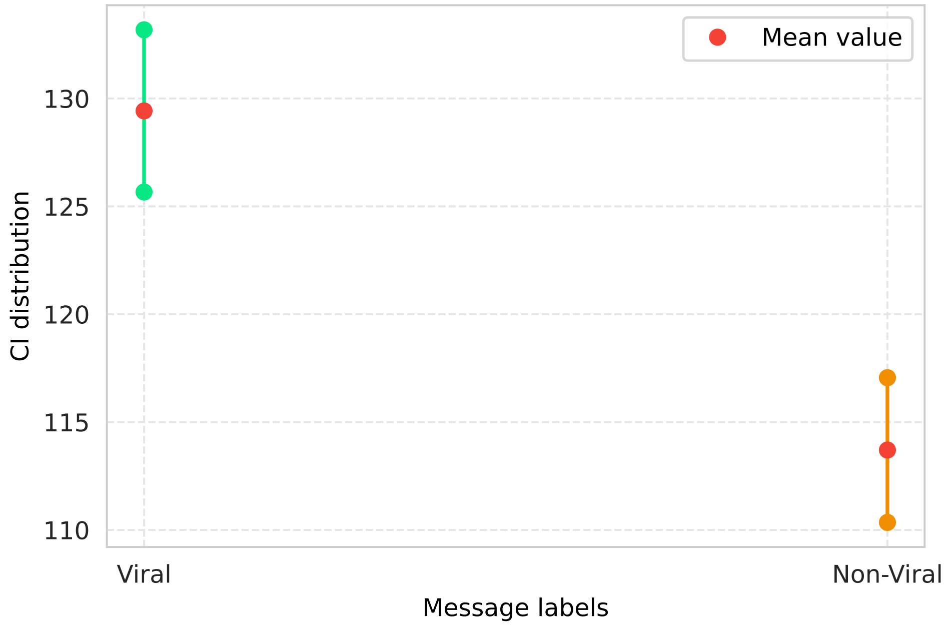
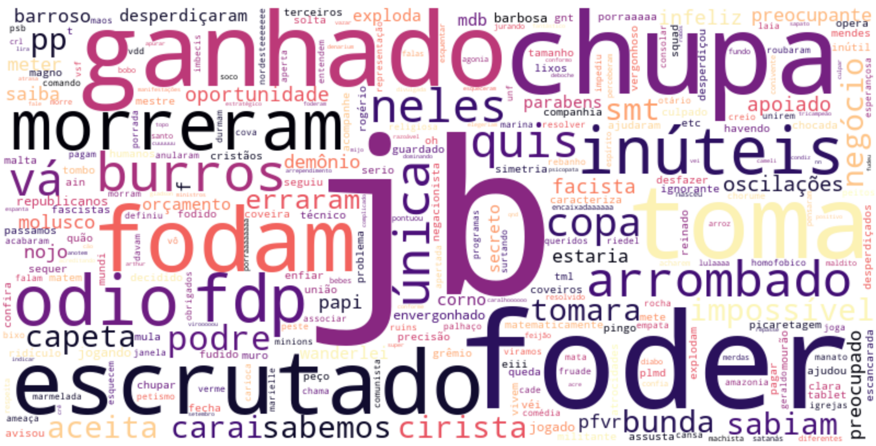


Figure 4.14: Message size Confidence Intervals.

Another striking feature of the non-viral word cloud was the presence of derogatory terms, particularly the word mollusc, an offensive nickname used by opponents of Lula. The widespread use of insult-based discourse in this category suggests that messages with more aggressive or offensive tones did not achieve the same level of virality as those focused on political events, election results, or misinformation narratives.

Takeaway. Understanding the linguistic characteristics of viral and non-viral content provides valuable insights into what drives engagement in OSMPs. Our analysis revealed that viral messages tended to be longer, featured a more complex vocabulary, and often referenced key events surrounding the election. In contrast, non-viral messages exhibited slightly greater diversity in POS tagging but were also more likely to contain aggressive or hostile language. These findings suggest that while structured and informative content was more likely to gain traction, messages with a more confrontational tone did not spread as widely. One possible reason for this could be who the author of the messages was, as verified accounts of important political figures or media outlets usually write more formal or polite messages using less aggressive terms. These accounts often have a high reach and engagement, which could contribute to their messages often going viral.

(a) Viral.



(b) Non-Viral.

Figure 4.15: Word clouds for the predictions of the Bertimbau+B model.

4.3.2 Impact of Textual and Topological Features on Virality Classification

We started our evaluation with the Baseline BERTimbau classifier, which used only the textual content of the messages as input. As described in Chapter 3, we conducted a grid search to optimize hyperparameter configurations, selecting the best-performing model based on validation loss. The lowest validation loss achieved was 0.1223. With this configuration, we retrained the model and evaluated its performance on the test set. To ensure the robustness of the evaluation, we applied bootstrapping with 300 iterations, computing 95

Although textual content serves as the primary means of communication in OSMPs like Twitter, engagement is influenced by hidden factors beyond the text itself. To examine whether structural information from the dissemination network could enhance classification, we introduced the Bertimbau+B model. This model extends the baseline classifier by incorporating a topological signal: whether a message appeared in the backbone of the dissemination network. Since backbone users are highly engaged and play a central role in content amplification, their sharing patterns could indicate a structured dissemination dynamic.

Following the same hyperparameter tuning process, the best Bertimbau+B configuration was selected based on validation loss, which resulted in a higher final value of 0.4493. This increase suggests that the model may have introduced greater complexity, possibly due to the integration of topological data. While a higher loss might indicate a more challenging optimization process, the true impact of the additional feature is better understood through classification metrics such as precision, recall, and F1-score. To ensure a fair comparison, we applied the same bootstrapping methodology used for the baseline model.

Metric	**BERTimbau (Baseline)**	**BERTimbau+B**
Accuracy	0.79 (0.74, 0.83)	0.73 (0.68, 0.79)
Precision	0.75 (0.68, 0.82)	0.66 (0.60, 0.73)
Recall	0.82 (0.76, 0.88)	0.90 (0.85, 0.95)
F1-Score	0.79 (0.73, 0.84)	0.76 (0.71, 0.81)
AUC-ROC	0.87 (0.82, 0.91)	0.84 (0.79, 0.88)

Table 4.9: Bootstrap results for the models (Mean and 95% Confidence Interval).

Table 4.9 presents the performance metrics for both models. To evaluate whether the observed differences were statistically significant, we first applied the Kolmogorov-Smirnov (KS) test to assess whether the paired differences in Accuracy, Precision, Recall, and F1-score followed a normal distribution. The results indicated a significant deviation from normality ($p < 0.05$), leading us to reject the null hypothesis that the differences were normally distributed. Given this, we proceeded with the Wilcoxon Signed-Rank Test, a non-parametric alternative to the paired t-test, to evaluate whether the median differences in these metrics were significantly different from zero. The results showed p-values below 0.05 for all evaluation metrics, confirming that the performance differences between the models were statistically significant and unlikely to be due to random chance.

