UNIVERSIDAD NACIONAL DE SAN MARTIN

RCSI

Sentiment analysis in social network X for the evaluation of the positioning of candidates in political elections

Análisis de sentimientos en la red social X para la evaluación del posicionamiento de candidatos en elecciones políticas

Denicia-Carral, María Claudia[1*]
Ballinas-Hernández, Ana Luisa[1]
Minquiz-Xolo, Gustavo Manuel[1]
Medina-Cruz, Héctor[1]

[1]Benemérita Universidad Autónoma de Puebla. Puebla, México

ABSTRACT

Social networks are among the most important means of political communication, through which opinions are published and generated on various topics; for this reason, they are an excellent means to analyze and understand events. In this work, sentiment analysis of X/Twitter posts about the elections of presidential candidates in Mexico in 2024 was carried out via the classification of sentiment polarity to measure the positioning of the participants. A methodology based on KDD was used in this research, and 151,821 posts were analyzed about four aspiring candidates for the presidency of Mexico from the MORENA party. The results showed that the candidates best positioned in the election obtained the highest number of posts with positive polarity, although the winning candidate did not coincide with the highest percentage of positive polarities. This result indicates that it is necessary to include other variables in addition to polarity to more accurately predict the winners of political contests.

Keywords: KDD Methodology; Data Mining; NLP; Sentiment Polarity

RESUMEN

Las redes sociales son uno de los medios de comunicación política más importantes, a través de ellas se publican y generan opiniones de una gran variedad de temas, por esta razón, son un medio excelente para realizar análisis y comprender sucesos. En este trabajo se realizó un análisis de sentimientos de publicaciones de X/Twitter sobre las elecciones de candidatos presidenciales de México en 2024 utilizando la clasificación de polaridad de sentimientos para medir el posicionamiento de los participantes. Se utilizó una metodología basada en KDD y se analizaron 151 821 publicaciones sobre cuatro aspirantes a candidatos a la presidencia de México del partido MORENA. Los resultados mostraron que los candidatos mejor posicionados en la elección son aquellos que obtuvieron la mayor cantidad de publicaciones con polaridad positiva, aunque el aspirante ganador no coincidió con el mayor porcentaje de polaridades positivas. Este resultado indica que es necesario incluir otras variables además de la polaridad para hacer una predicción más exacta de los ganadores de las contiendas políticas.

Palabras clave: Metodología KDD; Minería de Datos; PLN; Polaridad de Sentimientos

1. INTRODUCTION

In Mexico, presidential elections are determinant because the president-elect will have a significant effect on the direction of the country, its public policies, the fight against corruption, security, the economy, human rights, and the relationship with other nations. Presidential elections are held every 6 years, and different political parties carry out internal candidate selection processes. One of the most representative is MORENA, which governs at the national level for the first time in history, so continuity in the presidency is highly relevant for his permanence and leadership.

For the selection of the presidential candidate, the political parties in Mexico carry out an internal selection that involves a political campaign process culminating in a consultation among the militants to select their representative (Espejel Espinoza & Díaz Sandoval, 2022). The campaigns are carried out in person and digitally, with social networks being the greatest means of communication with citizens and how opinions and feelings are expressed toward the proposals.

Sentiment analysis involves extracting and analyzing opinions, sentiments, attitudes, and perceptions through text (Birjali et al., 2021). It is one of the techniques for making decisions in different areas; for example, in the investigation of Garg et al. (2020) web series preferences are determined through the analysis of Twitter comments and the prediction characteristics of the most popular series. Other authors have focused on the application of sentiment analysis in the medical field, such as those of Bharany et al. (2023); and Turón et al. (2023) who measured the impact of COVID-19 through posts from X, formerly known as Twitter (Antonakaki et al., 2021).

In the political field, sentiment analysis has increased interest in the scientific community. Oliveira et al. (2019) analyzed opinions expressed on Twitter regarding social programs in Brazil, while Ansari et al. (2020) developed a sentiment-based classification model to predict election results in India. Other authors who used sentiment analysis in the political field were Caetano et al. (2018), Oueslati et al. (2023), Sobkowicz & Kozłowski (2018), who concluded that sentiment analysis was used successfully to support management practices in the political environment.

Some political studies involve other variables. Ongo Nkoa et al. (2023) the impact of women in politics has been analyzed using social networks. (Russo et al., 2024)studied how social networks impact the construction of theories that generate political disturbances, especially in the case of Brazil in 2023. Flamino et al. (2023) observed the impact of influencers and the political bias that they can generate through social networks.

Social networks are the larger virtual spaces for giving and receiving opinions on any subject. In politics, the social network X/Twitter is a social network of preferences. In this regard, Gilardi et al. (2022) and Marín Dueñas et al. (2019) social networks have become the main means of political communication because, through social networks, opinions are expressed by both the main actors and the readers, who have become judges and commentators on the political scene.

Social networks have been explored by some authors; for example, in the study conducted by Chaudhry et al. (2021), research on the 2020 elections in the U.S. was conducted by analyzing sentiment on Twitter before, during, and after the elections. Some results coincided with those obtained in the election, but the positive sentiment was the predominant sentiment in the winning candidate. On the other hand, Noor et al. (2024) analyzed Canadian elections in 2021 and concluded that high positivity in posts does not always reflect electoral results.

Sarapugdi & Namkhun (2023) reported that sentiments expressed on X/Twitter may be biased on some occasions; however, they are excellent guides in the development of analysis research and prediction models if complemented with data from other social networks. The opinion bias may be caused by different factors. Cantini et al. (2022) showed that the use of boots on X/Twitter is one of the reasons why opinions can be altered to favor a candidate.

The year 2024 marked a turning point in Mexico's political landscape, as the country underwent a presidential transition that significantly impacted its governance direction; one of the most important political parties was the National Regeneration Movement (MORENA) (Lucca, 2019). This party conducts an internal presidential candidate selection process among the four candidates. Applicants widely use social networks to publicize their proposals, achievements, and events. One of the favorite networks in Mexico for the political sphere is X/Twitter, which is why applicants make several efforts to analyze what happens on it.

The objective of this work is to carry out a sentiment analysis through X/Twitter posts referring to each of the candidates for the presidential candidacy of the MORENA party to determine and analyze their popularity and acceptance among the inhabitants of Mexico through the determination of the polarity of posts made by users.

2. MATERIALS AND METHODS

This study gathered feedback from X/Twitter regarding the campaigns of MORENA party presidential candidacy contenders. A quantitative approach was employed, utilizing a sentiment analysis system to collect and examine numerical data associated with tweets. The level of research used in this study was descriptive since it sought to describe the public perception and acceptance of each applicant in cyberspace through sentiment analysis of the posts related to the applicants. The study employed a non-experimental approach as the researchers did not manipulate any variables and observed the phenomenon in its natural setting.

