JeDEM Issue 16 (1): 186-212, 2024
ISSN 2075-9517
http://www.jedem.org
Date of submission: 08.09.2023
Date of acceptance: 21.03.2024
Article DOI: 10.29379/jedem.v16i1.846

Unveiling political polarization on Twitter: Machine learning and sentiment analysis in presidential elections

David Valle-Cruz

Author ORCID Nr 0000-0002-5204-809
Unidad Académica Profesional Tianguistenco, UAEMéx
Paraje El Tejocote, San Pedro Tlaltizapán, Tianguistenco, Estado de México, davacr@uaemex.mx

Rodrigo Sandoval-Almazán

Author ORCID Nr 0000-0002-7864-6464
Department of Political Science, UAEMéx
Instituto Literario No. 100
Toluca, Estado de México, rsandovala@uaemex.mx

Asdrúbal López-Chau

Author ORCID Nr 0000-0001-5254-0939
CU UAEM Zumpango, UAEMéx
K.M. 3.5, Camino Viejo a Jilotzingo, Valle Hermoso, Zumpango, Estado de México, 55600, México, alchau@uaemex.mx

J. Ignacio Criado

Author ORCID Nr 0000-0002-9184-9696
Autonomous University of Madrid
Department of Political Science
Madrid, Spain, ignacio.criado@uam.es

Abstract: The year 2024 stands out as a pivotal year marked by significant political transformations across the globe. Some countries, such as Mexico and the United States, could be deeply affected by political polarization and echo chambers. This study employed sentiment analysis and machine learning techniques to investigate political polarization on Twitter during the 2018 Mexican presidential election. The findings reveal that the winning candidate exhib-

ited the highest level of polarization. This underscores the pivotal role of social media in elections. For some time now, social media platforms like Twitter have contributed to intensified political polarization and the creation of echo chambers. Further research is essential to understand the influence of polarization on voter decision-making and democratic procedures. Establishing ethical guidelines for using machine learning in policy analysis is critical to preserving the integrity of democratic processes while reaping the potential benefits of new technologies.

Keywords: political polarization, presidential elections, artificial intelligence, machine learning, sentiment analysis

1. Introduction

2024 stands out as a pivotal year marked by significant political transformations across the globe. Some countries, such as the US and Mexico, will be deeply affected by political polarization and echo chambers. These phenomena can distort reality, impede informed decision-making, and contribute to societal fragmentation. In an era of fake news, deepfake content, opaque advertising, and misinformation, society is witnessing a proliferation of infodemics, disinformation, and conspiracy theories that leave people deeply divided and trapped in echo chambers. This disturbing trend has led to new forms of nihilism in which emotional communication overrides rationality and makes consensus-building increasingly difficult (Han, 2022). Political polarization, the sharp division of public opinion into opposite sides, is exacerbated by the widespread use of the Internet and machine learning techniques in political campaigns, reinforcing existing beliefs and opinions (Barberá et al., 2015).

A striking manifestation of this polarization occurs on social media platforms, especially intensifying before elections, leading to clashes between candidates and voters (Lindqvist & Östling, 2008). Extreme party candidates, such as Donald Trump, have successfully used social media to influence the public and spread their ideologies (Valle-Cruz et al., 2023). Social media has emerged as an important arena for political participation, enabling the discussion and expression of political ideologies, as well as providing valuable data for the analysis of various political phenomena (Nulty et al., 2016). Harnessing the power of artificial intelligence and machine learning, sentiment analysis, and other algorithms has been a key help in extracting insights from social media posts (Morozov, 2018). Nevertheless, the dual purpose of the Internet has greatly exacerbated the problem of polarization and echo chambers, with Twitter emerging as an important virtual space for political discussion (Takikawa & Nagayoshi, 2017).

In the backdrop of the 2018 elections in Mexico, a pronounced air of polarization permeated the political landscape (Castro, 2023; Cornejo & Langston, 2024). The intensity of this polarization was notably driven by the opposition candidate's pursuit of the presidency for the third consecutive time, cultivating an atmosphere steeped in dissatisfaction and disenchantment among citizens toward the incumbent ruling party. The polarized electoral campaign ultimately culminated in the victory of the opposition candidate, Andrés Manuel López Obrador (AMLO), as the President (Hernández-Alcántara, 2019). This campaign was marked by heightened polarization, the dissemination of disinformation, and the promise of changing the political regime, emblematic of prevalent trends in contemporary political campaigns. The insights gained and lessons learned from navigating such phenomena have the potential to offer valuable perspectives that extend beyond Mexico's borders, providing a lens to comprehend analogous situations in electoral events globally.

Recognizing the urgency of understanding the nature and impact of political polarization in the upcoming presidential elections, this article uses machine learning techniques to analyze the dynamics of polarization on Twitter during the 2018 Mexican presidential election. Specifically, the authors explore the levels of positivity, negativity, and neutrality associated with the five candidates of the 2018 Mexican presidential election (Estrada & Rawnsley, 2021). The question that leads this research is: How was political polarization between candidates on Twitter in the Mexican presidential elections? To answer this question, the authors applied a novel methodology based on sentiment analysis and machine learning techniques (López-Chau et al., 2020; Valle-Cruz et al., 2020). The rest of the paper is as follows: The second section presents a literature review of political polarization and echo chambers in social media. The third section describes the methodology for analyzing candidates' tweets during the political election period. The fourth section presents the results. The fifth section shows the discussion. Moreover, the final section shows the conclusions, limitations, and future directions.

2. Literature review

In this section, the authors briefly review the literature, which is twofold. The first part focuses on two prominent phenomena of political campaigning on social media: echo chambers and political polarization. The second part explores related research on the application of machine-learning techniques in political elections.

2.1. Political polarization and echo chambers in social media

2.1.1. The impact of social media on political processes

Exploring the impact of social media on political processes necessitates a nuanced examination of several interconnected dimensions that collectively define the contemporary socio-political landscape. Key elements, such as public opinion, cultural stereotypes, and user types, provide valuable insights into the intricate influences wielded by social media. This exploration broadens its scope to encompass significant contextual factors, like culture, ideology, temporality, beliefs, war, revolution, and pandemics, where the role of social media is accentuated in shaping narratives and responses. Delving into how social communications affect interactions within social groups and institutions, influence political processes, and incite civic engagement offers a comprehensive understanding of social media's multifaceted impact on the political sphere. These interconnected topics collectively contribute to unraveling the complexities and nuances that characterize the intricate relationship between social media and political processes in contemporary society.

Bond et al. (2012) conducted one of the initial pioneering studies on the influence of social media on elections. Their research analyzed the messages transmitted to 61 million Facebook users during the 2010 US congressional elections. The study revealed a significant impact of these messages in terms of disseminating both online and offline behaviors within human social networks. In their study, Zhuravskaya et al. (2020) examine the impact of the Internet and social media on voting, street protests, and political attitudes. They identify various direct and indirect relationships between different technologies and regimes' censorship, surveillance, and propaganda tactics. Building on this research, Scheffler and Miller (2021) propose four distinct effects that social media can have: (1) a weakening effect on strong democratic regimes, (2) an intensifying effect on strong authoritarian regimes, (3) a radicalizing effect on weak democratic regimes, and (4) a destabilizing effect on weak authoritarian regimes.

Brito Adeodato (2022) conducted a recent study on the influence of social media on presidential elections in Argentina (2019), Brazil (2018), Colombia (2018), and Mexico (2018). Their research revealed significant statistical correlations between social media performance and voting outcomes. In a subsequent study, the same scholars utilized Machine Learning techniques on the same dataset and achieved a remarkable level of accuracy in predicting the final vote share of candidates. Additionally, their daily predictions proved to be competitive with, or even superior to, traditional polling methods (Brito & Adeodato, 2023).

In this regard, Halich's (2023) investigation into the impact of social communications on public opinion, cultural stereotypes, and user types, particularly in contexts like war, revolution, and pandemics, contributes to understanding how social communications influence interactions among social groups and institutions, including political processes, civic engagement, and youth culture. Alodat, Al-Qora'n, and Abu Hamoud's (2023) study explored the impact of social media on political participation among Jordanian youth, identifying social media as a significant and favorable factor influencing political engagement. Zhang, Kübler, & Dong's (2023) examination of Chinese public perceptions of the EU and China-EU relations focuses on the influence of social media use. Utilizing original survey data from China in 2020, the study establishes that social media use and socioeconomic factors are significant predictors of Chinese public perceptions, emphasizing the variation in effects across different platforms.