The Baseline BERTimbau achieved an F1-score ranging from 0.73 to 0.83 within the confidence interval, with an average close to 0.80. Accuracy followed a similar trend, showing that the model classified most instances correctly. Notably, recall was the highest-performing metric (>0.82), demonstrating that the baseline model was particularly effective in identifying viral tweets while minimizing false negatives. The Bertimbau+B model, which integrates topological information, showed a higher recall than the baseline model (0.90 vs. 0.82). This suggests that incorporating information from the backbone dissemination network helped capture patterns indicative of viral content. However, this gain in recall came at the cost of lower precision (0.66 vs. 0.75), indicating that the Bertimbau+B model was more likely to misclassify non-viral tweets as viral. This trade-off suggests that, while the model became more sensitive to viral content, it also introduced a higher rate of false positives.

To further assess the models' ability to distinguish between viral and non-viral content, we analyzed the AUC-ROC, a key metric that evaluates classification performance across different decision thresholds. Figure 4.16 presents the ROC curves for both models, calculated from the original test set. The results indicate that both classifiers achieved high AUC values, demonstrating strong separability between the two classes. The Baseline BERTimbau model achieved an AUC-ROC of 0.868, while BERTimbau+B obtained a slightly lower value of 0.841. This suggests that, in terms of overall classification performance, the baseline model had a marginal advantage in distinguishing viral from non-viral tweets. This trend aligns with the bootstrap results (Table 4.9), which reflect the mean AUC-ROC values obtained over 300 iterations during the test phase.

To determine whether this difference was statistically significant, we first applied the Kolmogorov-Smirnov (KS) test, which confirmed that the paired differences between AUC values deviated from a normal distribution. Consequently, we employed the Wilcoxon Signed-Rank Test, which revealed that the performance gap between the two models was statistically significant ($p < 0.05$). This confirms that, although both models demonstrated strong classification ability, the Baseline BERTimbau model consistently performed slightly better in terms of overall discriminative power.

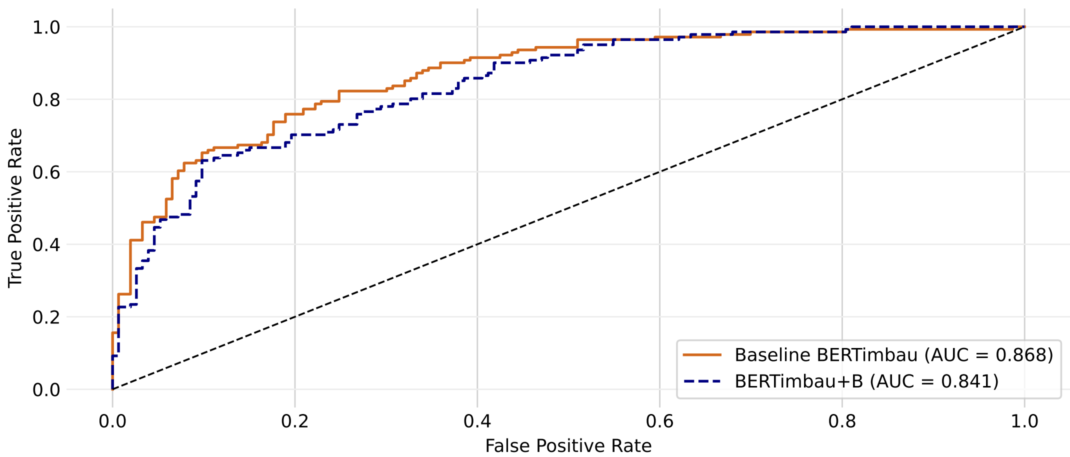


Figure 4.16: AUC-ROC curves for original test set of the two models.

This difference aligns with the previously observed metric evaluations for BERTimbau+B, which exhibited a higher recall at the expense of a slightly lower AUC-ROC. This result suggests a potential trade-off between overall classification performance and a more conservative approach to identifying viral content. Despite these variations, both models demonstrated similar discriminative capabilities, indicating that while integrating network topology enhances recall, it slightly decreases other metrics without substantially altering the global classification capacity.

It is important to emphasize that identifying and predicting viral content in social networks is an inherently complex task, highly dependent on the surrounding context. These findings highlight the potential benefits of incorporating information about a tweet's presence in the backbone of the dissemination network, which contributed to capturing more true positives. However, this improvement in recall came at the expense of other evaluation metrics. One possible explanation is that the backbone does not necessarily indicate global virality, but rather persistent sharing patterns within specific user groups. This characteristic could have influenced the classification process, making the model more biased toward predicting positive viral messages. Future research could explore strategies to mitigate this effect, such as integrating additional topological signals—like structural roles of users within the network or temporal patterns of message amplification—to enhance overall model performance.

Takeaway. The results suggest that incorporating topological information, specifically a tweet's presence in the backbone of the dissemination network, may enhance classification recall. However, the Baseline BERTimbau model achieved superior performance in terms of precision, F1-score, and accuracy, demonstrating its effectiveness in capturing both viral and non-viral instances more reliably. The AUC-ROC analysis confirmed that both models exhibited strong discriminative power. The primary distinction between the two models may stem from limitations in data collection, as we only captured a subset of the broader election-related debate. Consequently, the topological factor may not have fully surfaced the importance of dissemination content in this particular context.

Chapter 5

Conclusion

The role of online social media platforms (OSMPs) in shaping public discourse and influencing societal dynamics has grown significantly in the past decade. These platforms serve as key environments where individuals express opinions, engage in political debates, and contribute to the rapid dissemination of information. Understanding the mechanisms that drive information spread, particularly in highly polarized contexts such as elections, is a complex task that requires a multifaceted approach. In this research, we investigated the dissemination dynamics of the 2022 Brazilian elections on Twitter, focusing on both the textual and structural aspects of content diffusion. Our study was structured around three key perspectives: (i) the characterization of information diffusion networks, (ii) the analysis of textual properties and psycholinguistic features of disseminated messages, and (iii) the evaluation of classification models for distinguishing viral and non-viral content. By modeling dissemination networks and applying backbone extraction techniques, we identified groups of users who played a central role in spreading political content. This approach allowed us to highlight interactions that were more persistent and non-random, which can be indicative of coordinated dissemination patterns.