The population was the users of the social network X/Twitter in Mexico who published tweets related to the four presidential candidates of the MORENA party during the electoral campaign of the elections for presidential candidates in 2024. The research sample consisted of 151 821 posts (before tweets) collected from users of the social network in Mexico who published information related to the four candidates of the MORENA party during the electoral campaign. The four candidates for the candidacy were Marcelo Ebrard, Ricardo Monreal, Claudia Sheinbaum, and Adán Augusto López.

A post on X/Twitter is composed of different elements, including the user who wrote the post, the date and time of the post, the text (maximum 280 characters), which can include mentions to other users (@ is used to make mentions), hashtags that are tags to mark keywords or topics (# for hashtags), and the addition of multimedia content such as photos or videos. It is also possible to identify from where it was written, for example, a mobile device or a web page, how many times it was shared, among other things.

This work was carried out via data mining tools (Vyas & Uma, 2018); specifically, RapidMiner Studio was used in version 9.7 for post-extraction according to search criteria. The MeaningCloud extension for RapidMiner was used for text processing and polarization with sentiment analysis, that is, the detection of positive, negative, or neutral sentiment concerning a text, in addition to irony sentiment detection. The texts in an X/Twitter post are short and composed of a limited number of sentences; therefore, a local polarity identification process was carried out in each sentence to determine the global polarity of each post later.

The process developed is shown in Figure 1 and consists of 5 stages inspired by the KDD knowledge discovery methodology (Debuse et al., 2001). This begins with a data collection stage, followed by a preprocessing model that homogenizes the data, cleans it, and transforms it to be able to apply classification algorithms. Finally, the results obtained are analyzed to conclude the patterns discovered.
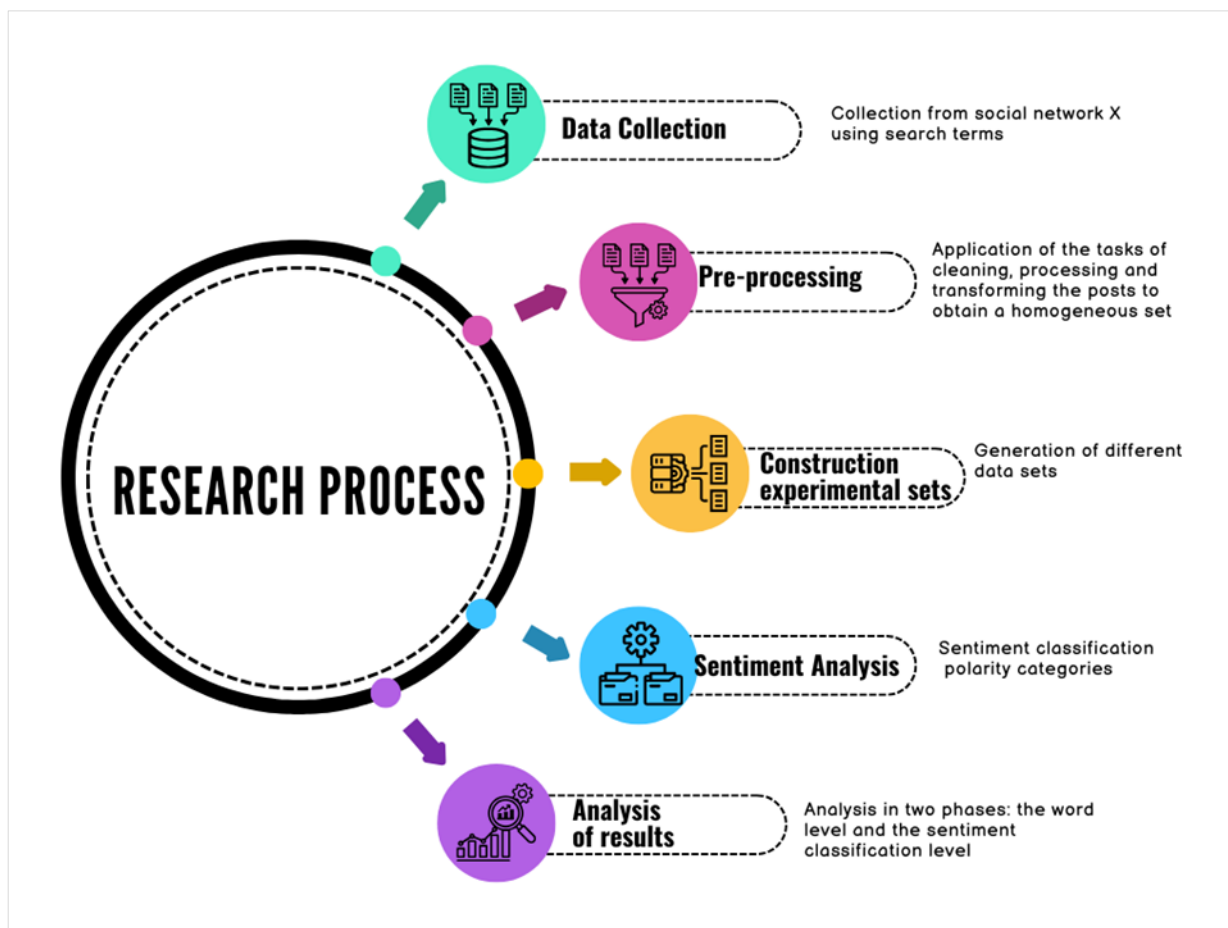
Figure 1. Stages of the research process

Below is a description of each of the research stages:

2.1. Data collection

The collection was carried out during the campaigns of the applicants and a period of one week, interspersing the days of the collection as follows: two days at the beginning of the campaign, three days during the campaign, and two days before the end of the campaign. The campaign period was from June 19, 2023, to September 5, 2023. The collection dates were the following: June 20 and 25, 2023; August 14, 18, and 20, 2023; and September 4 and 5, 2023. The search terms for each candidate are shown in Table 1. The names of the applicants and some variations and combinations were used, as were the hashtags used in the social network and the official accounts of each applicant.