Chang's (2019) research reveals that while the Internet significantly encouraged voting turnout in the 2014 election, this effect was not observed in the 2012 or 2016 presidential and legislative elections. The findings highlight the specific influence of the Sunflower Movement on young adults' political engagement in the 2014 election, underscoring the dynamic and diverse nature of the Internet's influence on offline political participation. Aksenov's (2019) examination of the evolving political process and media coverage, particularly within the British Broadcasting Corporation (BBC), suggests that the BBC may not be entirely unbiased in its representation of Syrian events. The study scrutinizes the quality of information provided by the BBC through social media, using the example of civic confrontation in the Syrian Arab Republic in the 2010s, providing insights into current media practices.

Park's (2019) findings show a positive association between the use of social media for political news and knowledge about political issues, although not with knowledge about political processes. The study also uncovers that professional media use for political news is significantly associated with political issue knowledge and political process knowledge. The impact of social media on political issue knowledge increases when combined with professional media news use. Additionally, the study reveals that engaging in political talk reinforces the positive association between social media use for news and political issue knowledge.

The pervasive influence of social media on political processes plays a pivotal role in shaping contemporary discourse. It fuels political polarization as individuals encounter tailored content that aligns with their existing beliefs, contributing to the formation of echo chambers where diverse perspectives are diminished, further deepening ideological divides within society.

2.1.2. Echo chambers

Echo chambers appeared in tandem with advances in information technology that were originally associated with mass media. These refer to the phenomenon of individuals driven by a particular ideology consuming information only from sources consistent with their beliefs. This creates an isolated communicative niche or ghetto that limits a wider view (Garrett, 2009). Although certain studies demonstrate interconnectivity and information flow across ideological boundaries on platforms, such as Twitter, some hardcore partisan communities remain largely segregated. Research has found that information sharing occurs primarily among people with similar ideological preferences and that liberals are likelier than conservatives to engage in cross-ideological dissemination (Barberá, 2015). Active participation within echo chambers significantly impacts community emotional behavior, with greater participation correlated with a passive approach (Del Vicario et al., 2016). Moreover, highly active users tend to shift more quickly in the negative direction compared to less active users. Filter bubbles and echo chambers have been identified in Twitter communities in some countries (Bruns, 2017).

Although the concept of echo chambers continues to be debated, there is evidence that individuals tend to select information that reinforces their pre-existing beliefs, which contributes to the formation of echo chambers (Justwan et al., 2018). This behavior prevents members of the filter bubble from accessing relevant information and discussions, leading to a systematic distrust of external sources (Nguyen, 2020). Attempting to fill the gaps between echo chambers can degrade network centrality and content assessment (Garimella et al., 2018).

Polarization and echo chambers have been extensively studied in the context of political campaigns and have provided different insights. Right-wing and populist ideologies have used Echo chambers to empower candidates and political parties in various countries (Boulianne et al., 2020). For instance, France's opinion leaders and information seekers tend to avoid echo chambers that spread disinformation and instead focus on building trust (Dubois et al., 2020). The "false consensus effect" highlights the tendency of individuals to perceive public opinion as biased in their favor and serves as an example of the existence of echo chambers (Luzsa & Mayr, 2021). Studies on various social media platforms, such as Facebook, Reddit, and Twitter, consistently reveal the existence of echo chambers while exploring possible mitigation strategies (Terren & Borge, 2021). Political homosexuality on Twitter has been studied and found higher ideological similarities in mutual follower networks compared to non-reciprocal networks (Colleoni et al., 2014). A highly polarized group characterized by the spread of scientific information and conspiracy theories has been identified on Facebook (Batorski & Grzywińska, 2018). Certain platform implementations of newsfeed algorithms have been shown to contribute to creating echo chambers (Cinelli et al., 2020).

Social media has been found to be associated with greater ideological distance between individuals despite exposure to content on the less favorable side of the political spectrum (Flaxman et al., 2016). The tendency of social media users to select information consistent with their beliefs and form polarized groups can have profound implications for information cascades and public debate on social issues; potentially, it can affect the electoral process (McLaughlin et al., 2020).

2.1.3. Political polarization

Verba and Nie (1987) proposed four dimensions of political participation. Voting, campaigning, contacting officials, and collective action. Researchers argue that the Internet promotes political participation, such as studies by Bakker & De Vreese (2011), McLeod et al. (1999), and Verba & Nie, (1987). McLeod et al. (1999), based on a survey of 389 people, found that TV news had a modest indirect effect on political participation. Boulianne (2009) analyzed 38 studies and provided evidence that the Internet has a positive impact on civic participation.

The Internet and machine learning techniques have improved the availability of political information but also strengthened existing opinions (Barberá, 2015). Twitter has emerged as a major platform for political discussion, leading to increasing polarization and the formation of echo chambers (Takikawa & Nagayoshi, 2017).

Polarization predates the Internet, with individuals perceiving their attitudes as rational but prejudiced against opposing groups (Ross & Ward, 1995). Increased polarization hinders consensus-building efforts between different groups (Elkind et al., 2017) and contributes to the nomination and selection of extreme party candidates (Westfall et al., 2015). The use of social media in politics has revolutionized ideological polarization and fuelled disputes between candidates and voters (Lindqvist & Östling, 2008). Social media provides an opportunity to meet different political points of view but can also have a polarising effect (Barberá, 2015). Furthermore, social media experiences can alienate individuals from politics, especially in times characterized by excessive partisanship (Tucker et al., 2018).

Polarization is characterized by extreme and contradictory positions, with moderate views rarely represented (Morales et al., 2015). Echo chambers, where individuals search for content that matches their opinions, further exacerbate the polarization in social media (Boutyline & Willer, 2017; Flaxman et al., 2016). Social media echo chambers amplify polarization through statements, feelings, and opinions. The spread of disinformation and the influence of political influencers also contribute to polarization. Exposure to disinformation can mobilize supporters and demobilize opponents (Recuero et al., 2020). Disinformation can come from politicians themselves or other sources, but the polarization of the elite can influence the political polarization of the masses (Abramowitz & Saunders, 2008; Hetherington, 2001). As represented by President Obama, the persuasive power of elected officials can shape public opinion. For example, Michael and Agur (2018) found that President Obama's net neutrality announcement on Twitter led to an increase in activity, sparking an outcry ahead of traditional media coverage.

In their recent study, Yarchi et al. (2021) introduced the term "affective polarization" to describe expressing emotions and attitudes. Their findings also indicate that political polarization on social media cannot be viewed as a singular phenomenon, as notable variations exist across different platforms. Similarly, Wagner (2021) and Hernandez et al. (2021) have examined polarization in the context of elections, while Lelkes (2021) has focused on the affective polarization within political parties. These studies highlight the ongoing research exploring the connection between polarization and social media.

2.2. Machine learning techniques and political elections

This section presents two main perspectives that emerge in the literature on machine learning and social media campaigns. One focuses on election prediction and management, and the other applies opinion mining, sentiment analysis, and sentiment data processing to analyze candidate-voter interactions.

2.2.1. Election control and prediction

The presence and political polarization of echo chambers in social media have significant implications for machine learning and its application in election campaigns. The literature on machine learning and social media election campaigns reveals two main perspectives: First, some scholars focus on election prediction and control, aiming to use machine learning algorithms to predict election outcomes and influence election campaign strategy interactions. These studies recognize the importance of understanding echo chamber dynamics and polarization in shaping voter behavior and preferences, which can inform the development of more accurate predictive models (Fujiwara et al., 2021; Garimella & Weber, 2017). Second, other researchers have studied the application of opinion mining, sentiment analysis, and affective computing techniques to analyze candidate-voter interactions (Sandoval-Almazan & Valle-Cruz, 2020; Valle-Cruz et al., 2021). These studies use machine learning algorithms to uncover patterns in social media data to gain insight into voter sentiment, identify influential factors, and explore the effects of echo chambers and polarization on political discourse. We aim to understand the impact. Both perspectives emphasize the need to consider the role of echo chambers and political polarization when applying machine learning to election campaigns. Understanding the nature of these phenomena will facilitate a more inclusive and balanced representation of political opinions and the development and implementation of algorithms that mitigate the potential negative effects of echo chambers on political debate and decision-making.