Our findings indicate that, throughout both electoral rounds, discussions on Twitter revolved around key political events, controversies related to the previous government, and reactions to election results. Users who consistently participated in content sharing were largely aligned with pro-Lula narratives, with the most prominent topics including government policies, the COVID-19 pandemic, and electoral integrity. We observed that while general network discussions were dominated by polarized and emotionally charged content, the backbone users exhibited a more balanced discourse, with a notable proportion of neutral sentiment. From a psycholinguistic perspective, the prevalence of attributes related to *home* and *family* in the first round, followed by *death, positive emotion, and religion* in the second round, reflects key aspects of the electoral debate. These shifts align with real-world political narratives, where religious discourse and references to the COVID-19 pandemic played a role in shaping voter opinions. The backbone extraction technique proved to be effective in uncovering stable patterns of interaction, shedding light on how specific user groups contributed to sustaining discussions over time.

In the context of content virality, we first performed a linguistic analysis to explore potential drivers of engagement. The results indicated that non-viral messages exhibited greater lexical diversity and a higher occurrence of aggressive language, whereas viral messages tended to be longer, contained more references to recent political events, and used more complex vocabulary. To further investigate the phenomenon of virality, we developed classification models based on fine-tuned transformers, employing grid search for hyperparameter optimization and AUC-ROC for performance validation. The baseline model, relying solely on textual features, achieved a strong F1-score (above 0.75 on average) and balanced performance across accuracy, precision, and recall.

We then extended the classification model by integrating topological information, specifically whether a tweet was shared within the backbone network. This addition introduced a layer of complexity to the classification task, reflecting the dynamic nature of social media environments where multiple factors influence content spread. The inclusion of structural features improved recall, indicating that the model became more sensitive to detecting viral content. However, this improvement came at a cost: precision and overall classification balance were negatively impacted compared to the baseline model. This trade-off highlights the intricate interplay between textual and network-based signals in predicting content virality, emphasizing that no single feature alone is sufficient for optimal classification. Despite the valuable insights gained, our study has some limitations. The dataset was collected based on a predefined set of keywords, which, while carefully selected, may not have captured the full spectrum of discussions related to the elections. Additionally, the construction of dissemination networks required substantial computational resources, limiting our ability to analyze larger time frames or alternative network structures. Future research could address these challenges by employing broader data collection strategies and exploring more refined network representations, such as incorporating user influence metrics or temporal patterns of message amplification.

Overall, this study contributes to the understanding of online political discourse and the dynamics of information spread by combining network analysis and NLP techniques. Our results demonstrate that while textual and structural features provide valuable signals for distinguishing viral content, their effectiveness varies depending on the specific classification objectives. The findings open avenues for further research into more sophisticated models that integrate multiple dimensions of content dissemination, potentially improving both predictive accuracy and interpretability in social media analysis.

5.1 Future Work

This master's thesis addressed the dissemination dynamics of political content on social media, particularly in the context of the 2022 Brazilian elections. However, several aspects remain open for further exploration. Future research could expand on our findings by addressing the following directions:

- **Exploring alternative backbone extraction methods and network representations**: While we employed a backbone extraction technique to filter out non-recurrent interactions and identify key users, alternative approaches could yield different perspectives on content dissemination. Future work could investigate different backbone models from the literature.
- **Incorporating additional topological features for virality prediction**: Our study focused on a single topological feature—whether a message appeared in the backbone network. Future research could extend this by including other network-derived features, such as the degree of the originating user, clustering coefficients, or temporal spread patterns. This could provide a richer understanding of how network structures influence content virality, improving classification performance.
- **Leveraging metadata and user attributes to enhance virality classification**: Beyond textual and structural features, metadata such as user engagement metrics (e.g., follower count, retweet rate, verification status), temporal posting patterns, and external link usage could serve as additional signals for predicting virality. Integrating these features in machine learning models may reveal how different user behaviors contribute to information spread.
- **Investigating the dissemination phenomenon in ephemeral OSMP events**: Many discussions on social media revolve around short-lived, high-intensity events, such as breaking news or viral hashtags. The dynamics of virality in ephemeral discussions may differ from those observed in prolonged election debates. Future studies could explore how network structures and sharing behaviors evolve in short-term viral events compared to sustained political discussions.
- **Applying causal inference to understand dissemination mechanisms**: Our classification models predict virality based on observed patterns, but they do not capture the underlying decision-making process behind content sharing. Causal inference techniques, such as Granger causality or structural causal models (SCMs), could help uncover the key factors driving content diffusion. These methods could

be used to analyze whether user interactions, post timing, or exposure to specific narratives causally influence virality.

- **Exploring alternative machine learning approaches for virality classification**: Our study employed a transformer-based classification model (BERTimbau) with a supervised learning approach. Future work could explore unsupervised and semi-supervised techniques, such as self-supervised contrastive learning, which might help improve generalization to unseen viral trends. Additionally, graph neural networks (GNNs) could be used to model both textual and topological features simultaneously, leveraging the full structure of dissemination networks for prediction.

- **Investigating the role of misinformation and coordinated behavior in content dissemination**: Given the importance of OSMPs in shaping public opinion, future research could analyze how misinformation spreads within dissemination backbones and whether coordinated campaigns play a role in boosting specific narratives. This could involve detecting bot activity, inauthentic engagement patterns, or anomalous clusters that artificially promote certain content.

By addressing these open questions, future research can deepen the understanding of the complex mechanisms driving content virality in social media. Expanding the methodological scope through alternative network modeling, additional feature integration, and advanced machine learning techniques can further enhance the predictive capabilities of virality classification models while improving interpretability and applicability to broader contexts.

Bibliography

[1] Akiko N. Aizawa. An information-theoretic perspective of tf-idf measures. *Inf. Process. Manag.*, 39(1):45–65, 2003.

[2] Nadia Alonso-López, Pavel Sidorenko Bautista, and Fábio Giacomelli. Beyond challenges and viral dance moves: Tiktok as a vehicle for disinformation and fact-checking in spain, portugal, brazil, and the usa. *Anàlisi*, 64:65–84, 06 2021.

[3] Zahra Aminolroaya and Ali Katanforoush. How iranian instagram users act for parliament election campaign? a study based on followee network. In *2017 3th International Conference on Web Research (ICWR)*, pages 1–6, 2017.

[4] Claudio Moisés Valiense De Andrade, Washington Cunha, Guilherme Fonseca, Ana Clara Souza Pagano, Luana De Castro Santos, Adriana Silvina Pagano, Leonardo Chaves Dutra Da Rocha, and Marcos André Gonçalves. Explaining the hardest errors of contextual embedding based classifiers. In Libby Barak and Malihe Alikhani, editors, *Proceedings of the 28th Conference on Computational Natural Language Learning*, pages 419–434, Miami, FL, USA, November 2024. Association for Computational Linguistics.