Table 1. Search terms for each applicant

Applicant	Search terms
Marcelo Ebrard	"Marcelo Ebrard", "Marcelo Luis Ebrard Casaubón", "candidato Marcelo Ebrard", la combinación de los términos "Marcelo Ebrard" y "MORENA", #Marcelo_Ebrard, @m_Ebrard
Claudia Sheinbaum	"Claudia Sheinbaum", "Claudia Sheinbaum Pardo", "Dra. Claudia Sheinbaum", "candidata Claudia Sheinbaum", la combinación "Claudia Sheinbaum" y "MORENA", #ClaudiaSheinbaum, @Claudiashein
Adán Augusto López	"Adán Augusto López", "Adán Augusto López Hernández", "Adán Augusto López H", "candidato Adán Augusto López", la combinación "Adán Augusto López" y "MORENA", #AdanAugustoLópez, @adan_augusto
Ricardo Monreal	"Ricardo Monreal", "Ricardo Monreal A" , "Ricardo Monreal Ávila", "candidato Ricardo Monreal", la combinación "Ricardo Monreal" y "MORENA", #RicardoMonreal, @RicardoMonrealA

4

Rev. Cient. Sist. Inform. 5(1): e763; (Jan-Jun, 2025). e-ISSN: 2709-992X

The posts were collected using the search Twitter operator RapidMiner, which uses a connector to the Twitter/X API. For the download, the query or search concept, collection date, limit of posts to download, and language are defined. Samples were collected the following day. For example, the collection for June 20 was on June 21, and the limit of posts per day was 20 000; however, in none of the cases was this amount obtained. The language used was Spanish, and the same search parameters were used for all collection days. The results contain an attribute called "create at" that allows us to distinguish the day the post was created; through this parameter, it was possible to filter past posts.

2.2. Pre-processing

The preprocessing tasks include cleaning, processing, and transforming the posts; for example, the text of the post was processed to eliminate or replace symbols and remove references, among other tasks. The characteristics of each post were filtered and eliminated, repeated posts were eliminated, and these were transformed into data tables for the execution of the sentiment analysis algorithms.

2.3. Construction experimental sets

With the collected information, different data sets were generated by applying filters, and five sets were developed to analyze the posts from different perspectives. On the one hand, the posts that mentioned the search terms were analyzed, as were those posts that were shared at least once, those that were published through a mobile device, and those that were published at times when a social-political event occurred importantly.

2.4. Sentiment analysis

At this stage, sentiment analysis was configured for each set of data using the MeaningCloud tool, which resulted in a classifier text of each tweet into five categories according to the sentiment: very positive, positive, very negative, negative, and neutral. Each tweet is assigned to one of these categories based on the words that comprise it; if the category of the tweet cannot be defined, it is marked as an unidentified category.

2.5. Analysis of the results and conclusions

The study concluded with a two-phase examination of the findings. First, an analysis was carried out at the word level to determine how each applicant is perceived; for the analysis, word clouds were made from the collected and processed tweets, and empty words, mentions, and hashtags were removed from each tweet. The second stage consisted of an analysis of the results obtained in the classification of feelings, as well as a comparison with the actual results of the presidential candidate election.

Although, the research focuses on the political campaign for the election of candidates of the MORENA party in Mexico. This methodology can be applied to other political events, in other languages, or with different social networks. The preprocessing stage should be adapted according to the language used, and the ensemble construction stage should be carried out according to the objectives of the political analysis to be pursued. The tools used can be replaced by tools or computer programs that have a function like that presented.

3. RESULTS

3.1. Analysis of the posts collection

The results of the data collected per day and per candidate are shown in Figure 2. The applicant Marcelo Ebrard obtained the highest number of posts, with a total of 80 897, followed by Claudia Sheinbaum, with a total of 37 302; Adán Augusto López, with a total of 19 916; and Ricardo Monreal, with a total of 13 706.

We observe that the presence of Marcelo Ebrard in the social network with an official account since 2009, together with his extensive political career in various fields, was decisive in the number of posts that are made about him. Whereas in the case of Claudia Sheinbaum, she is the applicant with the most recent presence in social networks through an official account since 2012, in addition to the fact that her political career is more recent than the Aspiring Ebrard, in the case of Adán Augusto López and Ricardo Monreal, which have been present through an official account since 2010 and 2009, respectively, but they have held smaller political positions than Ebrard and Sheinbaum.

Figure 2 also shows that in the last days of the campaigns, the number of posts increased compared with the number of posts at the beginning, but in general, it remained within a constant range, except for the applicant Claudia Sheinbaum, whose posts were on the rise.
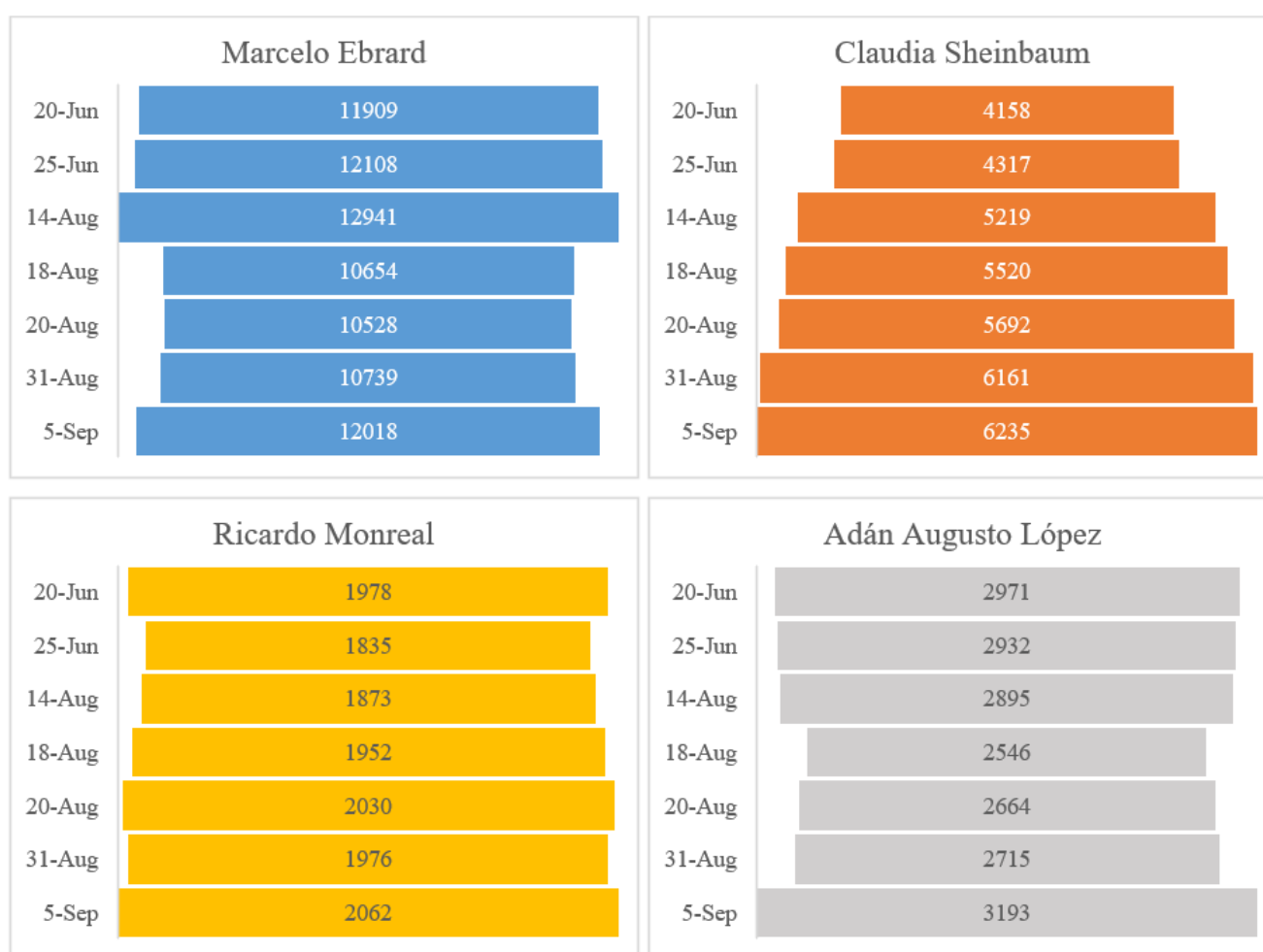

