Multi-Dimensional Insights: Annotated Dataset of Stance, Sentiment, and Emotion in Facebook Comments on Tunisia's July 25 Measures

Sanaa Laabar, Wajdi Zaghouani

Hamad Bin Khalifa University
College of Humanities and Social Sciences

sala29626@hbku.edu.qa, wzaghouani@hbku.edu.qa

Abstract

On July 25, 2021, Tunisian President Kais Saied announced the suspension of parliament and dismissal of Prime Minister Hichem Mechichi, a move that sparked intense public debate. This study investigates Tunisian public opinion regarding these events by analyzing a corpus of 7,535 Facebook comments collected from the official Tunisian presidency page, specifically the post announcing the July 25 measures. A team of three annotators labeled a subset of 5,000 comments, categorizing each comment's political stance (supportive, opposing, or neutral), sentiment (positive, negative, or neutral), emotions, presence of hate speech, aggressive tone, and racism. The inter-annotator agreement, measured by Cohen's kappa, was 0.61, indicating substantial consensus. The analysis reveals that a majority of commenters supported President Saied's actions, outnumbering those who opposed or took a neutral stance. Moreover, the overall sentiment expressed in the comments was predominantly positive. This study provides valuable insights into the complex landscape of public opinion in Tunisia during a crucial moment in the country's ongoing political transformation, highlighting the role of social media as a platform for political discourse and engagement.

Keywords: Sentiment Analysis, Hate Speech, Digital Humanities, Data Annotation

1. Introduction

Eleven years have passed since Mohammed Bouazizi, a Tunisian fruit and vegetable vendor, set himself on fire in the town of Sidi Bouzid, Tunisia. This incident sparked the Arab Spring, a series of revolutions that began in Tunisia and proceeded to engulf the Arab World, including Egypt, Libya, Yemen, and Syria. In 2011, Tunisia's smooth transition into a democracy was considered a success story and a beacon of hope for democracy to its Arab neighbors who had suffered from political turmoil and civil wars after their revolutions. The Jasmine revolution erupted as a strike back to the unanswered desperate call to change the country's rising unemployment, food inflation, corruption, lack of political freedom, and poor living conditions. Demonstrations and protests spread in every governorate in the country and President Ben Ali fled through a private presidential jet to Kingdom of Saudi Arabia where he sought political refuge until his passing in 2019 in Jeddah. Unfortunately, Tunisia's democratic and progressive state ceased to exist when President Kais Said won the elections and changed the country's direction in a sudden unprecedented decision two years after his election. On the night of the 25th of July 2021, Said suspended all the works of the Tunisian parliament and dismissed his head of government, Hichem Mechichi. After living in a state of democracy for 10 consecutive years, Tunisia had suddenly been catapulted into a different political reality. This action was termed as a coup against the revolution and the constitution by his opposers; however, supporters saw it as correcting the revolution. Within one night, Tunisia witnessed a clear and sharp division amongst its people between those who expressed their support for the 25th of July measures and those who opposed them. According to Yerkes and Mbarek, "the very fact that people can freely and publicly express their criticism of the government without fear of harm or retribution is a dramatic achievement" (Yerkes & Mbarek, 2021, p.1). Therefore, the political polarity that the country witnessed literally overnight is what motivated this research and prompted the investigation to determine the dominant political stance group. This fascinating division was evident all-over social media platforms. Interestingly, among these platforms, Facebook stood out, given that it is the most popular social media platform used in the country (Statista, 2021).

The political landscape in Tunisia underwent a fundamental transformation following the revolution. After the flight of President Ben Ali, Foued Mebazaa, then Speaker of Parliament, assumed the presidential role, as detailed by Zayani (2015). Subsequently, Prime Minister Mohamed Ghannoushi signaled his intent to establish a transitional government. However, escalating protests in Al-Kasbah led to his resignation. This period marked a pivotal transition towards progressive goals, including political reform and democratic transition, as envisioned by the Committee of Safeguarding the Revolution and the Commission for Political Reform.

Expectations for a reformed Tunisia burgeoned, setting the stage for growing disenchantment with political entities, particularly in the context of the failed promises and objectives of the revolution. The emergence of Ennahda, a moderate Islamist party

once banned, into power, and its coalition with secular parties, Congress for the Republic and the Democratic Front for Labor and Liberty, signified a significant political shift. However, the assassinations of secular opposition members Chokri Beleid and Mohamed Brahmi implicated Ennahda, exacerbating public discontent and weakening the coalition known as the Troika, amid mounting economic and social instability.

These dynamics, coupled with intensified political and ideological divisions, underscored the crumbling revolutionary objectives, as Zayani (2015) points out for over 10 years post-revolution, and led to a disillusionment with the traditional political framework, as Wolf (2019) observes. The context of perceived political failures catalyzed the public's disaffection with established political parties, influencing electoral outcomes. Mounting dissatisfaction for the political and economic status-quo grew gradually in the last 10 years prior to the 2021, 25th of July measures. Despite his lack of conventional political experience and affiliations, Kais Saied's ascension to the presidency with a significant electoral margin highlighted the populace's yearning for change. Saied's victory, as Allahoum (2019) notes, was notable given his minimal campaign efforts and outsider status in the political arena. This trust in Saied and his decisive actions can be seen as a direct response to the accumulated frustration and skepticism towards the conventional political parties and their historical legacies in Tunisia.

Consequently, as this political event is a recent event in the country's history, not enough research has been published on the topic. Hence, to study this event, a corpus that includes 5000-annotated comments retrieved from the Tunisian presidency's Facebook page through Facepager is created, specifically from the Facebook post that announced the 25th of July measures. The aim of this research is to document and analyze two groups based on their political stance toward the measures to determine Tunisians' public opinion. Moreover, our study examines the sentiment, emotion, and hate speech in the comments and their connection to a corresponding political stance. Additionally, it analyses and compares word frequencies between both groups using the tool AntConc.