[5] Arturo Arriagada and Francisco Ibáñez. "you need at least one picture daily, if not, you're dead": Content creators and platform evolution in the social media ecology. *Social Media + Society*, 6(3):2056305120944624, 2020.

[6] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer normalization, 2016.

[7] Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada A. Adamic. The role of social networks in information diffusion. *CoRR*, abs/1201.4145, 2012.

[8] Jack Bandy and Nicholas Diakopoulos. Curating quality? how twitter's timeline algorithm treats different types of news. *Social Media + Society*, 7(3):20563051211041648, 2021.

[9] Peiman Barnaghi, Parsa Ghaffari, and John G. Breslin. Opinion mining and sentiment polarity on twitter and correlation between events and sentiment. In *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 52–57, 2016.

[10] Fabiano Belém, Washington Cunha, Celso França, Claudio Andrade, Leonardo Rocha, and Marcos André Gonçalves. A novel two-step fine-tuning pipeline for cold-start active learning in text classification tasks, 2024.

[11] Fabrício Benevenuto and Philipe Melo. Misinformation campaigns through whatsapp and telegram in presidential elections in brazil. *Commun. ACM*, 67(8):72–77, August 2024.

[12] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11-7), 2016.

[13] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of comm. in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, oct 2008.

[14] Javier Borge-Holthoefer, Walid Magdy, Kareem Darwish, and Ingmar Weber. Content and network dynamics behind egyptian political polarization on twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, page 700–711, New York, NY, USA, 2015. Association for Computing Machinery.

[15] Maria Teresa Borges-Tiago, Flavio Tiago, and Carla Cosme. Exploring users' motivations to participate in viral communication on social media. *Journal of Business Research*, 101:574–582, 2019.

[16] Michael S. Boyd. (new) participatory framework on youtube? commenter interaction in us political speeches. *Journal of Pragmatics*, 72:46–58, 2014. Participation framework revisited: (new) media and their audiences/users.

[17] Kellyton Brito and Paulo Jorge Leitão Adeodato. Machine learning for predicting elections in latin america based on social media engagement and polls. *Government Information Quarterly*, 40(1):101782, 2023.

[18] Andriy Burkov. *The Hundred-Page Machine Learning Book*. 2019.

[19] Josemar Caetano, Samuel Guimarães, Marcelo M. R. Araújo, Márcio Silva, Júlio C. S. Reis, Ana P. C. Silva, Fabrício Benevenuto, and Jussara M. Almeida. Characterizing early electoral advertisements on twitter: A brazilian case study. In Frank Hopfgartner, Kokil Jaidka, Philipp Mayr, Joemon Jose, and Jan Breitsohl, editors, *Social Informatics*, pages 257–272, Cham, 2022. Springer International Publishing.

[20] Josemar Alves Caetano, Helder Seixas Lima, Mateus F. Santos, and Humberto Torres Marques-Neto. Using sentiment analysis to define twitter political users' classes

and their homophily during the 2016 american presidential election. *J. Internet Serv. Appl.*, 9(1):18:1–18:15, 2018.

[21] Laura Cervi and Carles Marín-Lladó. What are political parties doing on tiktok? the spanish case. *El Profesional de la información*, 30, 07 2021.

[22] François Chollet. *Deep Learning with Python.* Manning, second edition, 2021.

[23] Gregory W. Corder and Dale I. Foreman. *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach.* John Wiley & Sons, 2nd edition, June 2014.

[24] Leonardo Costa, Graça Rossetto, and Tiago Franklin Rodrigues Lucena. Personalização e positividade dos candidatos a prefeito uma análise do uso do instagram durante as eleições de 2020 em maringá-pr. *Compolítica*, 12:113–142, 04 2023.

[25] André Cristiani, Douglas Lieira, and Heloisa Camargo. A sentiment analysis of brazilian elections tweets. In *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*, pages 153–160, Porto Alegre, RS, Brasil, 2020. SBC.

[26] Jose Martins da Rosa, Renan Saldanha Linhares, Carlos Henrique Gomes Ferreira, Gabriel P. Nobre, Fabricio Murai, and Jussara M. Almeida. Uncovering discussion groups on claims of election fraud from twitter. In *Proc. of Social Informatics: 13th International Conference*, 2022.

[27] Morris H. DeGroot and Mark J. Schervish. *Probability and Statistics.* Addison Wesley, 3rd edition, 2002.

[28] Anastasia Denisova. Viral journalism. strategy, tactics and limitations of the fast spread of content on social media: Case study of the united kingdom quality publications. *Journalism*, 24(9):1919–1937, 2023.

[29] Suely Ferreira Deslandes. O ativismo digital e sua contribuição para a descentralização a política. *Cien. Saude Colet.*, 23(10):3133–3136, 2018.

[30] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics, 2019.

[31] Karina Di Nubila, Carlos A. Ballesteros-Herencia, Dunia Etura, and Virginia Martín-Jiménez. Technopopulism and politainment in brazil: Bolsonaro government's weekly youtube broadcasts. *Media and Communication*, 11(2):137–147, 2023.

[32] Jesse Dodge, Gabriel Ilharco, Roy Schwartz, Ali Farhadi, Hannaneh Hajishirzi, and Noah A. Smith. Fine-tuning pretrained language models: Weight initializations, data orders, and early stopping. *CoRR*, abs/2002.06305, 2020.

[33] Elizabeth Dubois and Devin Gaffney. The multiple facets of influence: Identifying political influentials and opinion leaders on twitter. *American Behavioral Scientist*, 58(10):1260–1277, 2014.

[34] Roman Egger and Joanne Yu. A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in Sociology*, 7, 05 2022.

[35] Rafael Evangelista and Fernanda Bruno. Whatsapp and political instability in brazil: targeted messages and political radicalisation. *Internet Policy Review*, 8(4):1–23, 2019.

[36] Anna Fang and Zina Ben-Miled. Does bad news spread faster? In *2017 International Conference on Computing, Networking and Communications (ICNC)*, pages 793–797, 2017.

[37] Juliana Feitosa, Luiz De Camargo, Eloisa Bonatti, Giovanna Simioni, and José Brega. explainable artificial intelligence in sentiment analysis of posts about covid-19 vaccination on twitter. In *Proceedings of the 29th Brazilian Symposium on Multimedia and the Web*, page 65–72, Porto Alegre, RS, Brasil, 2023. SBC.

[38] Emilio Ferrara, Herbert Chang, Emily Chen, Goran Muric, and Jaimin Patel. Characterizing social media manipulation in the 2020 u.s. presidential election. *First Monday*, 2020.

[39] Carlos Henrique Gomes Ferreira, Fabricio Murai, Ana Paula Couto da Silva, Jussara Marques de Almeida, Martino Trevisan, Luca Vassio, Idilio Drago, and Marco Mellia. Unveiling community dynamics on instagram political network. In *Proceedings of Web Science*, 2020.