Figure 2. Posts collected by date and by applicant

3.2. Analysis by textual content

The first analysis, which was carried out with Set 1, which corresponds to the original set, consisted of the construction of word clouds for each of the aspiring candidates. The objective was to identify the perceptions of internet users according to the frequency of the words they use in their posts. In the case of Marcelo Ebrard, the words that are most frequently used in the posts are "justice", "support", "socialist" and "aggressor"; in the case of Claudia Sheinbaum, she is perceived as a "good person", "beautiful", "triumph" and "transparent". For Adán Augusto, the most frequent words are "transformation", "security", "people" and "justice"; finally, in the case of Ricardo Monreal, they were "transformation", "security", "victory" and "leader" (Figure 3).

6

Rev. Cient. Sist. Inform. 5(1): e763; (Jan-Jun, 2025). e-ISSN: 2709-992X

Marcelo Ebrard is the most mentioned in the posts because he already had a political career and had a current position before the selection process; that is, his passage through social networks was not recent. The same occurred with Claudia Sheinbaum, whereas the other two remaining applicants did not have enough presence on social networks before the selection process.
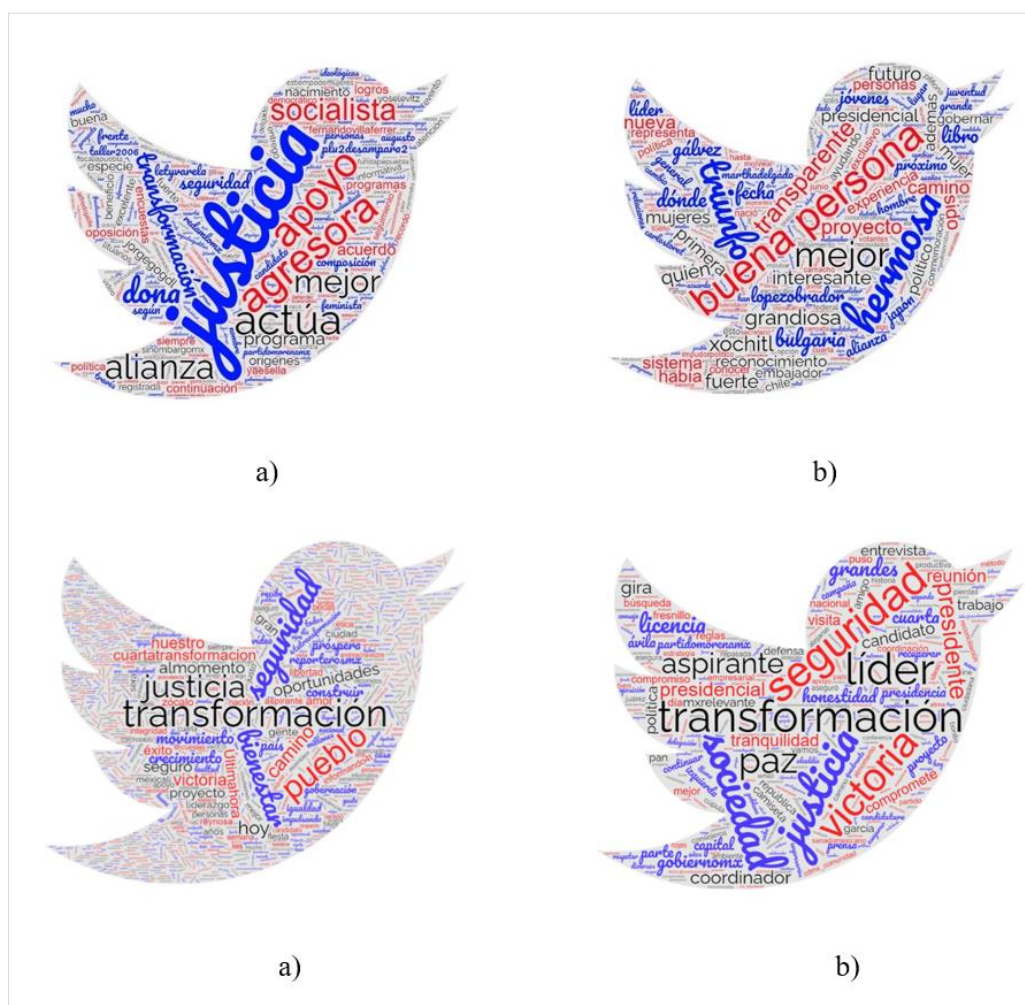
Figure 3. Word clouds of the posts collected for each candidate for the presidential candidacy of MORENA. a) Marcelo Ebrard, b) Claudia Sheinbaum, c) Adán Augusto López, d) Ricardo Monreal

An analysis of the vocabulary in the posts reveals that Adán Augusto's mentions are associated with a more diverse range of words. This suggests that his name appears in discussions covering wide topics, indicating a broader scope of context in which he is referenced.

3.3. Sentiment analysis

Sentiment analysis polarity determines the feelings expressed by users based on the relationships among words with positive, negative, and neutral aspects. The following datasets were created:

- Set 1: The collected tweets were without duplicating 73 291 posts.

- Set 2: The retweet attribute was only posts that had already been shared at least once were taken, and this set consisted of 65 834 posts.

- Set 3: This set considers only those tweets that were written and shared through a mobile device. This is an indication of the mobility of users and that they share events and their feelings in real-time, this set consists of 57 092 posts.

- Set 4: The posts shared in a relevant period were important, so this set was made of tweets that were shared only on the day of an important political episode for Mexico, since we consider that at this time, the users may have different opinions influenced by these episodes. The selected event was the "Reform of the Law to punish gender violence," which boomed from August 18 to 20, 2023; this set comprises 26 055 posts.