2. Related Work

The current body of literature demonstrates the evolving application of digital tools in media and communication studies, as evidenced by research from Mahmadi et al. (2017) and Saad & Sabrini-Zin (2022), which illustrate the application of these tools in extracting and analyzing digital content. The evolving landscape of news dissemination underscores the significant role of digital tools in adapting to the changing patterns of news consumption. Within this context, the United States elections emerge as a focal point, attracting extensive coverage and prioritization in news monitoring efforts. Mhamdi et al. (2017) illustrated this by employing Facepager to gather data from the Facebook pages of prominent news channels, namely Fox News, CNN, and ABC News. This data was subsequently analyzed using the same tool to understand the dynamics of digital news dissemination. Additionally, the analytical capabilities of digital tools extend beyond news analysis, as demonstrated by Saad & Sabrini-Zin (2022), who utilized AntConc to examine the lexical features of Robert Frost's poem \"Into My Own.\" These instances highlight the multifaceted applications of digital tools in both news analysis and literary studies, reflecting their integral role in contemporary research methodologies. The study demonstrated that the use of the tool facilitated enhanced insights compared to earlier research reliant solely on manual analysis. Despite the qualitative nature of the study's methodology, a descriptive approach was employed to quantitatively analyze the data extracted from the text. This methodological combination allowed for a more nuanced understanding of the textual information, showcasing the advantage of integrating digital tools in the analytical process.

To assess public sentiment regarding the July 25 measures via traditional news media, multiple surveys were executed. Shems FM, a local Tunisian News Agency, surveyed 1,707 individuals aged 18 and above from August 4 to 13. The findings showed substantial support: 87.3% of the respondents backed the decisions made on that date, 81.6% endorsed the suspension of the parliament, and 76.2% agreed with the dismissal of Prime Minister Hichem Mechichi. In a separate study by L'Economiste Maghrebin, 94.9% of participants expressed support for these exceptional measures (Marzouk, 2021). Additionally, Sigma Conseil conducted a poll reflecting a 72.2% approval rate for President Kais Saied and his policies, further indicating significant public endorsement of the actions taken on and following July 25 (2021).

Our methodology is informed by Zaghouani and Awad's (2016a) work on developing an Arabic punctuated corpus, which provides critical insights into annotation guidelines. Similarly, the comprehensive framework for annotating Arabic corpora for machine translation, as elucidated by Zaghouani et al. (2016), offers a robust model for our annotation processes. Our dataset creation approach draws upon Hawwari et al.'s (2016) meticulous annotation of Arabic morphological patterns, presenting a nuanced understanding of linguistic intricacies. Zaghouani et al.'s (2010) work on the revised Arabic Propbank further guides our dataset structuring, emphasizing the importance of detailed proposition annotation in political discourse analysis.

Bianchi et al. (2023) explore digital communication's nuances, shedding light on interactive dynamics relevant to our analysis of social media political framing. Biswas et al.'s (2023) examination of Twitter content for vaccine-related discussions exemplifies the potential of social media analytics in extracting meaningful insights from online discourse, paralleling our sentiment and sarcasm detection efforts as discussed by Farha et al. (2021).

Moreover, the foundational work by Obeid et al. (2016) on Arabic diacritization and the structured annotation processes outlined by Zaghouani and Awad (2016b) significantly influence our annotation guideline development. The creation of a multi-genre Arabic corpus by Bouamor et al. (2016) and the dataset focusing on political framing in the U.S. COVID-19 discourse by Shurafa et al. (2020) provide methodological blueprints for our data compilation and analysis efforts.

Our research is further contextualized within the broader discourse of social media data analysis. The extraction and examination of dialectal Arabic irony from Twitter, as conducted by Abbes et al. (2020), and the detection of propaganda in Arabic content, as explored by Alam et al. (2022), offer pertinent insights into the complexities of online political communication. The multi-dialect Twitter corpus analysis by Zaghouani and Charfi (2018a) enriches our understanding of language variety and demographic factors in social media interactions.

In synthesizing these diverse studies, our research aims to extend the existing scholarly discourse on political framing and annotation within the dynamic realm of social media, leveraging the rich corpus of Arabic language content and the multifaceted methodologies established in the aforementioned works.

3. Methodology

In this investigation, the Facepager application, conceived in 2019 by communication scientists Jakub Jünger of the University of Greifswald and Till Kelling of Ludwig-Maximilians University, served as the primary tool for data acquisition. Specifically, the study targeted Facebook comments extracted from a post on the official Tunisian presidency's Facebook page. This post pertained to the unprecedented suspension of parliament, a pivotal event in Tunisian political discourse. The official Facebook page of the Tunisian presidency, established in 2011 and becoming active in content posting since 2012 under President Moncef Marzouki, acts as a crucial digital platform for presidential communication. Its significance has been consistent through the administrations of Presidents Moncef Marzouki, Beji Caid Essebsi, and the incumbent president, Kais Said, marking it as a significant source of official presidential communications and public engagement. For the data extraction process, a MacBook Air running the macOS Big Sur operating system was utilized to operate the Facepager software.

This tool facilitated the efficient extraction of relevant comment data, which is typically exported into a .csv format suitable for analysis in spreadsheet applications like Macintosh's Numbers. However, to streamline the analytical workflow and enhance the ease of data annotation, the extracted Facebook comments were systematically transferred to a Google spreadsheet. This transition allowed for a more manageable and interactive engagement with the data.

Complementing the data collection process, the study incorporated the digital text analysis tool AntConc. This software, developed by Dr. Laurence Anthony, a professor and software engineer at Waseda University, Japan, is instrumental in performing a comprehensive analysis of textual data. AntConc's sophisticated functionalities support a deep dive into the linguistic and thematic elements of the extracted comments, facilitating a nuanced exploration of public sentiment and discourse patterns. By integrating these methodological tools, the research aims to meticulously parse the digital public sphere's reactions and interactions concerning the significant political development of parliament suspension in Tunisia, thus enabling a rich and insightful examination of the public discourse captured on this official digital platform.