Research on machine learning and social media election campaigns has yielded valuable insights from two main perspectives. The first perspective focuses on election management and prediction, using machine learning techniques to develop algorithms that predict winners under a variety of conditions (Elkind et al., 2017). This body of research also explored various voting algorithms (Faliszewski et al., 2018; Singh & Sawhney, 2018) and election manipulation by influencing voting decisions (Conitzer & Sandholm, 2002). In addition, content analysis, sentiment analysis, and studies of social media platforms, such as Twitter and Facebook, are useful for understanding the political competition during the election process (Jaidka & Ahmed, 2015).

The second subtopic deals with research focused on the Twitter platform and sentiment analysis. Research has shown that the effects of information cascades and that journalists' tweets have a longer lifespan compared to mainstream media tweets and play an important role in quickly disseminating information (Choy et al., 2012). Researchers have studied real-time sentiment analysis of tweets and sentiment change point detection (Srivastava et al., 2015). Furthermore, he proposed a new performance measure for sentiment analysis using natural language processing (NLP) techniques and machine learning (Srivastava & Bhatia, 2017), resulting in improved accuracy through Twitter sentiment mapping (Srivastava et al., 2015). In addition, studies conducted in various countries, such as Germany, Pakistan, and India, have explored the relationship between Twitter messages and election results, thereby revealing the impact of social media on elections (Ikiz et al., 2014; Ali et al., 2022).

However, despite extensive research into election prediction, the results remain controversial. Brito et al. (2021, 2023) conducted a systematic literature review and highlighted the limited success of commonly used volumetric and sentiment analysis on Twitter. They emphasized the need to explore more advanced machine-learning approaches. Their results showed strong correlations between social media metrics and vote gains in presidential elections in Argentina, Brazil, Colombia, and Mexico. These results highlight the importance of further investigating advanced machine-learning techniques to improve election prediction and analysis.

2.2.2. Opinion mining, sentiment analysis, and affective computing

Sentiment analysis has proven to be a valuable tool in a variety of application scenarios and offers great benefits in areas such as buzz monitoring. This approach involves monitoring and tracking consumer reactions to services and products, enabling businesses to assess product demand, assess customer experience, and effectively manage crisis situations (Staiano & Guerini, 2014). To improve the analysis of emotions, various frameworks have been developed that provide guidelines for identifying emotions in various forms of communication, such as physical expression, gestures, speech, and text (Balomenos et al., 2004).

Using machine learning techniques in sentiment analysis has led to promising results. For example, Clark, Morris, and Lomax (2018) use machine learning to estimate the proportion of Brexit votes cast in the UK referendum, measure political sentiment, and generate inferences about relevant outcomes, demonstrating electronic petition data's beneficial nature and versatility. In another study, Quan and Ren (2010) identified eight basic emotions using various classification methods, including decision trees, support vector machines (SVM), and Naive Bayes. As a result, they found that SVM achieved better accuracy than his 95%. In addition, they analyzed a manually annotated corpus of emotion categories, intensities, owners/goals, and verbal expressions, providing valuable insights into sentiment analysis. The latest research on sentiment analysis in elections through social media platforms has achieved accuracy ranging from 87.29% to 95% (Karamouzas, Mademlis, & Pitas, 2022; Patil & Kolhe, 2022; Olimpio da Silva, Losada, & Borondo, 2023).

Further research on sentiment analysis explored social media platforms and natural language processing techniques. Takikawa and Nagayoshi (2017) conducted an analysis of Japanese Twitter data, identified five different political communities, and observed the prevalence of xenophobic sentiment within right-wing communities. Koolagudi and colleagues (2009) proposed a Telugu emotional language corpus and used prosodic parameters to explore basic emotions. While recognizing the challenges associated with accurately defining emotion due to factors such as speaker diversity, linguistic differences, and semantic nuances, their research suggests that emotion in language is used to express subjective experiences, attitudes, and affective states.

In the field of text-based sentiment analysis, Chaffar and Inkpen (2011) introduced supervised learning techniques and machine learning approaches using diverse sentiment-annotated datasets, including headlines, fairy tales, and blogs. They used the 'Bag of Words' representation to compare three classification algorithms: Decision Trees, Naive Bayes, and Sequential Minimal Optimization SVM. The average accuracy of sentiment analysis achieved in their study was about 80%. Santander et al. (2017) examined the application of sentiment analysis to political contexts. They conducted research in Chile that focused on analyzing social media data for predicting election outcomes, particularly in presidential primaries. Their research showed the potential of using machine learning to analyze political communications. Moreover, the Corpus Regis case, in which a computer was elected governor of California, demonstrates the influence of machines influencing voters through perceptions of rationality and fairness, as shown in election polls (Genesereth, 2018). For several years, the analysis of various political phenomena has been approached through affective computing (Valle-Cruz et al., 2023), which combines research topics of emotion recognition and sentiment analysis and can be performed with unimodal or multimodal data (Wang et al., 2022).

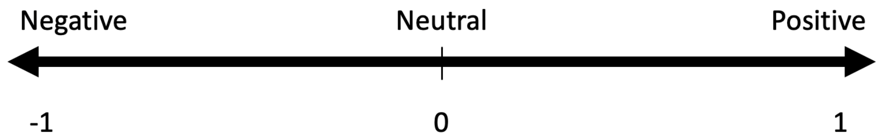
3. Method

In this section, we describe the methodology employed in our study. To carry out the polarity analysis of tweets, we utilize an unsupervised learning approach based on SenticNet (Cambria et al., 2020). On the other hand, the posts of each candidate can politically influence (polarize) opinions about themselves or other candidates, depending on their content. We apply a supervised machine learning approach to identify these influences (political polarization).

The dataset analyzed consisted of tweets posted by presidential candidates between January 2018 and October 2018. We used capitalized acronyms to refer to each candidate's tweets: MEADE (José Antonio Meade Kuribreña), ANAYA (Ricardo Anaya Cortés), AMLO (Andrés Manuel López Obrador), ZAVALA (Margarita Ester Zavala Gómez del Campo), BRONCO (Jaime Heliodoro Rodríguez Calderón). The document uses these acronyms to refer to candidates' data and results. The data collection process involved downloading 7,031 tweets, focusing exclusively on content from presidential candidates to ensure relevance and specificity, and excluding trending topics and unrelated user tweets. All repeated tweets, those containing only internet links and those with fewer than ten characters, were removed. This filtering was carried out automatically using a Python script. The curation ultimately yielded 6,519 tweets, enhancing the dataset's quality and suitability for comprehensive analysis in the study of political polarization.

The polarity of each tweet (n tweets per candidate) was determined by applying SenticNet to each word within the tweets. SenticNet returns a continuous value between -1.0 and 1.0, where values nearing -1.0 suggest negative polarity, values approaching 1.0 indicate positive polarity and values near 0.0 are considered neutral (see Figure 1). The sum of positive and negative polarities is calculated to derive the total polarity of a tweet.

Figure 1. The polarity range calculated with SenticNet



The average overall positive ($\overline{p}_j^{+}$) and negative ($\overline{p}_j^{-}$) polarities of each candidate were calculated during the study period.

$\overline{p}_j^{+}=\frac{\sum_{i=1}^{n} p_{ij}^{+}}{n}$ and $\overline{p}_j^{-}=\frac{\sum_{i=1}^{n} p_{ij}^{-}}{n}$

Then, for each candidate, the absolute value between the average overall positive and average overall negative was calculated to measure the polarity generated by the Twitter posts during the political campaign (pol_j).

$$pol_j =$$

Finally, total polarity was calculated as the sum of all absolute values ($Totalpol$), to know the percentage proportion of each of the candidates $\%pol_j$.