- Set 5: It is composed of posts on a relevant day, but through mobile devices, this set was made of 22 783 posts.

Table 2 shows the distribution of the datasets used in the development of this research, as the candidate with the most posts is Marcelo Ebrard, whereas Ricardo Monreal has the fewest posts.

Table 2. Distribution of collected posts

| Applicant	Posts collected	Posts pre-processing				
		Set 1	Set 2	Set 3	Set 4	Set 5
Marcelo Ebrard	80 897	39 742	37 427	35 190	11 149	9800
Claudia Sheinbaum	37 302	16 103	14 908	13 548	9118	8035
Adán Augusto López	19 916	9667	6549	4945	3121	2497
Ricardo Monreal	13 706	7779	6950	3409	2667	2451
Total	**151 821**	**73 291**	**65 834**	**57 092**	**26 055**	**22 783**

The results obtained for Set 1 are in Figure 4, where it is possible to see that the positive polarity is the predominant one in all the applicants. In the negative polarity, the applicant Claudia Sheinbaum goes to the head with some points. Adán Augusto López has the highest number of unclassified posts, those in which the polarity cannot be identified, which can be caused by the posts being very short or the words used not being placed in one of the categories. This phenomenon may occur due to the brevity of the posts or the inability to categorize the words used within the existing classifications. This phenomenon is common in social networks because occasionally abbreviations are used. Or the words are not written correctly or in full, for example, GPI, which means "thank you for inviting" or LOL, which means " laughing out loud". The use of new forms of language increases day by day daily in social networks, so this phenomenon will always happen with self-classification. The classified polarities are more than 90% accurate, and the strong positive polarities, marked with a + sign, have a higher percentage than the strong negative polarities.
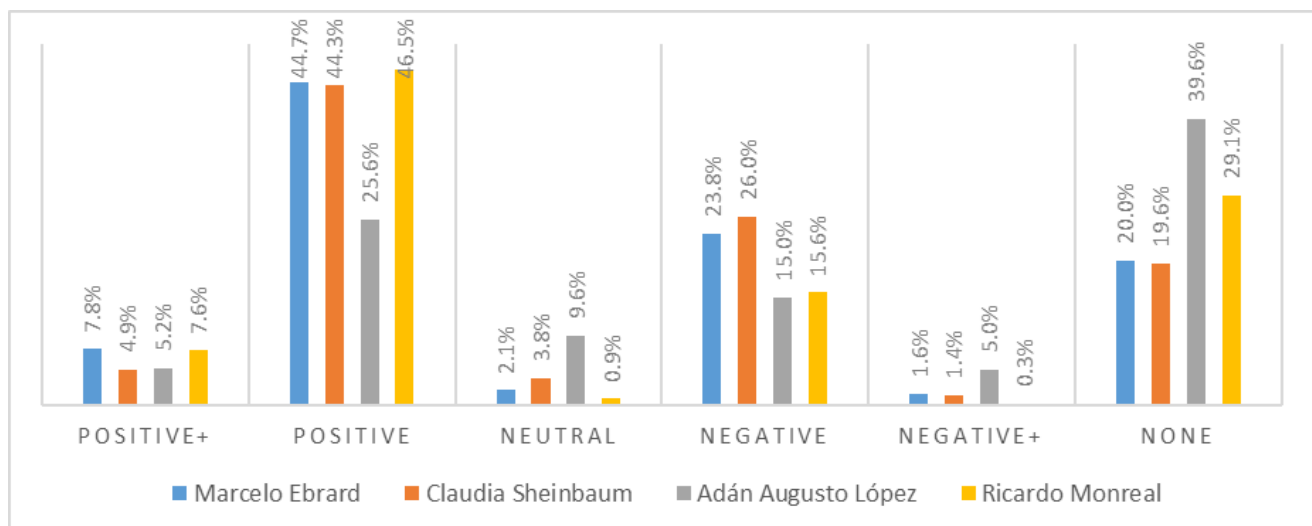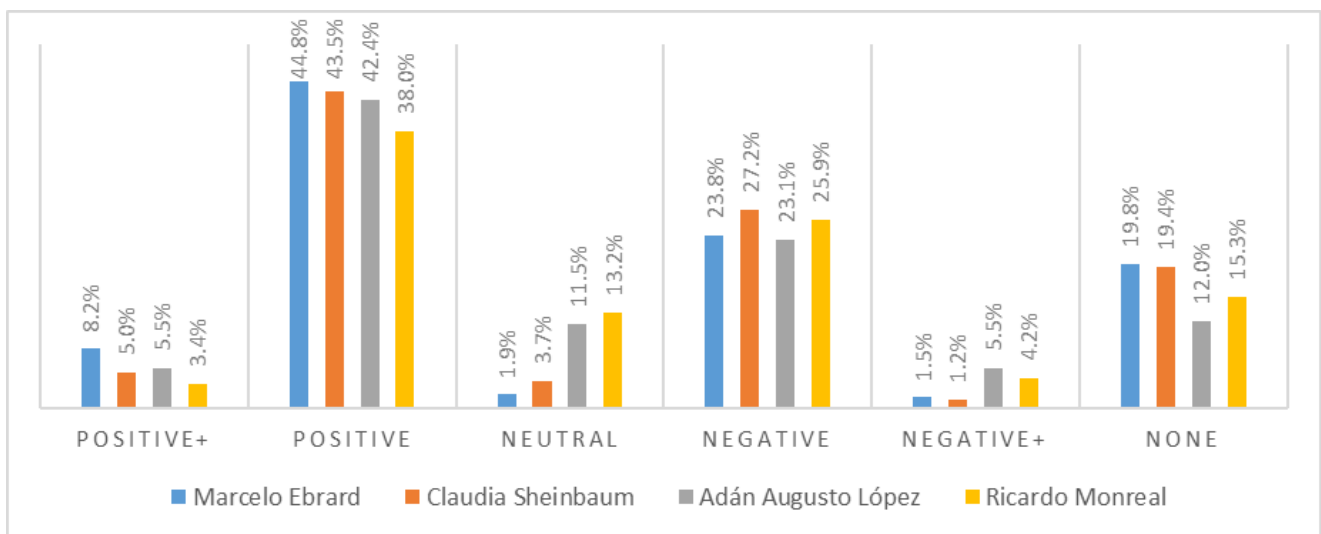