[40] Carlos HG Ferreira, Fabricio Murai, Ana PC Silva, Jussara M Almeida, Martino Trevisan, Luca Vassio, Marco Mellia, and Idilio Drago. On the dynamics of political discussions on instagram: A network perspective. *Online Social Networks and Media*, 25:100155, 2021.

[41] Jeffrey A Fine and Megan F Hunt. Negativity and elite message diffusion on social media. *Polit. Behav.*, 45(3):955–973, September 2023.

[42] E Fonseca, L Santos, Marcelo Criscuolo, and S Aluisio. Assin: Avaliacao de similaridade semantica e inferencia textual. In *Computational Processing of the Portuguese Language-12th International Conference, Tomar, Portugal*, pages 13–15, 2016.

[43] Kiran Garimella and Ingmar Weber. A long-term analysis of polarization on twitter. 2017.

[44] Carlos Henrique Gomes Ferreira, Fabricio Murai, Ana Paula Couto da Silva, Jussara Marques de Almeida, Martino Trevisan, Luca Vassio, Idilio Drago, and Marco Mellia. Unveiling community dynamics on instagram political network. In *Proceedings of the 12th ACM Conference on Web Science*, pages 231–240, 2020.

[45] Carlos Henrique Gomes Ferreira, Fabricio Murai, Ana P. C. Silva, Martino Trevisan, Luca Vassio, Idilio Drago, Marco Mellia, and Jussara M. Almeida. On network backbone extraction for modeling online collective behavior. *PLOS ONE*, 17(9):1–36, 09 2022.

[46] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[47] Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.

[48] Joshua Guberman, Carol E. Schmitz, and Libby Hemphill. Quantifying toxicity and verbal violence on twitter. In Darren Gergle, Meredith Ringel Morris, Pernille Bjørn, and Joseph A. Konstan, editors, *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2015, San Francisco, CA, USA, February 27 - March 2, 2016, Companion Volume*, pages 277–280. ACM, 2016.

[49] Andrew M. Guess, Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Edward Kennedy, Young Mie Kim, David Lazer, Devra Moehler, Brendan Nyhan, Carlos Velasco Rivera, Jaime Settle, Daniel Robert Thomas, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Beixian Xiong, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, 381(6656):398–404, 2023.

[50] Anna Guimaraes, Oana Balalau, Erisa Terolli, and Gerhard Weikum. Analyzing the traits and anomalies of political discussions on reddit. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01):205–213, Jul. 2019.

[51] Daniel Halpern, Sebastián Valenzuela, and James E. Katz. We Face, I Tweet: How Different Social Media Influence Political Participation through Collective and Internal Efficacy. *Journal of Computer-Mediated Communication*, 22(6):320–336, 11 2017.

[52] Trevor Hastie, Robert Tibshirani, and Jerome H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Edition*. Springer Series in Statistics. Springer, 2009.

[53] Michelle Hood and Amanda L. Duffy. Understanding the relationship between cyber-victimisation and cyber-bullying on social network sites: The role of moderating factors. *Personality and Individual Differences*, 133:103–108, 2018. Examining Personality and Individual Differences in Cyberspace.

[54] Ruben Interian and Francisco A Rodrigues. Group polarization, influence, and domination in online interaction networks: a case study of the 2022 brazilian elections. *Journal of Physics: Complexity*, 4(3):035008, September 2023.

[55] Anna Sophie Kümpel Jörg Haßler and Jessica Keller. Instagram and political campaigning in the 2017 german federal election. a quantitative content analysis of german top politicians' and parliamentary parties' posts. *Information, Communication & Society*, 26(3):530–550, 2023.

[56] Andressa Kappaun and Jonice Oliveira. Análise sobre viés de gênero no youtube: Um estudo sobre as eleições presidenciais de 2018 e 2022. In *Anais do XII Brazilian Workshop on Social Network Analysis and Mining*, pages 127–138, Porto Alegre, RS, Brasil, 2023. SBC.

[57] Siddhartha Krishnan, Jose Cadena, and Naren Ramakrishnan. The dynamics of competing cascades in social media : Applications to agenda setting. 2014.

[58] André Luiz Martins Lemos, Elias Cunha Bitencourt, and João Guilherme Bastos dos Santos. Fake news as fake politics: the digital materialities of youtube misinformation videos about brazilian oil spill catastrophe. *Media, Culture & Society*, 43(5):886–905, 2021.

[59] Juliana Lima, Maria Santana, Andreiwid Correa, and Kellyton Brito. The use and impact of tiktok in the 2022 brazilian presidential election. In *Proceedings of the 24th Annual International Conference on Digital Government Research*, dg.o '23, page 144–152, New York, NY, USA, 2023. Association for Computing Machinery.

[60] Renan S. Linhares, José M. Rosa, Carlos H. G. Ferreira, Fabricio Murai, Gabriel Nobre, and Jussara Almeida. Uncovering coordinated communities on twitter during the 2020 u.s. election. In *Proc. of ASONAM*, 2022.

[61] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692, 2019.

[62] Caio Machado, Beatriz Kira, Vidya Narayanan, Bence Kollanyi, and Philip Howard. A study of misinformation in whatsapp groups with a focus on the brazilian presidential elections. In *Companion Proceedings of The 2019 World Wide Web Conference*, WWW '19, page 1013–1019, New York, NY, USA, 2019. Association for Computing Machinery.

[63] Suman Maity, Aman Kharb, and Animesh Mukherjee. Language use matters: Analysis of the linguistic structure of question texts can characterize answerability in quora. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1):612–615, May 2017.

[64] Larissa Malagoli, Giovana Piorino, Carlos Ferreira, and Ana Silva. Twitter and the 2022 brazilian elections portrait: A network and content-driven analysis. In *Proceedings of the 30th Brazilian Symposium on Multimedia and the Web*, pages 283–291, Porto Alegre, RS, Brasil, 2024. SBC.

[65] Larissa G. Malagoli, Julia Stancioli, Carlos H. G. Ferreira, Marisa Vasconcelos, Ana Paula Couto da Silva, and Jussara M. Almeida. A look into covid-19 vaccination debate on twitter. In *Proceedings of the 13th ACM Web Science Conference 2021*, WebSci '21, page 225–233, New York, NY, USA, 2021. Association for Computing Machinery.

[66] Christopher D. Manning and Hinrich Schütze. *Foundations of statistical natural language processing*. MIT Press, 2001.

[67] Ricardo Martins, Marco Gomes, José João Almeida, Paulo Novais, and Pedro Henriques. Hate speech classification in social media using emotional analysis. In *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, pages 61–66, 2018.