$Totalpol = \sum_{j=1}^{5} pol_j$ and $\%pol_j = \frac{pol_j}{Totalpol}$

The sum goes from one to five, due to the number of candidates analysed.

The total polarity and the proportional percentage of each candidate summarise into a single number the polarities of all the candidate's posts during the campaign period.

To gauge the political polarization instigated by the candidates, we scrutinized the tweets shared by each of them. For every tweet, we determined whether the polarization was negative, positive, or neutral in relation to the same candidate and others. The manual labelling of tweets is a complex task requiring significant time investment. Defining clear rules for assigning labels is essential to achieve a high-quality dataset. The formulation of these rules was carefully considered in this research. To assign the polarity to each tweet, we stashed the following three criteria:

1) A tweet is considered positive for a presidential candidate (including the same author of the post) whether:
   a) It mentions at least a positive quality of someone.

   b) It gives merit to the ideas or proposals of someone.
   c) It is not sarcastic.
   d) It is not offensive in any way.
2) A tweet is considered negative for a presidential candidate if:
   a) It contains at least a negative criticism of someone (or his/her proposals).
   b) It mentions negative things about someone's past.
   c) It is offensive in some way.
3) If a tweet is neither positive nor negative, then it is considered neutral.
4) For each data set, about 10% (randomly chosen) of its tweets were labeled manually, as explained above.

Due to human factors, it is possible for two labelers not to agree on the assignment of labels for a given tweet. To avoid ambiguities, the decision was made to employ three labelers, so the majority determines the final label assigned. No cases of ties were detected. To assess the agreement (the degree to which two or more people agree on the same observed phenomenon) among the three people who labeled the tweets, we applied the procedure based on the Kappa Index designed by Cohen and proposed in Rangel, Sidorov, and Guerra (2014). This is as follows: the difference between the observed agreement rate and the expected agreement ratio is calculated by chance; if it is equal to zero, then the degree of agreement observed can be attributed entirely to chance; if the difference is positive, this indicates that the degree of agreement is greater than what would be expected. The estimation of the Kappa Index was 0.82, 0.81, and 0.81 for positive, negative, and neutral polarities, respectively. These values are considered a good agreement among the three people who labeled the tweets.

Pre-processing represents a prelude to extracting features for sentiment analysis. We implemented a combination of the methods of Rangel et al. (2014) and Paredes-Valverde et al. (2017) based on the following steps:

5) Convert to lowercase all the words.
6) Remove stopwords. These are words with no significant meaning.
7) Replace accented vowels with the corresponding unaccented vowels.
8) Replace URL, username, and emoji with special tags URL USERNAME and EMOTICON, respectively. URLs Begin with http://, and mentions to users begin with the symbol @. To identify emoticons in text, we use the table in https://unicode.org/emoji/charts/full-emoji-list.html.
9) Remove numbers, spaces, and punctuation symbols. It has been investigated that these factors do not contribute to the identification of sentiment within tweets.
10) Replace hashtags with the special tag HASHTAG. On Twitter, the # symbol is a way of marking something as belonging to a particular category. A category can contain one or more words. In the latter case, the first letter of each word is capitalized.
11) Get rid of HTML tags that remain after applying previous steps. These tags are between the symbols < and >.
12) Apply stemming. It consists of cutting the beginning or end of words, considering a list of common prefixes and suffixes that can be found in an inflected word.

Feature extraction is a fundamental step for generating high-performance predictive models. Although there are several text feature extraction techniques, such as the bag of words (BoW), word embeddings, and others. In this research, the term frequency-inverse document frequency (TF-IDF) technique was chosen. This is because BoW does not consider the frequency of terms in documents, which generally impairs classifier performance. On the other hand, word embeddings can generate different representations of words, leading to non-reproducible results. An intermediate representation is TF-IDF. This score reflects the importance of a term in a specific document relative to its frequency across all documents in the corpus.

We computed the TF-IDF based on the collected tweets. This statistic includes the frequency reversal coefficient, which measures the amount of information a word provides. Calculated using the following formula:

$$TF\text{-}IDF(t,d,D) = TF(t,d) \times IDF(t,D)$$

Where:

$TF(t,d)$ is the Term Frequency, representing the number of times term t appears in document d.

$IDF(t,D)$ is the Inverse Document Frequency, calculated as $\log (N/(DF(t,D)))$

where N is the total number of documents in the corpus and $DF(t,D)$ is the number of documents containing term t.

Through the application of TF-IDF, each tweet is represented by a sparse vector. Following the tagging, pre-processing, and vectorization of the tweets, we employed machine-learning techniques for sentiment analysis. According to the recent literature, the most common classification methods for sentiment analysis are Support Vector Machine (SVM), Bayesian Classifiers, Logistic Regression, and recently Deep Learning Techniques (Bhowmick et al., 2009; Karthik et al., 2018; Parveen & Pandey, 2016). These methods were employed for the analysis, excluding Deep Learning Techniques, which were not considered due to their substantial demand for labeled data.

The grid-search technique was used to select optimal parameter values for the classifiers, which involves performing a brute-force search over a range of values for each parameter.

For SVM, the best-found parameters were as follows: RBF kernel, gamma 0.01, C=10. For logistic regression, the best parameter value was C=2.5. For the decision tree, the split criterion was entropy, and the minimum number of samples per split was 2. The Naive Bayes classifier is parameter-free.

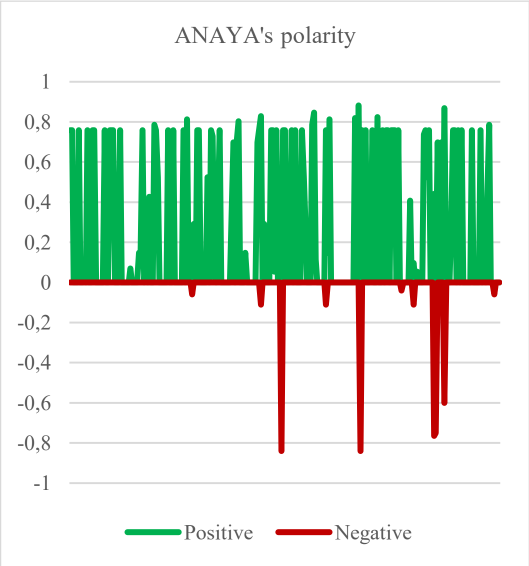
4. Results

This section presents the results of calculating Twitter posts' polarity of each candidate, the polarization generated in the political campaign, and the datasets' classification to understand the political polarization between candidates.

4.1. Towards a political polarization approach

To visualize and measure the polarity exercised by each candidate, we plotted daily polarity scores derived from Twitter posts. The results show the polarities of tweets posted by presidential candidates, with AMLO and BRONCO exhibiting large fluctuations between negative and positive polarities. MEADE and ANAYA exhibit predominantly positive polarity, while ZAVALA exhibits the lowest activity (Figure 2).

Figure 2. Posts' polarity per candidate



ANAYA emerged as the candidate with the highest average positive polarity throughout the campaign, followed by MEADE and ZVALA. AMLO ranked fourth in positive polarity, while BRONCO had the lowest average positive polarity. On the negative side, AMLO gave the most negative impression, followed by BRONCO, ZAVALA, MEADE, and ANAYA. A polarity approach that subtracts negative polarity from positive polarity was used to measure the impact of polarization in each candidate's publications. This allows us to understand the degree of polarization exhibited by each candidate, from positive to negative (Table 1).

Table 1. Percentage proportion of polarization per candidate.

Candidate	Average polarity		Absolute value (pol_j)	Polarization $\%pol_j$
	Positive ($\overline{p}_j^+$)	Negative ($\overline{p}_j^-$)		
BRONCO ($j = 1$)	0.1623	-0.0517	0.2140	17.2%
MEADE ($j = 2$)	0.1747	-0.0284	0.2032	16.3%
AMLO ($j = 3$)	0.1640	-0.2059	0.3698	29.7%
ZAVALA ($j = 4$)	0.1661	-0.0403	0.2064	16.6%
ANAYA ($j = 5$)	0.2329	-0.0171	0.2410	20.1%

Total	1.2434	100.0%

Source: Own elaboration

Regarding the political polarization during the political campaign, we found that the most polarising candidate on Twitter during the 2018 elections was AMLO (29.7%), in second place was ANAYA (20.1%), in third place BRONCO (17.2%), in fourth place ZAVALA (16.6%), and finally MEADE (16.3%).