3.1 Linguistic Variations in Tunisian Social Media Discourse

The Tunisian social media landscape is characterized by unique linguistic variations, particularly the use of "Arabizi" and code-switching between Arabic, French, and English. Arabizi refers to the romanization of Arabic text, where users employ Latin script to write Arabic words and phrases (Bjørnsson, 2010). This phenomenon is prevalent in informal online communication, especially among younger generations. Additionally, code-switching, the alternation between multiple languages within a single conversation or utterance, is a common practice in Tunisian social media discourse (Kebeya, 2013). These linguistic variations pose challenges for sentiment analysis, as they introduce non-standard orthography and complicate the identification of language-specific features. In our corpus, we observed several instances of Arabizi and code-switching, such as "sbe7 el khir" (good morning) and "vive la Tunisie" (long live Tunisia). Future research could explore methods to effectively handle these linguistic phenomena in sentiment analysis tasks.

3.2 Data Collection and Sample Representativeness

The Facebook comments analyzed in this study were collected on [insert date] from the official Tunisian Presidency Facebook page, specifically focusing on the post announcing the July 25 measures. A total of 7,535 comments were retrieved, out of which 5,000 were randomly selected for annotation. While the random sampling method aimed to ensure an unbiased representation of the overall comment population, it is important to acknowledge that Facebook users may not be entirely representative of the Tunisian general public. According to the World Bank (2021), around 66% of the Tunisian population had access to the internet in 2020, suggesting that a significant portion of the population may not be active on social media platforms. Additionally, the demographic composition of Facebook users in Tunisia may skew towards younger age groups and those with higher digital literacy. Future research could benefit from incorporating demographic data of the commenters, if available, to assess the

representativeness of the sample and potential biases in the data.

3.3 Clarification on Stance and Emotion Annotation

In our study, stance and emotion were annotated independently by the human annotators. Stance was categorized as either pro-Saied (supporting the July 25 measures), anti-Saied (opposing the measures), or neutral. Emotions, on the other hand, were annotated based on the presence of specific emotional cues in the comment text, such as joy, trust, anger, or fear. The phrase "stance count by predominant emotions" in the Results section refers to the distribution of stance categories (pro, anti, or neutral) within each emotion category. For example, among the comments annotated with the "joy" emotion, we examined the proportion of pro-Saied, anti-Saied, and neutral stances. This analysis aimed to uncover potential correlations between specific emotions and political stances. However, it is crucial to note that emotions were not used as determinants of stance; rather, they were treated as separate but potentially related dimensions of the public opinion landscape.

3.4 Data Collection

The data for this study were collected using the Facepager tool, specifically targeting comments from the official announcement regarding the temporary 30-day suspension of Tunisia's Assembly of the Representatives of the People. As depicted in Figure 1, this announcement was made on the official presidency page, attracting significant interaction, evidenced by the retrieval of 7,572 comments. Due to time constraints, a subset of 5,000 comments was selected for annotation.

Figure1: Presidency Facebook Page

Moreover, data collection utilized AntConc, a tool that necessitates a reference corpus to activate features such as "Word List" and "Key Word List." For this purpose, a corpus comprising 400 articles from Economic Arabic Newspapers (Al-Sluaiti & Atwell, 2003) was employed. Upon processing this corpus, AntConc generated a keyword list highlighting terms with unexpectedly high frequency compared to those in the reference corpus and identified collocates to reveal words frequently associated with the search terms, assessing the association's strength.

Additionally, manual annotation was conducted to differentiate between subjective sentiments and objective facts, incorporating thematic analysis to evaluate emotions and feelings with attention to subjectivity and complexity (Monkey Learn, 2022; Williams et al., 2019). Hate speech detection adhered to the criteria from Hate.org, focusing on malicious intentions towards specific groups (Crabb et al., 2019). Despite the inherent subjectivity in such analysis, the pre-trained annotators achieved a kappa score of 0.61, signifying substantial inter-annotator agreement. This was based on a blind sample of 200 comments, indicating a high consensus among the annotators in determining sentiment and emotion.

4. Results

The findings from this study provide a nuanced understanding of public sentiment, emotions, and stances expressed in the Tunisian political discourse, particularly following the July 25 measures.

4.1 Sentiment Analysis

This section illustrates a quantification of sentiments extracted from the Facebook comments. According to Google sheets findings, the sentiment results in Figure 2 show a higher value in "Very Positive" by 2,711 comments and "Positive" by 1,180 comments. On the other hand, the "Very Negative" sentiment is expressed in 307 comments, and the "Negative" sentiment in 112 comments, which are significantly less in comparison. The sentiment analysis reveals a predominantly positive public opinion towards the July 25 measures, with the combined "Very Positive" and "Positive" sentiments accounting for a substantial majority of the analyzed comments. This finding suggests a broad base of support for President Saied's actions among the Facebook users who engaged with the official presidency page.