Figure 4. Polarity results obtained in Set 1

Figure 5 shows the results of Set 2, in which the posts were shared at least once. The experiment considered that because a shared post had more impact, the behavior of the polarities was among the applicants, but the effect decreased. The number of applicants with fewer posts increased the number of strong negative polarities, indicating that negative posts were shared in the case of these applicants.
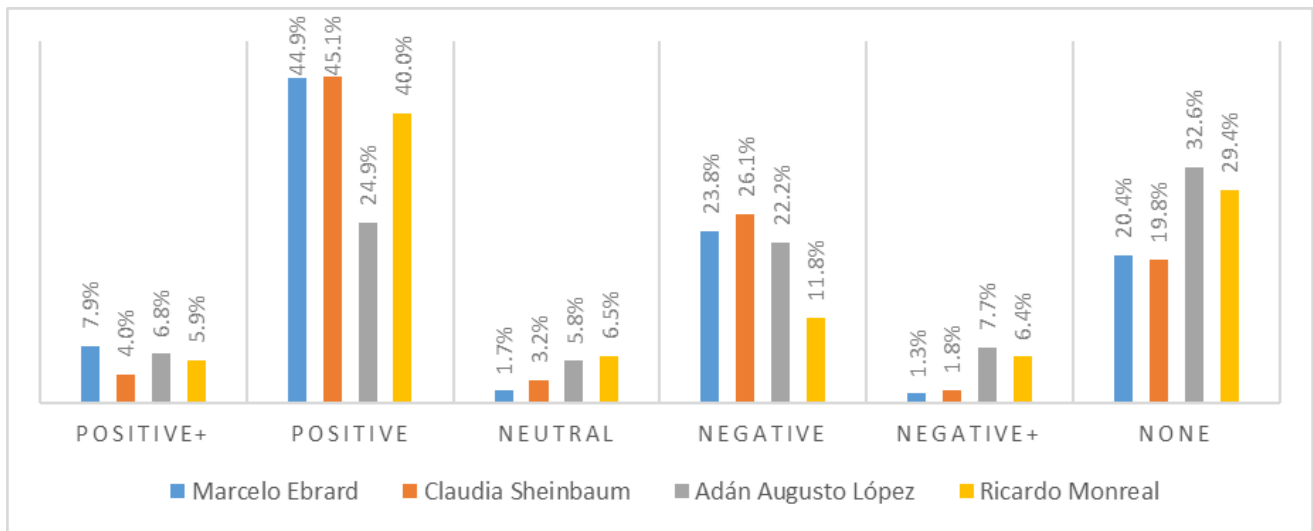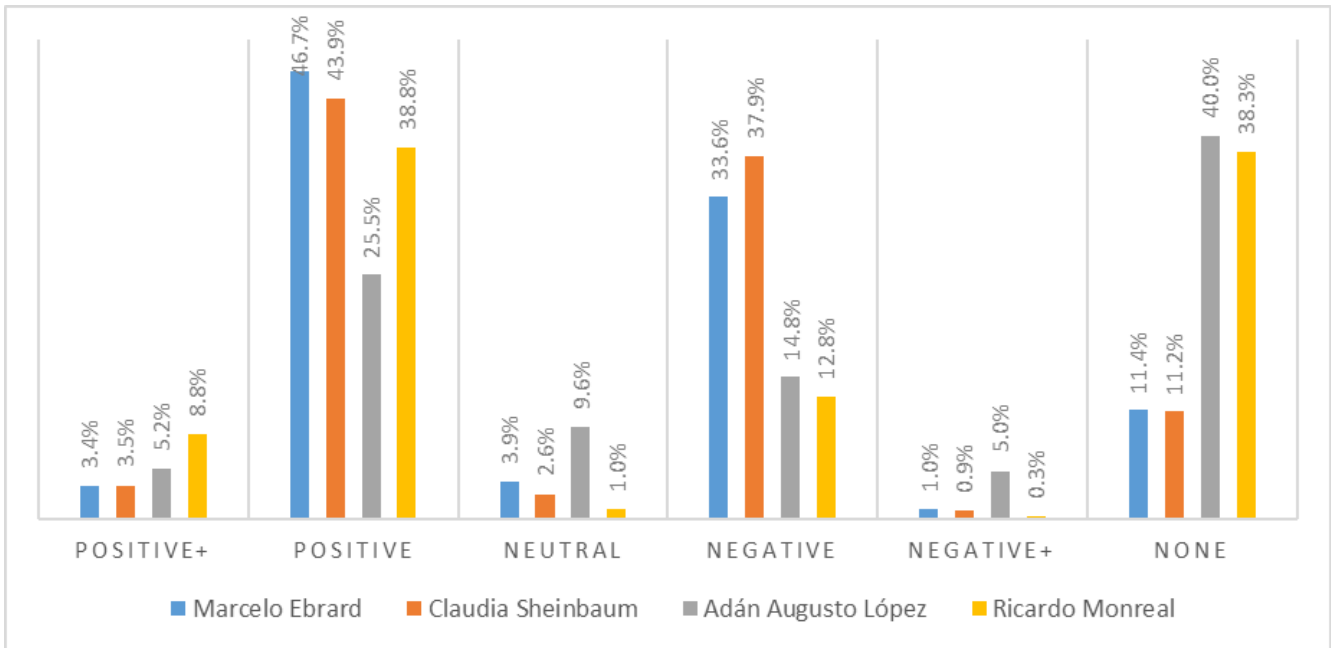
Figure 5. Polarity results obtained in Set 2