4.2. Effect of polarization on candidates

We used four classification methods: Decision Tree, Naive Bayes, Logistic Regression, and Support Vector Machine to assess the effect of one candidate's polarization on other candidates. Table 2 shows the accuracy achieved by each classification method for predicting the sentiment of the presidential candidate dataset. Accuracy is a commonly employed metric to report the performance of classifiers in this type of study. This metric is used to measure a classification model's effectiveness or correctness. It represents the proportion of correctly classified instances out of the evaluated instances. The Decision Tree classifier was considered the best for the MEADE and ANAYA datasets, while Naive Bayes gave the best results for the BRONCO, AMLO, and ZAVALA datasets. These classifiers were chosen to facilitate data analysis and polarity prediction of both the candidate's emotions and those directed at others. The NA values displayed in Table 2 correspond to sets containing few documents of one class and many of the other classes. In other words, the few tweets belonging to a single taxonomy do not influence other candidates.

Table 2. Classification Accuracy for each dataset

DATASET	METHOD	BRONCO	MEADE	AMLO	ZAVALA	ANAYA
MEADE	Decision Tree*	100%	76%	100%	100%	100%
	Naive Bayes	100%	72%	100%	100%	100%
	Logistic Regression	NA	NA	100%	NA	NA
	Support Vector Machine	NA	72%	100%	NA	NA
BRONCO	Decision Tree	48%	100%	100%	100%	100%
	Naive Bayes*	60%	100%	100%	100%	100%

	Logistic Regression	56%	100%	100%	100%	100%
	Support Vector Machine	56%	100%	100%	100%	100%
AMLO	Decision Three	100%	96%	68%	100%	100%
	Naive Bayes*	100%	96%	88%	100%	100%
	Logistic Regression	NA	NA	88%	NA	NA
	Support Vector Machine	NA	NA	88%	NA	NA
ZAVALA	Decision Tree	100%	100%	100%	92%	100%
	Naive Bayes*	100%	100%	100%	92%	100%
	Logistic Regression	NA	NA	NA	NA	NA
	Support Vector Machine	NA	NA	100%	96%	NA
ANAYA	Decision Tree	92%	100%	96%	100%	64%
	Naive Bayes*	92%	100%	88%	100%	64%
	Logistic Regression	NA	NA	NA	NA	NA
	Support Vector Machine	92%	NA	88%	NA	68%

**Best Accuracy*

Source: own elaboration

Table 3 shows the effect of the polarization generated by each candidate on other candidates. We used the Kappa index to measure the polarization estimate for each candidate. The speech by the AMLO candidate on Twitter was highly influential and led to significant polarization. The candidate ZAVALA was the worst performer, as she had the lowest level of positive polarization estimation. All candidates, except for MEADE, performed similarly in negative polarization; MEADE had the lowest negative Kappa estimation. According to our results, the most neutral speech was ANAYA. However, he was the candidate who most often attacked AMLO.

Table 3. Effect of polarization on candidates.

Dataset and accuracy	Effect of polarization to	BRONCO	MEADE	AMLO	ZAVALA	ANAYA	Kappa Estimation for the Polarization of each candidate		
							Positive	Negative	Neutral
MEADE (100%)	Positive	0	222	1	0	0	0.76	0.63	0.71
	Negative	0	0	0	0	0			
	Neutral	261	39	260	261	261			
BRONCO (100%)	Positive	200	1	2	0	0	0.63	0.73	0.77
	Negative	0	0	0	0	0			
	Neutral	61	259	258	261	261			
AMLO (>96%)	Positive	0	1	95	0	0	0.80	0.79	0.65
	Negative	0	0	0	0	0			
	Neutral	261	260	165	261	261			
ZAVALA (100%)	Positive	0	0	0	12	0	0.52	0.75	0.63
	Negative	0	0	0	0	0			
	Neutral	261	261	261	248	261			

ANAYA (>92%)	Positive	0	0	0	0	238	0.64	0.76	0.80
	Negative	8	0	5	0	0			
	Neutral	252	261	256	261	22			

Source: own elaboration

We found that, except for one dataset (ANAYA), there were not many direct attacks between candidates on Twitter, despite the polarising effect during the political campaign. As expected, most of the candidate's tweets expressed themselves positively, but these tweets influenced how voters perceived other candidates. These reinforced or challenged voters' existing perceptions. About the candidates' Twitter posts, which explicitly targeted other candidates, we only found three positive messages from other candidates towards AMLO and two towards MEADE. We also found eight negative messages from ANAYA to BRONCO and five negative messages from ANAYA to AMLO. Analyzing the effect of polarization on candidates, we found that the neutrality of Twitter's discourse predominated, but their posts generated polarization among the voters.

5. Discussion

Polarization and echo chambers play a key role in shaping voter behavior and influencing election outcomes. Understanding the impact of polarization on voting behavior and outcomes is critical to understanding its broader implications for democratic processes. By examining the polarization phenomena, it could be possible to gain insight into the mechanisms by which polarization influences election outcomes and democratic decision-making. This knowledge could allow us to assess the potential impact of polarization on the functioning of democratic systems, including public debate, policymaking, and the impact on the overall state of democratic institutions. Moreover, examining the effects of polarization can provide strategies to mitigate its negative impacts and promote a more inclusive and constructive policy environment. By exploring the complexities of polarization and how it relates to voting behavior, researchers can help promote informed and active civic participation, strengthen democratic norms, and improve electoral strategies.

The manuscript underscores the significance of understanding polarization in politics. It reveals Twitter's influence on polarization and opinion formation, emphasizing the need for effective political strategies. In this regard, this research has identified three key lessons. Firstly, sentiment analysis can be broadened to encompass a wider range of emotions. Secondly, these emotions can offer insights into the political polarization observed in the Mexican Presidential elections, as our research has specifically highlighted the polarization effect among multiple candidates. Lastly, our study provides empirical evidence that social media platforms can be utilized to measure political polarization.

Research results continue to support the notion of a polarizing effect on candidates, as evidenced by the mixed negative and positive perceptions of AMLO among voters. Despite the polarization, AMLO emerged as the winning candidate in the presidential election. This is consistent with previous research showing that the most dichotomous candidates often win elections (Glowacki et al., 2018). The political polarization and echo chambers that arose during the election campaign resulted in fierce support within certain political factions, resulting in strong support for candidates.

AMLO strategically used political polarization as a central aspect of its electoral strategy. While he succeeded in securing the unconditional support of one segment of the population, he also faced determined opposition from another (Hernández-Alcántara, 2019). This polarizing dynamic extends to social media platforms as well, with supporters vigorously defending polarized political positions (McLaughlin et al., 2020; Pronin et al., 2002). In contrast, other candidates, such as ANAYA and BRONCO, were perceived as less important and elicited a more neutral response from voters. However, AMLO adherents showed higher levels of polarization and engagement within the echo chamber. AMLO's divisive election editorials likely affected voters who were undecided or lacked a firm ground. This influence may have contributed to his victory.

Artificial intelligence techniques, such as machine learning and sentiment analysis, these techniques can be used to analyze candidates' phrasing and determine their overall political leanings. This analytical approach could help to understand voter responses by determining whether candidates are promoting polarization or seeking voter unity. However, ethical considerations of privacy, bias, and transparency are paramount when using these technologies. Establishing ethical guidelines for using machine learning in policy analysis and ensuring adherence to them is critical, especially in the context of emerging technologies, such as generative artificial intelligence. Besides, machine learning techniques can help predict, explain, and model physical and social phenomena. However, it is crucial to consider that these models may not capture the entirety of reality, as they do not account for other missing aspects in the data used for their calibration.