4.2 Emotions and Feelings Analysis

The analysis of feelings and emotions, as shown in Figure 3, reveals that "Trust" and "Joy" are the predominant emotions, reflecting a generally favorable public sentiment towards the political developments. The value of "Trust" is expressed in 2,883 comments, whereas "Joy" is present in 832 comments. "Neutral" emotions are found in 597 comments, and "Anger" is detected in only 353 comments. The prevalence of "Trust" and "Joy" in the analyzed comments underscores the public's confidence in and enthusiasm for the measures taken by President Saied. These positive emotions align with the overall supportive stance towards the July 25

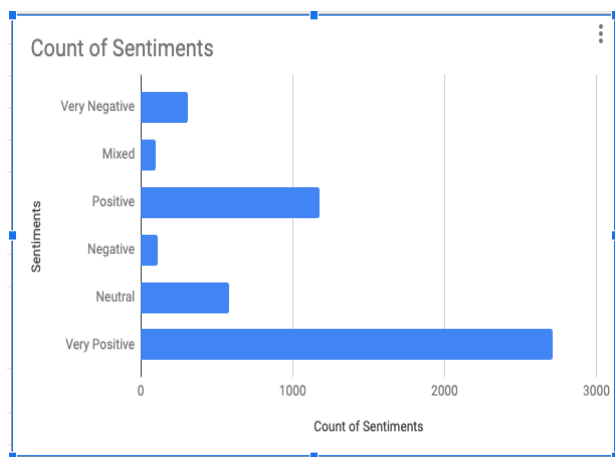events, as evidenced by the sentiment analysis results.
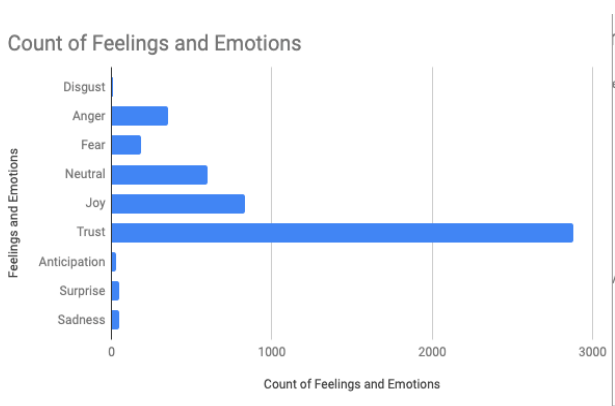

Figure 2: Count of Sentiments

Figure 3: Count of Feelings and Emotions

4.3 Stance Analysis

The predominant emotions are reflected in the stance count, as shown in Figure 4, where a significant majority of comments supported the July 25 measures, with neutral and opposing stances being less prevalent.

The count of stance indicates that most comments, 78.6% (3,922 comments), prove to be "pro" the 25th of July measures. Moreover, the "Neutral" stance, at 12.9% (646 comments), is more common than the "Against" stance, which accounts for only 8.4% (422 comments).

The stance analysis confirms the overwhelming public support for President Saied's actions, with the "pro" stance dominating the discourse on the official presidency Facebook page.

The relatively low proportion of "Against" comments suggests that the opposition to the measures was limited, at least within the scope of this study's dataset.
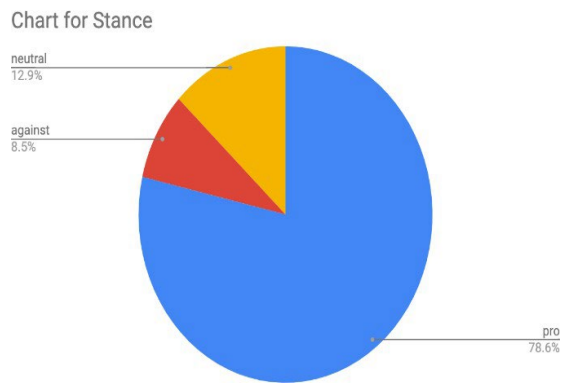

Figure 4: Count of Stance

4.4 Hate Speech, Aggressive Tone, and Racism Analysis

Despite the overall supportive sentiment, annotators assessed the presence of hate speech, aggressive tone, and racism in each comment, assigning a binary value of 'yes' or 'no'. Figure 5 shows the analysis, which revealed that a significant majority, 83.2% (4,076 comments), did not exhibit these negative tones, while only 16.6% (924 comments) did.

Notably, among the comments identified with hate speech, aggressive tone, or racism, the majority were supportive ('Pro') in stance, totaling 835 comments. Conversely, 79 comments were opposed ('Against'), and a mere 10 comments were categorized as 'Neutral'. This distribution underscores the predominance of such negative tones in supportive comments within the dataset.
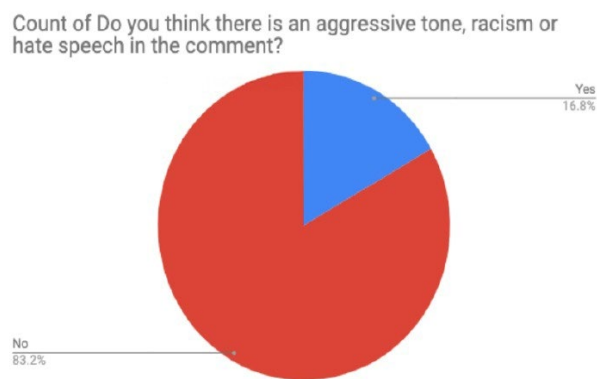

Figure 5: Count of Hate Speech, Aggressive Tone, and Racism

The analysis of hate speech, aggressive tone, and racism provides a more nuanced perspective on the public discourse surrounding the July 25 measures. While the overall sentiment was largely positive and supportive, the presence of these negative elements, particularly within the "pro" stance comments, highlights the potential for polarization and the expression of extreme views, even among supporters of President Saied's actions.

26

4.5 AntConc Analysis

The complexity of the public discourse is further explored in AntConc's word list analysis. By examining the corpus using AntConc, the findings show that the word "Tunis تونس" occurred 1,836 times, "people 1,234 الشعب" times, "President 1,106 الرئيس" times, and "with you 826 معاك" times. However, in the keyword list, it is noticeable that the word "Coup D'etat انقلاب" only occurred 284 times, and the word "army 191 الجيش" times.