[68] Philip May. Machine translated multilingual sts benchmark dataset. 2021.

[69] Tanushree Mitra, Scott Counts, and James Pennebaker. Understanding anti-vaccination attitudes in social media. In *Proc. of ICWSM*, 2016.

[70] Virginia Morini, Laura Pollacci, and Giulio Rossetti. Toward a standard approach for echo chamber detection: Reddit case study. *Applied Sciences*, 11(12), 2021.

[71] A.C. Müller and S. Guido. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, 2016.

[72] Kevin P. Murphy. *Machine learning - a probabilistic perspective.* Adaptive computation and machine learning series. MIT Press, 2012.

[73] Kevin P. Murphy. *Machine learning : a probabilistic perspective.* MIT Press, Cambridge, Mass. [u.a.], 2013.

[74] Debashis Naskar, Sanasam Ranbir Singh, Durgesh Kumar, Sukumar Nandi, and Eva Onaindia de la Rivaherrera. Emotion dynamics of public opinions on twitter. *ACM Trans. Inf. Syst.*, 38(2):18:1–18:24, 2020.

[75] Carlos Navarrete, Mariana Macedo, Rachael Colley, Jingling Zhang, Nicole Ferrada, Maria Eduarda Mello, Rodrigo Lira, Carmelo Bastos-Filho, Umberto Grandi, Jerome Lang, and César A. Hidalgo. Understanding political divisiveness using online participation data from the 2022 french and brazilian presidential elections, 2023.

[76] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, Feb 2004.

[77] Leonardo Nizzoli, Serena Tardelli, Marco Avvenuti, Stefano Cresci, and Maurizio Tesconi. Coordinated behavior on social media in 2019 uk general election. *Proceedings of the International AAAI Conference on Web and Social Media*, 15(1):443–454, May 2021.

[78] Gabriel Nobre, Carlos Ferreira, and Jussara Almeida. Beyond groups: Uncovering dynamic communities on the whatsapp network of information dissemination. In *SocInfo' 2020*, 2020.

[79] Gabriel Peres Nobre, Carlos H.G. Ferreira, and Jussara M. Almeida. A hierarchical network-oriented analysis of user participation in misinformation spread on whatsapp. *Information Processing and Management*, 59(1), jan 2022.

[80] Gabriel Peres Nobre, Kecia Aline Marques Ferreira, Ismael Santana Silva, and Glívia Angélica Rodrigues Barbosa. Characterization of public opinion on political events in brazil based on twitter data. pages 105–116. Springer Nature, aug 2018.

[81] Oluwabusayo Okunloye, Kerk Kee, R. Glenn Cummins, and Weiwu Zhang. The linguistic and message features driving information diffusion on twitter: The case of revolutionnow in nigeria. *International Journal of Communication*, 17(0), 2023.

[82] Geovana Oliveira, Otávio Venâncio, Vinícius Vieira, Jussara Almeida, Ana Silva, Ronan Ferreira, and Carlos Ferreira. Um framework para análise bidimensional de disseminação de informações em plataformas de mídias sociais. In *Proceedings of*

the 30th Brazilian Symposium on Multimedia and the Web, pages 301–309, Porto Alegre, RS, Brasil, 2024. SBC.

[83] Diogo Pacheco. Bots, elections, and controversies: Twitter insights from brazil's polarised elections. In *Proceedings of the ACM Web Conference 2024*, WWW '24, page 2651–2659, New York, NY, USA, 2024. Association for Computing Machinery.

[84] Beatriz Paiva, Beatriz Barbosa, Ana Silva, and Mirella Moro. O debate do feminismo no twitter: Um estudo de caso das eleições brasileiras de 2022. In *Anais do XII Brazilian Workshop on Social Network Analysis and Mining*, pages 103–114, Porto Alegre, RS, Brasil, 2023. SBC.

[85] Jason Phang, Thibault Févry, and Samuel R. Bowman. Sentence encoders on stilts: Supplementary training on intermediate labeled-data tasks. *CoRR*, abs/1811.01088, 2018.

[86] Simon J.D. Prince. *Understanding Deep Learning*. The MIT Press, 2023.

[87] Livy Real, Erick Fonseca, and Hugo Goncalo Oliveira. The assin 2 shared task: a quick overview. In *International Conference on Computational Processing of the Portuguese Language*, pages 406–412. Springer, 2020.

[88] Raquel Recuero. fraudenasurnas: estratégias discursivas de desinformação no twitter nas eleições 2018. *Revista Brasileira de Linguística Aplicada*, 20, 08 2020.

[89] Julio Reis, Philipe Melo, Fabiano Belém, Fabricio Murai, Jussara Almeida, and Fabricio Benevenuto. Helping fact-checkers identify fake news stories shared through images on whatsapp. In *Proceedings of the 29th Brazilian Symposium on Multimedia and the Web*, page 159–167, Porto Alegre, RS, Brasil, 2023. SBC.

[90] Filipe N. Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabricio Benevenuto, Oana Goga, Krishna P. Gummadi, and Elissa M. Redmiles. On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on facebook. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 140–149, New York, NY, USA, 2019. Association for Computing Machinery.

[91] Brian Richards. Type/token ratios: what do they really tell us? *Journal of child language*, 14:201–9, 07 1987.

[92] Daniel M. Romero, Brendan Meeder, and Jon Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th International Conference on World*

Wide Web, WWW '11, page 695–704, New York, NY, USA, 2011. Association for Computing Machinery.

[93] Gonzalo A. Ruz, Pablo A. Henríquez, and Aldo Mascareño. Sentiment analysis of twitter data during critical events through bayesian net. class. *Future Gen. Computer Systems*, 106:92–104, 2020.

[94] Maria Santana, Juliana Lima, Andreiwid Correa, and Kellyton Brito. Engajamento no tiktok dos candidatos às eleições brasileiras de 2022 – resultados iniciais. In *Anais do XII Brazilian Workshop on Social Network Analysis and Mining*, pages 151–162, Porto Alegre, RS, Brasil, 2023. SBC.

[95] Daiana Santos and Lilian Berton. Analysis of twitter users' sentiments about the first round 2022 presidential election in brazil. In *Anais do XX Encontro Nacional de Inteligência Artificial e Computacional*, pages 880–893, Porto Alegre, RS, Brasil, 2023. SBC.

[96] Patrícia Santos, Claudio Penteado, Laura Almeida, and Denise Goya. Democracia sob ataque: polarização política e produção de conteúdos hostis no twitter nas eleições de 2022. *Revista Debates*, 1:22, 04 2023.

[97] Daniel Scanfeld, Vanessa Scanfeld, and Elaine L. Larson. Dissemination of health information through social networks: Twitter and antibiotics. *American Journal of Infection Control*, 38(3):182–188, 2010.