It could be of utmost importance to recognize and understand the profound impact that polarization has on voter decisions, democratic processes, and the potential consequences that arise from them. Equally important is the careful consideration required when integrating generative artificial intelligence and applications like ChatGPT into the political arena. While these technological advances offer promising opportunities to improve voter communications, refine political marketing strategies, and improve election proposals, they also come with risks. These risks include the potential to exacerbate polarization and facilitate the spread of misinformation. A thorough understanding of these dynamics is, therefore, required to achieve a harmonious relationship between using artificial intelligence for political analysis and ensuring democratic processes. The impact of polarization on voting decisions and democratic processes cannot be underestimated. As the polarization within society increases, voters tend to align themselves along ideological lines, facilitating constructive dialogue and making compromises difficult to reach. This phenomenon creates an echo chamber where another perspective is ignored, thwarting the democratic ideal of making informed decisions from multiple perspectives. Moreover, polarization fosters an “us vs. them” mentality, which hinders political institutions' functioning and effective governance.

Still, it is important to recognize the risks associated with embedding AI into the political arena. One such risk is the potential for polarization amplification. AI algorithms can inadvertently reinforce existing prejudices and preferences and unintentionally perpetuate echo chambers. Moreover, the rapid spread of misinformation via AI-powered platforms can undermine public confidence in democratic processes, jeopardize electoral fairness, and distort political debate. Achieving a healthy democracy requires a balance between reaping the benefits of AI and maintaining the integrity of the democratic process. Policymakers, technologists, and society at large must prioritize transparency, accountability, and ethical standards in artificial intelligence development and deployment. Implementing rigorous fact-checking mechanisms, creating regulatory frameworks, and promoting interdisciplinary collaboration will help mitigate the risks associated with AI-generated content. In addition, emphasis should be placed on promoting the public's digital literacy and critical thinking skills to enable them to navigate complex information environments and make informed decisions effectively.

Drawing from Mexico's experience, global politicians can learn diverse political strategies to influence public opinion more effectively. Positive strategies emphasizing constructive proposals and solutions can strengthen the emotional connection with citizens. Avoiding negative political tactics that foster polarization and confrontation and instead adopting a collaborative and respectful approach may contribute to a healthier public debate. However, it may not necessarily benefit them in electoral outcomes. Regarding neutral strategies, presenting information objectively and equitably promoting transparency and impartiality can build public trust by providing a more balanced view of issues and solutions. However, every country has unique political conditions, and every political campaign has a distinct nature. It is difficult to generalize about political strategies in every case because this research does not specifically focus on this topic. Regardless, this research provides strong evidence that social media plays a significant role in promoting polarization. Combining this finding with the emerging field of generative artificial intelligence in political campaigns can create a powerful combination that can potentially influence campaigns and change how voters behave. This research serves as an initial warning about the potential consequences of this combination.

6. Conclusions, limitations, and future work

This study explored the political polarization surrounding the 2018 Mexican presidential election by analyzing Twitter messages. The election showed significant political polarization on Twitter, with the winning candidate, AMLO, exhibiting the highest level of polarization. The findings highlight the influential role that social media platforms, especially Twitter, play in fostering and amplifying political divisions and could be useful in understanding upcoming presidential elections. However, while direct attacks between candidates on Twitter were rare, the impact that candidates' tweets had on voters' perceptions of their rivals was striking. These tweets could amplify or challenge existing perceptions, highlighting the potential impact of social media on the voting decisions of its users. Despite the prevalence of neutral narratives on Twitter, polarization still has a clear impact on voter divisions.

The winning candidate could strategically use political polarization as a central element of its electoral strategy, successfully garnering firm support from one segment of the population while facing opposition from another. This shows that polarization can be a powerful tool for electoral success. The effects of polarization go beyond voter decisions and have far-reaching implications for the functioning of democratic processes and political institutions. Increasing polarization hinders effective governance by making constructive dialogue, compromise, and consideration of different perspectives difficult.

Incorporating AI into the world of politics carries inherent risks, such as increased polarization and the spread of misinformation. Maintaining democratic principles requires introducing transparency, accountability, and robust fact-checking mechanisms. Additionally, improving the population's digital literacy is critical to enabling individuals to navigate complex information environments and make informed decisions. Finding a balance between reaping the benefits of AI and maintaining democratic values is critical to fostering a healthy democracy. Promoting ethical AI development, creating regulatory frameworks, fostering interdisciplinary collaboration, and prioritizing digital skills are key to addressing AI's complex dynamics and risks in politics.

The results of this research contribute to advances in sentiment analysis, machine learning, and political campaign research in several ways. First, we introduced a new method adapted to the Mexican context, utilizing programming languages, such as Python and R, to classify and analyze the data using SenticNet and machine learning techniques. Additionally, we conducted individual candidate analysis, identifying tweets that mentioned or referenced other candidates, thereby enhancing the accuracy of sentiment analysis. Furthermore, the study demonstrates that social media discourse can be examined to analyze the impact of political polarization.

Moreover, the research's focus on content and link analysis distinguishes it from other studies that primarily concentrate on predicting message volume, retweets, network building, and campaign outcomes. One significant contribution of this research is tied to the methodology employed for analyzing polarity on Twitter, aiming to comprehend political polarization on social media. Simultaneously, the study furnishes evidence indicating that the most polarizing candidate stands a higher chance of winning the election. By delving into the behaviors of politicians on social media platforms, this study offers empirical evidence of political polarization within the realm of political contestation.

Furthermore, our research underscores the pivotal role of social media in political contestation, emphasizing the necessity of evaluating its impact on shaping political narratives and mobilizing support. As these platforms continue to gain popularity and influence, a comprehensive assessment of their role becomes crucial. The empirical evidence provided by our findings advocates for further exploration in this domain.

The identified limitations in this study prompt important considerations for future research endeavors. This research involved the analysis of Twitter data. While Twitter is not the exclusive factor in political polarization, its widespread use by politicians highlights its significance in reflecting diverse perspectives. Specifically, there is a need to delve into various dimensions of bias, encompassing demographic, ideological, and cultural factors. A thorough exploration of these dimensions is essential to understand their potential impact on the outcomes of models designed for detecting political polarization. This avenue of investigation will contribute to a more comprehensive and nuanced understanding of the challenges associated with bias in polarization detection models.

References

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? The Journal of Politics, 70(2), 542–555.

Bakker, T. P., & De Vreese, C. H. (2011). Good news for the future? Young people, Internet use, and political participation. Communication Research, 38(4), 451–470.

Balomenos, T., Raouzaiou, A., Ioannou, S., Drosopoulos, A., Karpouzis, K., & Kollias, S. (2004). Emotion analysis in man-machine interaction systems. International Workshop on Machine Learning for Multimodal Interaction, 318–328.

Barberá, P. (2015). How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the. Job Market Paper, New York University, 46. http://rubenson.org/wp-content/uploads/2015/10/barbera-tpbw.pdf

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? Psychological Science, 26(10), 1531–1542.

Batorski, D., & Grzywińska, I. (2018). Three dimensions of the public sphere on Facebook. Information, Communication & Society, 21(3), 356–374.

Bhowmick, P. K., Basu, A., & Mitra, P. (2009). Reader Perspective Emotion Analysis in Text through Ensemble based Multi-Label Classification Framework. Computer and Information Science, 2(4), 64–74.

Boulianne, S. (2009). Does Internet use affect engagement? A meta-analysis of research. Political Communication, 26(2), 193–211.

Boulianne, S., Koc-Michalska, K., & Bimber, B. (2020). Right-wing populism, social media and echo chambers in Western democracies. New Media & Society, 22(4), 683–699.

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. Political Psychology, 38(3), 551–569.

Brito, K. D. S., Silva Filho, R. L. C., & Adeodato, P. J. L. (2021). A systematic review of predicting elections based on social media data: research challenges and future directions. IEEE Transactions on Computational Social Systems, 8(4), 819–843.

Bruns, A. (2017). Echo chamber? What echo chamber? Reviewing the evidence. 6th Biennial Future of Journalism Conference (FOJ17), 14.

Cambria, E., Li, Y., Xing, F. Z., Poria, S., & Kwok, K. (2020). SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis. Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 105–114.

Chaffar, S., & Inkpen, D. (2011). Using a heterogeneous dataset for emotion analysis in text. Canadian Conference on Artificial Intelligence, 62–67.

Choy, M. J., Cheong, M. L. F., Ma, N. L., & Koo, P. S. (2012). A Sentiment Analysis of Singapore Presidential Election 2011 using Twitter Data with Census Correction.