The results of Set 3, the posts written and shared by mobile devices shown in Figure 6, represented approximately one-third of the total posts collected. The analysis of this set is relevant because mobile devices allow the news to reach the user in real-time. And the reactions are instantaneous, which can motivate users to express a more sincere opinion since it is not influenced by other media. An increase was observed in the neutral comments of the two candidates with the least presence in social networks, and a balance between the positive and negative percentages helped to mark the trend toward the candidates. The number of unclassified posts decreased concerning Set 2.
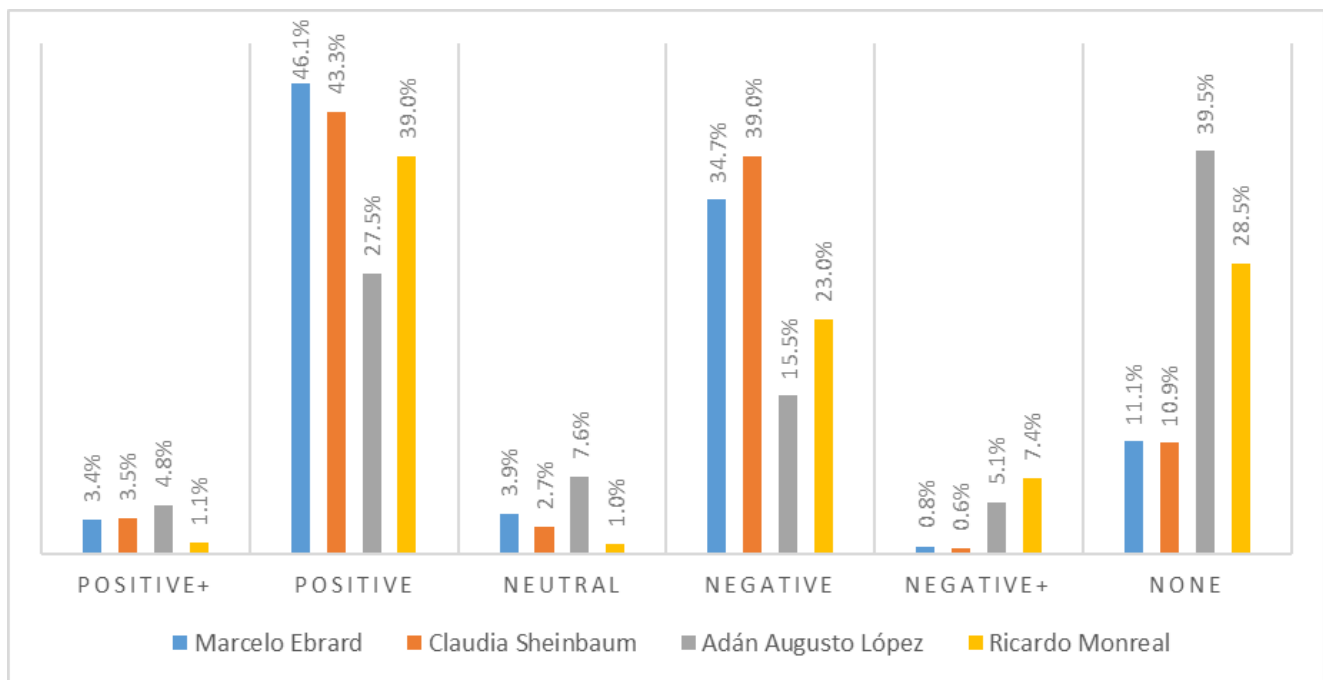

Figure 6. Polarity results obtained in Set 3

The posts were relevant political moments because of set 4 (Figure 7). It shows that for the candidates Claudia Sheinbaum and Marcelo Ebrard, the number of posts with negative polarity increases, and the number of neutral posts decreases. The polarity is affected by these political events, but the polarity remains stable.

Figure 7. Polarity results obtained in Set 4

The polarities of the posts of a relevant political event from mobile devices are in Figure 8, which corresponds to the experiments performed with Set 5. This behavior is like that shown in Set 4 as the number of posts increases and that of positive posts remains, which leads us to confirm that the impact of a political event does not depend on the medium through which it is disseminated.

Figure 8. Polarity results obtained in Set 5

The results obtained throughout the experiments show the aspiring candidates Marcelo Ebrard and Claudia Sheinbaum as pointers, giving the first more positive posts in each experiment, with a difference of less than 3 points, and the second of negative posts, greater when the difference is between 2 and 3 points. The CNN news agency (https://cnnespanol.cnn.com/2023/09/06/encuesta-morena-candidata-orix/) reported that the candidate Claudia Sheinbaum was the winner of the poll for the election of the presidential candidate of MORENA, with 39.4% of the votes, followed by Marcelo Ebrard with 25.8%, Adán

Augusto with 11.2%, and Ricardo Monreal with 5.9%, and the remaining percentage determined by two candidates who joined the race, with 17.8% of the vote.

4. DISCUSSION

Sentiment analysis has been used in various situations involving political events, obtaining multiple conclusions about its usefulness in predicting election results. According to our results, we agree with Sarapugdi & Namkhun (2023) that the prediction of an election is not predictable through the polarity obtained in the posts that talk about a particular candidate. The results may be biased for several reasons; for example, the user's perceptions become most negative when a relevant political event has been published.

The textual analysis and the methods used in the preprocessing of the posts indicated that there are many repeated or identical posts, which is related to the work of Cantini et al. (2022) about the use of bots to generate posts. Tendance of any subject, since these tools publish the same information through false profiles. Cheng et al. (2020) show that bots influence public opinion because, through the propagation of posts on a specific topic, they manage to skew the polarity on a topic; in this sense, we have more indications, but future work will be necessary to detect and validate our claims. For the detection of bots, methods based on machine learning can be used; for example, in the work of Aljabri et al. (2023) review of current techniques can be found, and we plan to explore these techniques in the future to demonstrate the impact of bots in political posts in Mexico.

In different studies analyzed, notable variation can be observed in the number of posts obtained and in the periods; for example, Caetano et al. (2018) analyzed almost 5 million posts referring to the election of states in the United States in 2016, which was obtained from August to November 2016, positive sentiment was superior for the candidate "Hillary Clinton" than for "Donald Trump"; however, she was not the winner of the election. Even when the number of posts is greater than that presented in this study, the result of the polarity of feelings did not coincide with the final results of the election.

In the work presented by Sarapugdi & Namkhun (2023), a prediction of the winners of the election of Thailand in 2023. More than 48 thousand X/Twitter posts were analyzed (five candidates), and the results revealed that factors such as the previous popularity of a candidate interfere with the user's feelings. Although the number of posts per candidate was not mentioned, an analysis of these results is presented. We agreed with the results since the candidate Marcelo Ebrard had higher previous popularity in social networks, which influences the number of posts obtained.

In the research by Ansari et al. (2020) for the elections in India 2019, 3896 posts were analyzed to demonstrate that the sentiment expressed in Twitter posts serves to classify preferences toward a political party. Social networks such as X/ / Twitter to evaluate the positioning of the candidates in an election is a good option to collect for collecting opinions from the voters. It shows that the leading candidates in the election coincide with the official results, which proves the results by authors such as Gilardi et al. (2022), who affirm that X/ / Twitter is the social network preferred by politicians and voters to express their opinions.

CONCLUSIONS

Based on the results of the perceptions of users of the social network X/ / Twitter regarding the aspiring presidential candidates in Mexico of the MORENA party, from the classification of feelings of the posts made in the period that lasted the campaigns, it can be concluded that: 1) The polarity expressed in user posts were positive for all candidates. 2) The comparative analysis revealed that the applicant Marcelo Ebrard had the highest number of posts with positive polarity in the experiments, followed by Claudia Sheinbaum.

3) The number of posts impacts on the election results, candidates with more posts were positioned better in the election results. 4) The number of negative posts increases when political events are published.

Our findings suggest that determining the winner in a competition requires more than analyzing the sentiment expressed by participants. It is crucial to consider various technological and social elements that could influence voters' ultimate decisions. The results obtained are valuable for understanding the impact of social media campaigns and their influence on public opinion, which can help improve political strategies and decision-making in future elections. The adoption of this method of analysis can be used to direct the content of a political campaign and increase the number of posts by observing the behavior of the users; however, other aspects, such as the number of followers of an account, the detection of bots and the timely monitoring of the content that is published, must be considered.