[98] M. Ángeles Serrano, Marián Boguñá, and Alessandro Vespignani. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*, 106(16):6483–6488, April 2009.

[99] Sarah Silva and Elaine Faria. Análise de sentimentos expressos no twitter em relação aos candidatos da eleição presidencial de 2022. In *Anais do XII Brazilian Workshop on Social Network Analysis and Mining*, pages 79–90, Porto Alegre, RS, Brasil, 2023. SBC.

[100] Valfredo Lima da Silva. Uso das redes sociais como forma de disseminação da informação: Um estudo de caso nas bibliotecas do instituto federal de educação, ciência e tecnologia da bahia (ifba). Master's thesis, Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), 2014.

[101] Rafael Silva Barbon and Ademar Takeo Akabane. Towards transfer learning techniques—bert, distilbert, bertimbau, and distilbertimbau for automatic text classification from different languages: A case study. *Sensors*, 22(21), 2022.

[102] Veronika Solopova, Oana-Iuliana Popescu, Christoph Benzmüller, and Tim Landgraf. Automated multilingual detection of pro-kremlin propaganda in newspapers and telegram posts. *Datenbank-Spektrum*, 23:1–10, 03 2023.

[103] Fábio Souza, Rodrigo Nogueira, and Roberto Lotufo. 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)*, 2020.

[104] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[105] Chi Sun, Xipeng Qiu, Yige Xu, and Xuanjing Huang. How to fine-tune bert for text classification? In Maosong Sun, Xuanjing Huang, Heng Ji, Zhiyuan Liu, and Yang Liu, editors, *Chinese Computational Linguistics*, pages 194–206, Cham, 2019. Springer International Publishing.

[106] Yla Tausczik and James Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54, 2010.

[107] Sebastián Valenzuela, Ingrid Bachmann, and Matías Bargsted. *The Personal Is the Political? What Do WhatsApp Users Share and How It Matters for News Knowledge, Polarization and Participation in Chile*, pages 26–46. 01 2023.

[108] José Van Dijck and Thomas Poell. Understanding social media logic. *Media and Communication*, 1:2–14, 06 2013.

[109] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca. *Python Deep Learning: Exploring Deep Learning Techniques and Neural Network Architectures with Pytorch, Keras, and TensorFlow, 2nd Edition*. Packt Publishing, 2019.

[110] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008, 2017.

[111] Otavio R. Venâncio, Carlos H. G. Ferreira, Jussara M. Almeida, and Ana Paula C. da Silva. Unraveling user coordination on telegram: A comprehensive analysis of political mobilization during the 2022 brazilian presidential election. *Proceedings of the International AAAI Conference on Web and Social Media*, 18(1):1545–1556, May 2024.

[112] Jun-Ming Xu, Kwang-Sung Jun, Xiaojin Zhu, and Amy Bellmore. Learning from bullying traces in social media. In *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 3-8, 2012, Montréal, Canada*, pages 656–666. The Association for Computational Linguistics, 2012.

[113] Shlomo Yitzhaki. Relative deprivation and the gini coefficient. *The quarterly journal of economics*, pages 321–324, 1979.

Appendix A

Most Disseminated Topics Details

In this appendix, we detail some of the data presented for the characterization of the network in Chapter 4. The tables for the top 20 most disseminated topics of the three different sets analyzed are displayed next, this time with the original most discriminative words found in each topic in Portuguese.

Table A.1: Top discussion topics found on Twitter in Portuguese.

ID	# Tweets	# Retweets	Most Discriminative Words	Description
1	587	48,978	inácio, silva, luiz, presidente, biden, eleito, vitória, novo, brasil, luis	Discusses President Lula's victory in the election in the second round and the possibility of his upcoming victory during the first round. Cites Joe Biden, president of the United States, who was one of the first international figures to recognize Lula's election.
2	473	30,496	eleicoes2022, vira, virou, lulanoprimeiroturno13, nordeste, eleições2022, lulapresidente1, eleicao2022, virada, lulinha	Regarding the turnaround in votes that Lula had, when the votes from the northeast began to be counted.
3	76	28,892	apoiadores, comemorando, virada, vitória, festa, comemorar, ruas, brasília, contra, petista	Refers to Lula's victory and the voters' celebration.
4	167	24,892	via, terceira, culpa, simone, chance, segundo, ter, ciro, vcs, votaram	Mentions the discourse of a third way of opposition to Lula and Bolsonaro, quoting candidates Ciro Gomes and Simone Tebet.
5	97	22,882	história, vezes, vez, reeleição, desde, 1º, presidente, reeleger, mandato, consegue	Topic that debates about the possibility of reelection of Bolsonaro and the fact of him being the first Brazilian president to not be re-elected.
6	145	18,158	luto, mil, pandemia, 700, covid, mortes, mortos, pessoas, perdeu, durante	Issues and fatalities that occurred during Bolsonaro's administration in the COVID-19 pandemic period.
7	293	14,738	acabou, pesadelo, adeus, venceu, tchau, lulapresidente2022, fim, bem, acima, finalmente	Electoral opponents of the Bolsonaro government celebrating the election results.
8	65	13,181	deputado, federal, paulo, ferreira, votado, mg, nikolas, eleito, paraná, senador	Mentions the State's Elections for House of Representatives and Senate.
9	42	12,755	eleita, federal, mulher, primeira, paulo, histórico, câmara, todas, damares, contra	Comments on the electoral victory of women for the position of congresswomen.
10	25	11,790	burra, gerais, votando, outras, votar, nele, suficiente, regiões, minas, direito	Criticizes voters for their decision to vote on polemic candidates from far right, including Bolsonaro.
11	124	11,604	obrigado, obrigada, parabéns, deus, boa, democracia, país, todos, senhor, acima	People celebrating and thanking the Brazilian democracy regime with Lula's election.
12	60	11,202	perdeu, neymar, cair, igual, perder, caindo, pau, copa, desse, nessa	Mentions terms related to the World Cup, which took place close to the election period.
13	40	11,181	lo, lulapresidente2022, bora, vira, grande, lulapresidente1, vitória, momento, luiz, ouvir	Talks about Lula's victory and his first speech.
14	53	11,140	urgente, faltam, apenas, menos, falta, mil, vitória, lulanoprimeiroturno13, eleições2022, dar	Refers to the first round when Lula led with 48.43% of the votes and almost was elected and the victory of Lula in second round.
15	51	10,758	tava, fraude, ganhando, ganhou, boa, né, falando, cima, the, virou	Debates about the turnaround, with some users using the discourse of electoral fraud.
16	168	10,030	zema, minas, nikolas, estranho, gerais, algo, mg, ganhando, errado, vota	Discussion of the voting outcomes for the state of Minas Gerais, debating on how the senate and governor votes were for far right candidates, but the most voted for president in the region was Lula.
17	477	9,902	votou, simone, branco, nulo, cu, votaram, ciro, tomar, vcs, pau	Critics on null votes and about votes for the third and fourth place candidates of the presidential election.
18	53	9,704	estados, lidera, nordeste, todos, gerais, liderando, minas, bahia, região, mato	Comments on the regions of Brazil that Lula was leading the dispute.
19	263	9,454	rua, vc, ta, alguma, fica, oq, alguém, qualquer, tweet, gente	A topic with common used words in tweets in Portuguese, commenting the event.
20	41	9,367	primeira, mulher, eleita, federal, petista, paulo, nova, algo, apoiar, sp	Discusses the first trans women elected for different Brazilian states as congresswomen.