To observe the differences in words associated with President Kais Saied and the most popular Islamic political party in the country, the researcher input common words from the corpus, such as "سعيد" ("Said," the last name of President Kais Saied), "Ennahdha النهضة," and "Brotherhood خوانجية." The collocates linked to President Saied convey positive sentiments with phrases like "elevate you يرفعك," "protect him يحفظو," and "we support نساند," indicating support. Conversely, collocates related to the Ennahdha party "النهضة" and the Brotherhood "خوانجية" are negative, showing aggression and hate speech with terms like "they destroy it يهلكوها," "they kill us يقتلونا," and "arrest them يعتقلوهم." Negative words such as "terrorists إرهابيين," "they stole it سرقوه," and "they destroyed it دمروه" are also associated with these groups, highlighting a stark contrast in sentiment.

4.6 Keyword and Collocates Analysis

The results of the keyword list analysis, presented in Table 1, display an interesting set of key terms that were predominant in the corpus. Keywords that apparently favor Saied and his decisions rank highest in keyness and frequency, while the word "Coup" has the lowest value.

Furthermore, the collocates match the percentage of stance and carry positive sentiments, emotions of joy, and evident support towards President Kais Saied, as expected. On the other hand, the term "Coup" seems to carry a majority of collocates that have neutral stances or sentiments, with the remaining collocates expressing positive sentiments that indicate a supportive stance towards President Saied.

4.7 Corpus and Annotation Challenges

Finally, in contextualizing the corpus size and annotation challenges, it should be noted that the corpus comprises 7,535 Facebook comments, of which 5,000 were annotated across five dimensions: comment text, sentiment, stance, emotions, and hate speech. The annotations were organized in two Excel tabs, providing a detailed view of the average pairwise annotation agreements between the three annotators for sentiment and emotion categories. A recurrent issue in the annotation process was the determination of sentiment and emotion intensity levels.

Divergences in annotation were particularly notable in the assessment of sentiments and emotions. Annotators often disagreed on whether a comment should be classified as negative or very negative and in distinguishing between emotions like trust and joy. These discrepancies were largely due to differences in how annotators perceived the commenter's intentions and emotional responses to President Saied's actions. For example, the comment "# لا للانقلاب No to Coup" was deemed negative by one researcher, who argued that the comment lacked the linguistic intensity to be categorized as very negative. Conversely, a second annotator interpreted the same comment as very negative, illustrating the subjective nature of interpreting sentiment and emotion in textual analysis.

Rank	Freq	Keyness	Effect	Keyword
1	2652	+ 4592.17	0.069	و (and)
2	1836	+ 4275.02	0.0485	تونس (Tunis)
3	1234	+ 2856.08	0.0328	الشعب (The people)
4	1106	+ 2143.7	0.0294	الرئيس (The president)
5	826	+ 1949.72	0.0221	معاك (With you)
6	814	+ 1906.63	0.0218	يا (O')
7	699	+ 1649.11	0.0187	تحيا (Long live)
8	606	+ 1429.17	0.0163	شكرا (Thank you)
9	611	+ 1415.56	0.0164	قيس (Kais)
10	469	+ 1105.47	0.0126	ربي (My God)
11	629	+ 1046.08	0.0169	الله (Allah)
12	415	+ 977.98	0.0112	سيدي (Mr.)
13	325	+ 765.61	0.0088	سيادة (His excellency)
14	536	+ 746.46	0.0144	سعيد (Said)
15	284	+ 668.92	0.0077	انقلاب (Coup D'etat)

Table 1: Keywords List

5. Discussion

The findings from this study offer valuable insights into the complex landscape of public opinion and sentiment in the aftermath of the July 25 measures taken by President Kais Saied in Tunisia. The analysis of Facebook comments from the official Tunisian presidency page reveals a multifaceted

discourse, with a predominant sentiment of support for the president's actions.

5.1 Sentiment Analysis

The sentiment analysis results clearly demonstrate that the majority of Tunisians who commented on the July 25 measures, specifically the suspension of parliament ordered by President Kais Saied, were in support of his decision. The stance analysis shows that the pro-decision group constitutes the majority, while those against the measures form a minority, even smaller than the neutral group. This finding provides readers with a clear understanding that the opposition to the measures was limited within the analyzed dataset. Furthermore, the sentiment analysis reveals a dominant score in the "Very Positive" category, followed by "Positive" and "Neutral," with "Negative" and "Very Negative" sentiments scoring significantly lower. This observation leads to the conclusion that the general sentiment surrounding the July 25 measures is highly positive, with those expressing negative sentiments being considered a minority.

5.2 Emotions and Feelings Analysis

The emotions and feelings analysis aligns with the sentiment findings, indicating that many Tunisians placed high trust in Saied's decision, as evidenced by the large number of comments scoring in the "Trust" value. The prominence of the "Joy" value also suggests that a significant portion of Tunisians were genuinely happy about these sudden decisions, which may explain the large crowds that took to the streets in the middle of the night to celebrate. Conversely, the values of "Anger" and "Fear" are far less prevalent in the analyzed comments, implying that individuals harboring these feelings were fewer in number. These latter values would primarily be associated with those who opposed the decision, further reinforcing the correlation between the emotions count and the stance count results.

5.3 Hate Speech, Aggressive Tone, and Racism Analysis

The analysis of hate speech, aggressive tone, and racism yields an intriguing finding. While the overall incidence of these negative expressions is lower, it is noteworthy that they are more frequent in comments from the pro-decision faction. Despite the corpus showing minimal hate speech or racism, with most comments expressing joy and trust towards the president's decisions, the pro-decision group, despite being the majority, displayed more hate speech than those against the measures. This observation suggests that alongside the positive feelings, the pro-decision group also harbored resentment towards other political entities.

5.4 AntConc Analysis

The AntConc findings further corroborate the sentiment and stance results, with the lexical analysis revealing a predominance of words conveying positive support for the president and the nation. The frequent occurrence of terms such as "Tunis," "President," "Saied," and "People" underscores this

sentiment. The analysis of collocates introduces a nuanced dichotomy reflective of Tunisia's past and present political landscape. Collocates associated with the term "Saied" carry positive connotations, indicating broad support and approval of President Saied's measures.