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2020). Echo chambers on social media: A comparative analysis. ArXiv Preprint ArXiv:2004.09603.

Clark, S. D., Morris, M. A., & Lomax, N. (2018). Estimating the outcome of UKs referendum on EU membership using e-petition data and machine learning algorithms. Journal of Information Technology & Politics, 15(4), 344–357.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. Journal of Communication, 64(2), 317–332.

Conitzer, V., & Sandholm, T. (2002). Complexity of manipulating elections with few candidates. AAAI/IAAI, 314–319.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. Scientific Reports, 6(1), 1–12.

Dubois, E., Minaeian, S., Paquet-Labelle, A., & Beaudry, S. (2020). Who to trust on social media: How opinion leaders and seekers avoid disinformation and echo chambers. Social Media+ Society, 6(2), 2056305120913993.

Elkind, E., Faliszewski, P., Skowron, P., & Slinko, A. (2017). Properties of multiwinner voting rules. Social Choice and Welfare, 48(3), 599–632.

Faliszewski, P., Slinko, A., Stahl, K., & Talmon, N. (2018). Achieving fully proportional representation by clustering voters. Journal of Heuristics, 24(5), 725–756.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. Public Opinion Quarterly, 80(S1), 298–320.

Fujiwara, T., Müller, K., & Schwarz, C. (2021). The effect of social media on elections: Evidence from the United States.

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018). Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. Proceedings of the 2018 World Wide Web Conference, 913–922.

Garimella, V., & Weber, I. (2017). A long-term analysis of polarization on Twitter. Proceedings of the International AAAI Conference on Web and Social Media, 11(1).

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. Journal of Computer-Mediated Communication, 14(2), 265–285.

Genesereth, M. (2018). AI Amusements: Computer Elected Governor of California Corpus Legis trounces human opponents in state election. AI Matters, 4(1), 32.

Glowacki, M., Narayanan, V., Maynard, S., Hirsch, G., Kollanyi, B., Neudert, L.-M., Howard, P., Lederer, T., & Barash, V. (2018). News and political information consumption in Mexico: Mapping the 2018 Mexican presidential election on Twitter and Facebook. The Computational Propaganda Project.

Han, B.-C. (2022). Infocracia. La digitalización y la crisis de la democracia. Taurus.

Hernández-Alcántara, C. (2019). Campañas electorales presidenciales pragmáticas en México 2018. Política y comunicación. Revista Mexicana de Ciencias Políticas y Sociales, 64(235), 327–352.

Hetherington, M. J. (2001). Resurgent mass partisanship: The role of elite polarization. American Political Science Review, 619–631.

Ikiz, O. O., Sobaci, M. Z., Yavuz, N., & Karkin, N. (2014). Political use of Twitter: The case of metropolitan mayor candidates in 2014 local elections in Turkey. Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance, 41–50.

Jaidka, K., & Ahmed, S. (2015). The 2014 Indian general election on Twitter: an analysis of changing political traditions. Proceedings of the Seventh International Conference on Information and Communication Technologies and Development, 1–5.

Justwan, F., Baumgaertner, B., Carlisle, J. E., Clark, A. K., & Clark, M. (2018). Social media echo chambers and satisfaction with democracy among Democrats and Republicans in the aftermath of the 2016 US elections. Journal of Elections, Public Opinion and Parties, 28(4), 424–442.

Karthik, V., Nair, D., & Anuradha, J. (2018). Opinion mining on emojis using deep learning techniques. Procedia Computer Science, 132, 167–173.

Koolagudi, S. G., Maity, S., Kumar, V. A., Chakrabarti, S., & Rao, K. S. (2009). IITKGP-SESC: speech database for emotion analysis. International Conference on Contemporary Computing, 485–492.

Lindqvist, E., & Östling, R. (2008). Political Polarization and the Size of Government (Issue 749). Research Institute of Industrial Economics (IFN). http://hdl.handle.net/10419/81331

López-Chau, A., Valle-Cruz, D., & Sandoval-Almazán, R. (2020). Sentiment Analysis of Twitter Data Through Machine Learning Techniques. In Software Engineering in the Era of Cloud Computing (pp. 185–209). Springer.

Luzsa, R., & Mayr, S. (2021). False consensus in the echo chamber: Exposure to favorably biased social media news feeds leads to increased perception of public support for own opinions. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 15(1).

McLaughlin, B., Holland, D., Thompson, B. A., & Koenig, A. (2020). Emotions and Affective Polarization: How Enthusiasm and Anxiety About Presidential Candidates Affect Interparty Attitudes. American Politics Research, 48(2), 308–316.

McLeod, J. M., Scheufele, D. A., & Moy, P. (1999). Community, communication, and participation: The role of mass media and interpersonal discussion in local political participation. Political Communication, 16(3), 315–336.

Michael, G., & Agur, C. (2018). The bully pulpit, social media, and public opinion: A big data approach. Journal of Information Technology & Politics, 15(3), 262–277.

Morales, A. J., Borondo, J., Losada, J. C., & Benito, R. M. (2015). Measuring Political Polarization: Twitter shows the two sides of Venezuela. Chaos: An Interdisciplinary Journal of Nonlinear Science, 25(3), 33114. https://arxiv.org/pdf/1505.04095.pdf;

Morozov, E. (2018). Big Data. Ubu Editora LTDA-ME.

Nguyen, C. T. (2020). Echo chambers and epistemic bubbles.

Nulty, P., Theocharis, Y., Popa, S. A., Parnet, O., & Benoit, K. (2016). Social media and political communication in the 2014 elections to the European Parliament. Electoral Studies, 44, 429–444.

Paredes-Valverde, M. A., Colomo-Palacios, R., Salas-Zárate, M. del P., & Valencia-García, R. (2017). Sentiment analysis in Spanish for improvement of products and services: a deep learning approach. Scientific Programming, 2017.

Parveen, H., & Pandey, S. (2016). Sentiment analysis on Twitter Data-set using Naive Bayes algorithm. 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (ICATCCT), 416–419.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. Personality and Social Psychology Bulletin, 28(3), 369–381.

Quan, C., & Ren, F. (2010). Sentence emotion analysis and recognition based on emotion words using Ren-CECps. International Journal of Advanced Intelligence, 2(1), 105–117.

Rangel, I. D., Sidorov, G., & Guerra, S. S. (2014). Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. Onomazein, 29, 31–46.

Recuero, R., Soares, F. B., & Gruzd, A. (2020). Hyperpartisanship, disinformation and political conversations on Twitter: The Brazilian presidential election of 2018. Proceedings of the International AAAI Conference on Web and Social Media, 14, 569–578.

Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In Advances in experimental social psychology (Vol. 27, pp. 255–304). Elsevier.

Sandoval-Almazan, R., & Valle-Cruz, D. (2020). Sentiment Analysis of Facebook Users Reacting to Political Campaign Posts. Digital Government: Research and Practice, 1(2), 1–13.

Santander, P., Elórtegui, C., González, C., Allende-Cid, H., & Palma, W. (2017). Redes sociales, inteligencia computacional y predicción electoral: el caso de las primarias presidenciales de Chile 2017. Cuadernos. Info, 41, 41–56.

Singh, P., & Sawhney, R. S. (2018). Influence of twitter on prediction of election results. In Progress in Advanced Computing and Intelligent Engineering (pp. 665–673). Springer.

Srivastava, R., & Bhatia, M. P. S. (2017). Real-time unspecified major sub-events detection in the twitter data stream that cause the change in the sentiment score of the targeted event. International Journal of Information Technology and Web Engineering (IJITWE), 12(4), 1–21.

Srivastava, R., Kumar, H., Bhatia, M. P. S., & Jain, S. (2015). Analyzing Delhi assembly election 2015 using textual content of social network. Proceedings of the Sixth International Conference on Computer and Communication Technology 2015, 78–85.

Staiano, J., & Guerini, M. (2014). Depechemood: a lexicon for emotion analysis from crowd-annotated news. ArXiv Preprint ArXiv:1405.1605.