FINANCING

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest related to the development of the study.

AUTHORSHIP CONTRIBUTION

Conceptualization; Data Curation; Formal Analysis; Research; Methodology; Visualization; Validation; Writing - original draft; Writing - revision and editing: Denicia-Carral, M. C., Ballinas-Hernández, A. L., Minquiz-Xolo, G. M. and Medina-Cruz, H.

REFERENCES

Aljabri, M., Zagrouba, R., Shaahid, A., Alnasser, F., Saleh, A., & Alomari, D. M. (2023). Machine learning-based social media bot detection: a comprehensive literature review. *Social Network Analysis and Mining*, *13*(1), 20. https://doi.org/10.1007/s13278-022-01020-5

Ansari, M. Z., Aziz, M. B., Siddiqui, M. O., Mehra, H., & Singh, K. P. (2020). Analysis of Political Sentiment Orientations on Twitter. *Procedia Computer Science*, *167*, 1821-1828. https://doi.org/10.1016/j.procs.2020.03.201

Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert Systems with Applications*, *164*, 114006. https://doi.org/10.1016/j.eswa.2020.114006

Bharany, S., Alam, S., Shuaib, M., & Talwar, B. (2023). *Sentiment Analysis of Twitter Data for COVID-19 Posts* (pp. 457-466). https://doi.org/10.1007/978-981-19-6004-8_37

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. https://doi.org/10.1016/j.knosys.2021.107134

Caetano, J. A., Lima, H. S., Santos, M. F., & Marques-Neto, H. T. (2018). Using sentiment analysis to define twitter political users' classes and their homophily during the 2016 American presidential election. *Journal of Internet Services and Applications*, *9*(1), 18. https://doi.org/10.1186/s13174-018-0089-0

Cantini, R., Marozzo, F., Talia, D., & Trunfio, P. (2022). Analyzing Political Polarization on Social Media by Deleting Bot Spamming. *Big Data and Cognitive Computing*, *6*(1), 3. https://doi.org/10.3390/bdcc6010003

Chaudhry, H. N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z. I., Shoaib, U., & Janjua, S. H. (2021). Sentiment Analysis of before and after Elections: Twitter Data of U.S. Election 2020. *Electronics*, *10*(17), 2082. https://doi.org/10.3390/electronics10172082

Cheng, C., Luo, Y., & Yu, C. (2020). Dynamic mechanism of social bots interfering with public opinion in network. *Physica A: Statistical Mechanics and its Applications*, *551*, 124163. https://doi.org/10.1016/j.physa.2020.124163

Espejel Espinoza, A., & Díaz Sandoval, M. (2022). *Tendencias organizacionales y democracia interna en los partidos políticos en México. Los casos del PAN, PRI, PRD, PT, PVEM, MC y MORENA* (1.ª ed.). Facultad de Estudios Superiores Acatlán, UNAM.

Flamino, J., Galeazzi, A., Feldman, S., Macy, M. W., Cross, B., Zhou, Z., Serafino, M., Bovet, A., Makse, H. A., & Szymanski, B. K. (2023). Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01550-8

Garg, P. K., Pandey, M., & Arora, M. (2020). Sentiment Analysis for Predicting the Popularity of Web Series. En *Communications in Computer and Information Science* (1.ª ed., pp. 133-140). Springer Nature. https://doi.org/10.1007/978-981-15-5830-6_12

Gilardi, F., Gessler, T., Kubli, M., & Müller, S. (2022). Social Media and Political Agenda Setting. *Political Communication*, *39*(1), 39-60. https://doi.org/10.1080/10584609.2021.1910390

Lucca, J. B. (2019). Teoría y política en la génesis de MORENA como nuevo partido. *Estudios Políticos*, *49*. https://doi.org/10.22201/fcpys.24484903e.2020.49.72396

Noor, H. M., Turetken, O., & Akgul, M. (2024). Social Media, Sentiments and Political Discourse – An Exploratory Study of the 2021 Canadian Federal Election. *ACM Transactions on Social Computing*, *7*(1-4), 1-23. https://doi.org/10.1145/3665450

Oliveira, D. J. S., Bermejo, P. H. de S., Pereira, J. R., & Barbosa, D. A. (2019). The application of the sentiment analysis technique in social media as a tool for social management practices at the governmental level. *Revista de Administração Pública*, *53*(1), 235-251. https://doi.org/10.1590/0034-7612174204

Ongo Nkoa, B. E., Ondoua Beyene, B., Ngo Nsoa Simb, J. F., & Ngnouwal Eloundou, G. (2023). Does social media improve women's political empowerment in Africa? *Telecommunications Policy*, *47*(9), 102624. https://doi.org/10.1016/j.telpol.2023.102624

Oueslati, O., Hajhmida, M. Ben, Ounelli, H., & Cambria, E. (2023). Sentiment Analysis of Influential Messages for Political Election Forecasting. En *International Conference on Computational Linguistics and Intelligent Text Processing* (1.ª ed., pp. 280-292). La Rochelle. https://doi.org/10.1007/978-3-031-24340-0_21

Russo, R., Blikstein, P., & Literat, I. (2024). Twisted knowledge construction on X/Twitter: an analysis of constructivist sensemaking on social media leading to political radicalization. *Information and Learning Sciences*, *125*(9), 693-719. https://doi.org/10.1108/ILS-12-2023-0210

Sarapugdi, W., & Namkhun, S. (2023). A Social Analysis of Thailand's 2023 Election Through Twitter Feeds. *2023 15th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 208-212. https://doi.org/10.1109/ICITEE59582.2023.10317682

Sobkowicz, A., & Kozłowski, M. (2018). Sentiment Analysis in Polish Web-Political Discussions. En *Human Language Technology. Challenges for Computer Science and Linguistics* (pp. 363-377). Springer Nature. https://doi.org/10.1007/978-3-319-93782-3_26

Turón, A., Altuzarra, A., Moreno-Jiménez, J. M., & Navarro, J. (2023). Evolution of social mood in Spain

13

Rev. Cient. Sist. Inform. 5(1): e763; (Jan-Jun, 2025). e-ISSN: 2709-992X

throughout the COVID-19 vaccination process: a machine learning approach to tweets analysis. *Public Health*, *215*, 83-90. https://doi.org/10.1016/j.puhe.2022.12.003

Vyas, V., & Uma, V. (2018). An Extensive study of Sentiment Analysis tools and Binary Classification of tweets using Rapid Miner. *Procedia Computer Science*, *125*, 329-335. https://doi.org/10.1016/j.procs.2017.12.044