In contrast, the analysis of "Ennahdha" and "Brotherhood" reveals a marked hostility towards these entities. The term "خوانجية" (Brotherhood), often used derogatorily in Tunisian political discourse, frequently appears in collocates with negative sentiment. This term is unfavorably associated with the Ennahdha party and its affiliates, who typically reject this nomenclature. The name "Ennahdha" (نهضة), representing the country's most prominent Islamic political party, is widely recognized and used in a non-derogatory manner by the general population. The linguistic evidence suggests a pervasive disdain for this party, extending beyond the events of July 25 and reflecting longstanding political tensions. Ennahdha's significant role in Tunisia's political arena, coupled with its visibility and cohesive group identity, has evidently fueled the negative sentiment captured in the study's corpus. This trend of animosity, primarily directed at Ennahdha and its representatives, indicates a polarized political sentiment within the Tunisian populace.

5.5 Keyword and Collocates Analysis

The exploration of keywords and collocates reinforces the narrative of widespread support for President Saied and the measures taken on July 25. Keywords such as "president" (الرئيس), "his excellency" (سيادة), "the people" (الشعب), and "Tunis" highlight a positive sentiment towards the July 25 measures, focusing on support for Tunisia and President Saied. Positive collocates associated with "The President" suggest widespread approval. Conversely, the term "Coup," used by the opposition, ranks low in the keyword list, indicating that those against the measures are a minority. While some collocates of "Coup" show negative sentiment, others are neutral to positive, suggesting varied perceptions even among supporters, who use the term either critically or in support.

5.6 Implications and Future Directions

The findings of this study have significant implications for understanding the complex dynamics of public opinion in the context of Tunisia's ongoing political transformation. The overwhelming support for President Saied's measures, as expressed in the analyzed Facebook comments, suggests a strong public mandate for his actions and a desire for change in the country's political landscape. However, the presence of hate speech and aggressive tone, particularly among the pro-decision group, highlights the potential for polarization and the need for fostering a more inclusive and respectful public discourse.

Future research could explore the evolution of public sentiment over time, as the political situation in Tunisia continues to unfold. Longitudinal studies could provide valuable insights into how opinions and

emotions shift in response to specific events and policy decisions. Additionally, the incorporation of demographic data, if available, could help identify any differences in sentiment and stance across various segments of the Tunisian population.

Moreover, the linguistic analysis could be expanded to include a more comprehensive examination of the Tunisian dialect and its unique features, such as the use of Arabizi and code-switching. This would enable a more nuanced understanding of the language used in online political discourse and its potential impact on sentiment analysis and opinion mining.

6. Limitations

This study faced two primary limitations: temporal constraints and corpus size. The restricted timeframe necessitated limiting the sample to 5,000 annotated Facebook comments, constraining the breadth of analysis and precluding broad generalizations. Consequently, this research should be regarded as an initial, exploratory, small-scale study. Despite these limitations, the pilot nature of this work lays the groundwork for future, more expansive research endeavors. Importantly, the manual annotation performed in this study serves as a valuable precursor to the development of machine learning tools and algorithms tailored for detecting sentiment polarity and stance within Tunisian dialect corpora, thereby enhancing the methodological approach to analyzing this linguistic context. Moreover, due to the limitation of time and corpus size, the researcher's annotated notes only included 35 comments that were written in 'Arabizi', the Arabic chat alphabet, of the Tunisian dialect. Evidently, the number of comments was ostensibly small and therefore the researchers we not able to include it in the linguistic part of the data analysis. However, the 'Arabizi' comments were included in the stance and emotion data collection and analysis.

7. Conclusion

The research indicated a prevalent support among Tunisian Facebook users for the July 25 measures, with "Trust" and "Joy" being the predominant emotions expressed. Opposition to these measures was comparatively minor, with neutral stances more common than outright negative ones. Notably, the "pro" faction, while largely supportive, exhibited a greater tendency towards hate speech and aggressive tones, though these instances were relatively few. The prevailing sentiment among the comments was "Very Positive," reflecting a broad endorsement of the actions taken.

Linguistic analysis, including word lists and collocates, pointed to significant backing for President Saied, alongside notable criticism of the "Ennahdha" party, underscoring a clear political divide. This study highlights the necessity for more nuanced inquiries into the shifting sentiments of Tunisians regarding President Saied's policies, particularly through the lens of symbolic and emotive elements like emojis. The July 25 measures emerge as a crucial juncture in

Tunisia's political and democratic trajectory, meriting deeper examination of the public's reaction and its implications for the country's future.

The findings of this study contribute to the growing body of research on public opinion and sentiment analysis in the context of political events, particularly in the Middle East and North Africa region. The study's focus on Tunisia's July 25 measures provides valuable insights into the complex dynamics of public discourse and the role of social media in shaping political narratives. The prevalence of support for President Saied's actions, as evidenced by the analyzed Facebook comments, underscores the importance of understanding public sentiment in times of political upheaval.

However, the study also reveals the potential for polarization and the presence of hate speech and aggressive tones, even among supporters of the measures. This finding highlights the need for further research into the factors contributing to the spread of such negative sentiments and the development of strategies to promote a more inclusive and respectful public discourse.

8. Data Availability Statement

The annotated dataset can be obtained by contacting the authors to facilitate future research and reproducibility. The users of the dataset must adhere to the terms and conditions outlined in the repository. To request the dataset for research purposes, please fill the following form:

https://forms.gle/S9fZtYjAyLAqFsH19

The dataset is released under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, allowing for its free use, distribution, and adaptation, provided the original work is properly credited.

The data are not publicly available due to the sensitive nature of the comments and the potential for misuse or misinterpretation outside the context of this research. Access to the data will be granted to researchers. Requestors will be required to sign a data sharing agreement that specifies the conditions under which the data can be used, including measures to protect the privacy and confidentiality of the individuals whose comments are included in the dataset. The annotated dataset will be made available in a de-identified format, with any personally identifiable information removed to ensure the anonymity of the commenters.