Takikawa, H., & Nagayoshi, K. (2017). Political polarization in social media: Analysis of the "Twitter political field" in Japan. 2017 IEEE International Conference on Big Data (Big Data), 3143–3150.

Terren, L., & Borge, R. (2021). Echo Chambers on Social Media: A Systematic Review of the Literature. Review of Communication Research, 9, 1–39. https://doi.org/10.12840/ISSN.2255-4165.028

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. Political Polarization, and Political Disinformation: A Review of the Scientific Literature (March 19, 2018).

Valle-Cruz, D., Lopez-Chau, A., & Sandoval-Almazan, R. (2021). How much do Twitter posts affect voters? Analysis of the multi-emotional charge with affective computing in political campaigns. DG. O2021: The 22nd Annual International Conference on Digital Government Research, 1–14.

Valle-Cruz, D., López-Chau, A., & Sandoval-Almazán, R. (2020). Impression analysis of trending topics in Twitter with classification algorithms. Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, 430–441.

Valle-Cruz, D., Sandoval-Almazán, R., & López-Chau, A. (2023). Affective Polarization in the US: Multi-Emotional Charge Analyzed Through Affective Computing. In Handbook of Research on Artificial Intelligence Applications in Literary Works and Social Media (pp. 192–219). IGI Global.

Verba, S., & Nie, N. H. (1987). Participation in America: Political democracy and social equality. University of Chicago Press.

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Gao, S., Sun, Y., Ge, W., Zhang, W., & others. (2022). A systematic review on affective computing: Emotion models, databases, and recent advances. Information Fusion, 83, 19–52.

Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. Perspectives on Psychological Science, 10(2), 145–158.

About the Authors

David Valle-Cruz

David Valle-Cruz (PhD) is a professor at the Tianguistenco Professional Academic Unit at the Autonomous University of the State of Mexico, a member of the Mexican National Researchers Level 1, and the Innovation and Artificial Intelligence Laboratory: i-Lab. His studies encompass a Bachelor's degree in Computer Engineering, a Master's in Computer Science, and a PhD in Economic-Administrative Sciences. David has undertaken research stays at the Center for Technology in Government (CTG), SUNY Albany, NY, and at the Computer Science and Multi-Agent Systems Laboratory of CINVESTAV, Guadalajara, Mexico. He is part of the editorial board of the Government Information Quarterly journal and editor of the "Handbook of Research on Applied Artificial Intelligence and Robotics for Government Processes" published by IGI Global. His worldwide impact activities include being a National Expert for the National Science Foundation's Digital Society project and an evaluator of Digital Government for the United Nations. The Digital Government Society has recognized David's academic work for excellence and innovation. His scientific publications are found in leading journals like Government Information Quarterly, Cognitive Computation, International Review of

Administrative Sciences, Public Policy and Administration, First Monday, Information Polity, and the International Journal of Public Sector Management, among others. David is an author and co-author of scientific chapters published by Oxford University Press, Springer, and IGI Global. His research interests focus on applying AI and data science for strategic decision-making.

Rodrigo Sandoval-Almazán

Dr. Rodrigo Sandoval-Almazan, Ph.D., is an Associate Professor at the Political Sciences and Social Sciences Department at the Autonomous University of the State of Mexico. He serves as the Director of the Artificial Intelligence and Innovation Lab for Mexico (i-Lab Mexico) and is a research fellow at the Center for Technology in Government, State University of New York (SUNY at Albany), as well as at the Lab Research Group Innovation, Technology, and Public Management, Universidad Autónoma de Madrid, Spain. He is also an Associate Editor of the Government Information Quarterly (GIQ). Dr. Sandoval-Almazan's research focuses on artificial intelligence in the public sector, open government, open data, transparency policies, digital government, social media metrics, and e-bureaucracy. He has published articles in top scholarly journals such as Government Information Quarterly, Information Polity, International Journal of Public Sector Management, and First Monday. Dr. Sandoval-Almazan was elected as a member of the Mexican Academy of Science in 2018 and received the SCOPUS prize in 2023 for being the most cited scholar of social sciences at the Autonomous University of the State of Mexico. He is currently a Member of the Mexican National Council of Social Sciences CONACYT, Level 2, since 2016.

Asdrúbal López-Chau

Asdrúbal López Chau (PhD) is a Communications and Electronics Engineer (1997) from the Higher School of Mechanical and Electrical Engineering at National Polytechnic Institute in Mexico (ESIME-IPN); Master of Science in Computer Engineering (2000) from the Computer Research Center of the IPN (CIC-IPN) in Mexico; and Doctor in Computer Science (2013) from the Center for Research and Advanced Studies of the IPN (CINVESTAV-IPN) in Mexico. His current research lines are applied computing, sentiment analysis, machine learning, and image processing. Dr. Asdrúbal is the author of 1 book about microcontrollers, He has participated in 17 articles in journals indexed in the Journal Citation Report (JCR); also, he has published more than 18 book chapters of the Springer Verlag publishing house in Germany and has participated as author of articles in more than a dozen international conferences. He has directed more than 30 undergraduate and postgraduate degree projects. He is currently a reviewer of articles for magazines indexed in JCR. He is an organizer and collaborator of national and international conferences. Since 2011, Dr. Asdrúbal has been a full-time research professor at the Autonomous University of the State of Mexico (UAEM) in Zumpango, where he has led six basic science research projects. He currently collaborates in the basic core of the Master's Degree in Computer Science at the CU UAEM Valle de México. In 2013, he was distinguished with the first place for the best doctoral thesis in the area of Artificial Intelligence by the Mexican Society of Artificial Intelligence; in 2014, he was awarded the prize for carrying out exceptional basic research work, granted by the UAEM through Clause 89. In October 2014, he won first place in the XXVII national doctoral thesis competition on computing and informatics organized by the National Association of Informatics Education Institutions. Since January 2014, He has been an uninterrupted member of the Member of the Mexican National Council of Social Sciences CONACYT, Level 1. Dr. Asdrúbal López Chau is part of the Laboratory of Applied Computing Technologies at CU UAEM Zumpango. More information on his web page: www.al-chau.com

J. Ignacio Criado

J. Ignacio Criado (PhD) is a senior lecturer / associate professor in Political Science and Public Administration at the Department of Political Science and International Relations, director of the Lab Research Group Innovation, Technology and Public Management, Universidad Autónoma de Madrid, Spain, and research fellow at Center for Technology in Government, State University of New York (SUNY at Albany). Also, he has been visiting fellow at Oxford Internet Institute, the University of Oxford, visiting researcher in different international institutions, including Monash University, the University of Manchester, the London School of Economics, or the European Institute of Public Administration, and visiting professor at Digital and Mobile Governance Lab, SIRPA, Fudan University (Shanghai), Utrecht School of Governance (Netherlands), El Colegio de México and CIDE (Mexico), among others. His research interests include algorithmic governance and artificial intelligence in the public sector; open government and policies for transparency, participation, and public innovation; and digital government, social media, and big data in public administration. His articles have been published in peer-reviewed leading journals, including Government Information Quarterly, Social Science Computer Review, Information Polity, International Journal of Public Sector Management, Public Policy and Administration, Local Government Studies, International Journal of Public Administration, International Journal of Electronic Government Research, International Journal of Electronic Governance, Transforming Government, or First Monday. He serves as an editorial board member for Government Information Quarterly, International Journal of Public Sector Management, or International Journal of Public Administration in the Digital Age, and reviews for many others. He has collaborated in training courses and consultancy projects for several local, regional, national, and international organizations and governments. He has evaluated for the European Awards for e-Government of the European Commission, the Independent Report Mechanism, Open Government Partnership, or national research and foresight agencies (ANEP, Spain, CONACYT, Mexico, NWO, Netherlands, etc.), and co-authored the Ibero-American Interoperability Framework for CLAD. He co-founded and worked for a University spin off for five years and directed the first Spanish University Gov Lab. He has served at Universidad Autónoma de Madrid as academic secretary (Department of Political Science and I.R.) (2008-2013), Vice-dean of research and innovation (2014-2018) (School of Law and Political Science), and Deputy Vice-Rector for Open Innovation and Data (2020-2021). He is an elected Board Member of the Digital Government Society (2020-) and the Spanish Government Open Government Forum (2018).