Table A.2: Top 20 discussion topics found on Twitter for backbones in Portuguese.

ID	# Tweets	# Retweets	Most Discriminative Words	Description (New Topics Only)
1	587	8,005	inácio, silva, luiz, presidente, biden, eleito, vitória, novo, brasil, luis	
2	473	5,865	eleicoes2022, vira, virou, lulanoprimeiroturno13, nordeste, eleições2022, lulapresidente1, eleicao2022, virada, lulinha	
3	76	3,893	apoiadores, comemorando, virada, vitória, festa, comemorar, ruas, brasília, contra, petista	
4	79	3,684	diferença, cai, milhão, menos, caiu, mil, 46, apenas, eleições2022, votos	Highlights the difference between Lula's and Bolsonaro's votes
5	53	3,112	urgente, faltam, apenas, menos, falta, mil, vitória, lulanoprimeiroturno13, eleições2022, dar	
6	97	2,987	história, vezes, vez, reeleição, desde, 1º, presidente, reeleger, mandato, consegue	
7	42	2,732	eleita, federal, mulher, primeira, paulo, histórico, câmara, todas, damares, contra	
8	56	2,418	deve, minutos, prf, 19, 10, próximos, neste, noite, globo, campanha	Discuss the projections by major news agencies, which estimate that Lula would surpass Bolsonaro in votes
9	41	2,179	primeira, mulher, eleita, federal, petista, paulo, nova, algo, apoiar, sp	
10	177	2,126	2º, datafolha, 1º, turno, segundo, presidencial, eleições2022, governador, vão, precisar	DataFolha survey indicating a high likelihood of second-round runoffs for the presidential race
11	53	1,976	estados, lidera, nordeste, todos, gerais, liderando, minas, bahia, região, mato	
12	167	1,874	via, terceira, culpa, simone, chance, segundo, ter, ciro, vcs, votaram	
13	65	1,838	deputado, federal, paulo, ferreira, votado, mg, nikolas, eleito, paraná, senador	
14	45	1,776	vantagem, sobre, segue, apuradas, 47, 90, milhão, urnas, liderança, quase	After the majority of voting machine results were cleared, Lula was leading the race, sparking widespread discussion among voters
15	83	1,667	nordeste, chegando, sempre, norte, pará, eleicoes2022, eleicao2022, bahia, eleições2022, região	Tweets celebrating the Northeast region votes were being counted, which significantly impacted the voting results in favor of Lula
16	69	1,652	saúde, quer, educação, povo, bem, liberdade, vida, viver, governar, porque	Concerns about education and health issues
17	40	1,565	lo, lulapresidente2022, bora, vira, grande, lulapresidente1, vitória, momento, luiz, ouvir	
18	64	1,526	virada, gostoso, lulinha, lulapresidente, lulanoprimeiroturno13, calma, eleicoes2022, deus, virou, nordeste	Tweets with the use of "He who laughs last, laughs best" to comment on Lula's victory in the election results.
19	293	1,454	acabou, pesadelo, adeus, venceu, tchau, lulapresidente2022, fim, bem, acima, finalmente	
20	37	1,443	amazonas, pandemia, durante, votam, nele, vergonha, parece, muitos, liderando, lidera	Controversial outcomes arose from the presidential election in the state of Amazonas, due to Bolsonaro's actions during the COVID-19 crisis[1]

Table A.3: Top 20 discussion topics found on Twitter for persistent users in Portuguese.

ID	# Tweets	# Retweets	Most Discriminative Words	Description (New Topics Only)
1	587	1,828	inácio, silva, luiz, presidente, biden, eleito, vitória, novo, brasil, luis	
2	473	1,322	eleicoes2022, vira, virou, lulanoprimeiroturno13, nordeste, eleições2022, lulapresidente1, eleicao2022, virada, lulinha	
3	76	886	apoiadores, comemorando, virada, vitória, festa, comemorar, ruas, brasília, contra, petista	
4	79	786	diferença, cai, milhão, menos, caiu, mil, 46, apenas, eleições2022, votos	
5	97	615	história, vezes, vez, reeleição, desde, 1º, presidente, reeleger, mandato, consegue	
6	53	582	urgente, faltam, apenas, menos, falta, mil, vitória, lulanoprimeiroturno13, eleições2022, dar	
7	42	576	eleita, federal, mulher, primeira, paulo, histórico, câmara, todas, damares, contra	
8	56	557	deve, minutos, prf, 19, 10, próximos, neste, noite, globo, campanha	
9	177	541	2º, datafolha, 1º, turno, segundo, presidencial, eleições2022, governador, vão, precisar	
10	41	470	primeira, mulher, eleita, federal, petista, paulo, nova, algo, apoiar, sp	
11	45	439	vantagem, sobre, segue, apuradas, 47, 90, milhão, urnas, liderança, quase	
12	69	424	saúde, quer, educação, povo, bem, liberdade, vida, viver, governar, porque	
13	53	399	estados, lidera, nordeste, todos, gerais, liderando, minas, bahia, região, mato	
14	65	398	deputado, federal, paulo, ferreira, votado, mg, nikolas, eleito, paraná, senador	
15	49	347	bahia, 30, milhões, apurados, ainda, pessoas, dia, votando, falta, votaram	Talks about Lula having the majority of votes in the state of Bahia, a strong electoral college in the Northeast region, during the counting
16	40	343	lo, lulapresidente2022, bora, vira, grande, lulapresidente1, vitória, momento, luiz, ouvir	
17	83	314	nordeste, chegando, sempre, norte, pará, eleicoes2022, eleicao2022, bahia, eleições2022, região	
18	37	307	amazonas, pandemia, durante, votam, nele, vergonha, parece, muitos, liderando, lidera	
19	64	300	virada, gostoso, lulinha, lulapresidente, lulanoprimeiroturno13, calma, eleicoes2022, deus, virou, nordeste	
20	167	251	via, terceira, culpa, simone, chance, segundo, ter, ciro, vcs, votaram	