Researchers interested in replicating or building upon the findings of this study are encouraged to contact the authors to discuss data access and collaboration opportunities.

9. Ethical Statement

This study was conducted in accordance with the ethical guidelines and regulations set forth by Hamad Bin Khalifa University. The study involved the analysis

of publicly available data from the official Facebook page of the Tunisian presidency.

As such, the research did not involve any direct interaction with human subjects and did not require informed consent from the individuals whose comments were included in the dataset. However, the researchers recognize the potential for harm and the need to protect the privacy and confidentiality of the commenters, even in the context of publicly available data.

- To mitigate potential risks and ensure the ethical conduct of the research, the following measures were taken:
- The data were collected and analyzed in an anonymous and de-identified format, with any personally identifiable information removed from the dataset.
- The researchers have taken steps to secure the data and prevent unauthorized access, including storing the data on encrypted drives and limiting access to authorized personnel only.
- The findings of the study are reported in aggregate form, without singling out or identifying any individual commenters.
- The researchers have strived to present the findings in a balanced and objective manner, avoiding any stigmatization or stereotyping of individuals or groups based on their opinions or political affiliations.
- The researchers are committed to the ethical and responsible conduct of research and have taken these measures to ensure that the study complies with the highest standards of academic integrity and human subjects protection. Any concerns or questions about the ethical aspects of this study should be directed to the authors.

10. Acknowledgments

11. References

Abbes, I., Zaghouani, W., El-Hardlo, O., & Ashour, F. (2020). DAICT: A dialectal Arabic irony corpus extracted from Twitter.

Alam, F., Mubarak, H., Zaghouani, W., Martino, G. D. S., & Nakov, P. (2022). Overview of the WANLP 2022 shared task on propaganda detection in Arabic.

Al-Ghadir, A., Azmi, A., & Hussain, A. (2021). A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. *Information Fusion, 67*, 29-40. https://doi.org/10.1016/j.inffus.2020.10.003

Al-Sulaiti, L., & Atwell, E. S. (2003). *The Design of a Corpus of Contemporary Arabic (CCA)*. The University of Leeds.

Aldayel, A., & Magdy, W. (2019). Your stance is exposed! Analyzing possible factors for stance detection on social media. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), Article 1. https://doi.org/10.1145/3359307

Allahoum, R. (2019, October 11). Tunisia's Kais Saied: 'He's just not interested in power'. *Aljazeera*. https://www.aljazeera.com/news/2019/10/11/tunisias-kais-saied-hes-just-not-interested-in-power

Baraniak, K., & Sydow, M. (2021). A dataset for sentiment analysis of entities in news headlines (SEN). *Procedia Computer Science, 192*, 3627-3636. https://doi.org/10.1016/j.procs.2021.09.136

Bianchi, R., Weber, A., Yyelland, B., Ghanam, R., Kittaneh, K., & Zaghouani, W. (2023). SYNCHRONOUS CONFERENCING SOFTWARE-ASSISTED TEACHING & LEARNING.

Biswas, M. R., Mohsen, F., Shah, Z., & Zaghouani, W. (2023). Potentials of ChatGPT for Annotating Vaccine Related Tweets.

Bothe, C., Weber, C., Magg, S., & Wermter, S. (2019). EDA: Enriching emotional dialogue acts using an ensemble of neural annotators. *arXiv preprint arXiv:1912.00819*.

Bouamor, H., Zaghouani, W., Diab, M., Obeid, O., Oflazer, K., Ghoneim, M., & Hawwari, A. (2016). On Arabic Multi-Genre Corpus Diacritization.

Crabb, J., Yang, S., & Zubova, A. (2019, December 10). Classifying hate speech: An overview. *Medium*. https://towardsdatascience.com/classifying-hate-speech-an-overview-d307356b9eba

Espace Manager. (2021) Sigma conseil: Kaïs Saïed caracole loin en tête avec 72,2% de confiance et 90% d'intentions de vote. *Espace Manager*.

https://www.espacemanager.com/si gma-conseil-kais-saied-caracole- loin-en-tete-avec-722-de-confiance- et-90-dintentions-de-vote.html.

Farha, I. A., Zaghouani, W., & Magdy, W. (2021). Overview of the WANLP 2021 shared task on sarcasm and sentiment detection in Arabic.

Gambino, O., Calvo, H., & García-Mendoza, C. (2018). Distribution of emotional reactions to news articles in Twitter. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*

(LREC 2018). https://doi.org/979-10-95546-00-9

Hardalov, M., Arora, A., Nakov, P., & Augenstein, I. (2021). A survey on stance detection for mis- and disinformation identification. *arXiv preprint arXiv:2103.00242.*

Hasan Saad, A., & Sarbini-Zin, M. L. (2022). Corpus-assisted analysis of Robert Frost's poem, "Into My Own" using AntConc. *PENDETA, 13*(1), 1-9.

Hawwari, A., Zaghouani, W., Diab, M., O Gorman, T., & Badran, A. (2016). Ampn: a semantic resource for Arabic morphological patterns.

Hendrickson, C., & Galston, W. (2017, April 28). Why are populists winning online? Social media reinforces their anti-establishment message. *Brookings.* https://www.brookings.edu/blog/techtank/2017/04/28/why-are-populists-winning-online-social-media-reinforces-their-anti-establishment-message/

Jungherr, A., Posegga, O., & An, J. (2021). Populist supporters on Reddit: A comparison of content and behavioral patterns within publics of supporters of Donald Trump and Hillary Clinton. *Social Science Computer Review, 40,* 089443932199613. https://doi.org/10.1177/089443932199613 0089443932199613.

Jurafsky, D. (1997). *Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual, draft 13* (Technical Report 97-01, pp. 225–233). University of Colorado Institute of Cognitive Science.

Aljazeera. (2019). Kais Saied wins Tunisia election with 72 percent. https://www.aljazeera.com/news/2019/10/14/tunisia-presidential-election-kais-saied-declared-winner

Tunisian Presidency. (2022). Tunisian presidency Facebook comments 25th of July measures. [Publisher unknown]. URL

Lai, M., Cignarella, A., Hernández Farías, D., Bosco, C., Patti, V., & Rosso, P. (2020). Multilingual stance detection in social media political debates. *Computer Speech & Language, 63,* 101075. https://doi.org/10.1016/j.csl.2020.101075

MacAvaney, S., Yao, H.-R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLoS ONE, 14*(8), e0221152. https://doi.org/10.1371/journal.pone.0221 152

Marzouk, H. (2021, August 17). Baromètre politique aout 2021 Tunisiens soutiennent mesures exceptionnelles 25 Juillet. *L'Economiste Maghrébin.* https://www.leconomistemaghrebin.com/2021/08/17/barometre-politique-aout-2021-949-tunisiens-soutiennent-mesures-exceptionnelles-25-juillet/

Mhamdi, C., Al-Emran, M., & Salloum, S. A. (2018). Text mining and analytics: A case study from news channels posts on Facebook. In *Intelligent Natural Language Processing: Trends and Applications* (pp. 399-415). Springer.

Mutlu, E., Oghaz, T., Jasser, J., Tutunculer, E., Rajabi, A., Tayebi, A., Ozmen, O., & Garibay, I. (2020). A stance data set on polarized conversations on Twitter about the efficacy of hydroxychloroquine as a treatment for COVID-19. *Data in Brief, 33,* 1-11. https://doi.org/10.1016/j.dib.2020.106401

Obeid, O., Bouamor, H., Zaghouani, W., Ghoneim, M., Hawwari, A., Alqahtani, S., Diab, M., & Oflazer, K. (2016). Mandiac: A web-based annotation system for manual Arabic diacritization.

Saif, H., & Alani, H. (2012). Semantic sentiment analysis of Twitter. In *The Semantic Web--ISWC 2012* (Vol. 7649, pp. 508-524). https://doi.org/10.1007/978-3-642-35176-1_32

Sandoval-Almazan, R., & Valle-Cruz, D. (2018). Facebook impact and sentiment analysis on political campaigns. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1-7)

Shems FM. (n.d.). Sondage: 87,3% des Tunisiens approuvent les décisions de Kais Saied, annoncées le 25 juillet. https://www.shemsfm.net/amp/fr/actualites_tunisie-news_news-nationales/309135/sondage-87-3-des-tunisiens-approuvent-les-decisions-de-kais-saied-annoncees-le-25-juillet.

Shurafa, C., Darwish, K., & Zaghouani, W. (2020). Political framing: US COVID19 blame game.

Statista. (2022). Tunisia: Share of social media users by platform. https://www.statista.com/statistics/1196465/share-of-active-social-media-users-by-platform-tunisia/.

Vashisht, G., & Thakur, S. (2014). Facebook as a corpus for emoticons-based sentiment analysis. *International Journal of Emerging Technology and Advanced Engineering, 4,* 904-908.

Wang, J., & Chen, K. (2021). Aggregating user-centric and post-centric sentiments from social media for topical stance prediction. *ROCLING.*

Williams, L., Arribas-Ayllon, M., Artemiou, A., et al. (2019)

Comparing the utility of different classification schemes for emotive language analysis. *Journal of Classification, 36,* 619–648. https://doi.org/10.1007/s00357-019-9307-0

31

Wolf, A. (2019). In search of 'consensus': The crisis of party politics in Tunisia. *The Journal of North African Studies, 24*(6), 883-886. https://doi.org/10.1080/13629387.2019.1675249

Zamani, N. A. M., Abidin, S. Z., Omar, N. A. S. I. R. O. H., & Abiden, M. Z. Z. (2013). Sentiment analysis: Determining people's emotions in Facebook. In *Proceedings of the 13th International Conference on Applied Computer and Applied Computational Science* (pp. 111- 116).

Zayani, M., & Downing, J. (2015). *Networked Publics and Digital Contention* (1st ed.). Georgetown University's Center of International and Regional Studies, School of Foreign Services in Qatar.

Shems FM. (2021). Sondage: 87,3% des tunisiens approuvent les décisions de Kais Saied, annoncées le 25 juillet. https://www.shemsfm.net/fr

Pak, A., & Paroubek, P. (2021, May). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*.

Baj-Rogowska, A. (2017, December). Sentiment analysis of Facebook posts: The Uber case. In *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)* (pp. 391-395). IEEE.

Abercrombie, G., & Batista-Navarro, R. T. (2020, May). ParlVote: A corpus for sentiment analysis of political debates. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 5073-5078).

Yerkes, S., & Mbarek, N. (2021, January 14). After ten years of progress, how far has Tunisia really come? Carnegie Endowment for International Peace. https://carnegieendowment.org/2021/01/14/after-ten-years-of-progress-how-far-has-tunisia-really-come-pub-83609

Zaghouani, W., & Awad, D. (2016a). Building an Arabic Punctuated Corpus.

Zaghouani, W., & Awad, D. (2016b). Toward an Arabic punctuated corpus: Annotation guidelines and evaluation.

Zaghouani, W., & Charfi, A. (2018a). Arap-tweet: A large multi-dialect twitter corpus for gender, age, and language variety identification.

Zaghouani, W., Diab, M., Mansouri, A., Pradhan, S., & Palmer, M. (2010). The revised Arabic propbank.

Zaghouani, W., Habash, N., Obeid, O., Mohit, B., Bouamor, H., & Oflazer, K. (2016). Annotation Guidelines and Framework for Arabic Machine Translation Post-Edited Corpus.