

Analyzing the Impact of Emotional Content in Tweets on User Engagement and Online Behavior

Soufien Jaffali

Department of Management Information Systems, College of Business and Economics,
Qassim University

P.O. Box: 6640, Buraidah: 51452 • Kingdom of Saudi Arabia

s.jaffali@qu.edu.sa

Abstract:

The widespread adoption of social media platforms has provided a valuable source of data for understanding public sentiment and engagement towards various topics, including healthcare and vaccination. In this paper, we analyze the impact of sentiment in tweets related to COVID-19 vaccination in Saudi Arabia on user engagement metrics, such as likes, replies, and retweets. We used web crawling techniques and applied sentiment analysis using BERT (Bidirectional Encoder Representations from Transformers) transformer models. Descriptive analysis of the data revealed the distribution of sentiment labels and user engagement metrics. We found that positive sentiment was associated with higher average likes and replies, while neutral sentiment showed a lower engagement level. Furthermore, we conducted statistical tests, including ANOVA (Analysis of Variance), to determine significant differences in user engagement based on sentiment labels. The results indicated a significant difference in likes count among different sentiment categories. This study provides insights into tailoring messaging strategies, monitoring sentiment trends, and collaborating with influencers to enhance the success of vaccination campaigns. The results emphasize the need for incorporating sentiment analysis into public health strategies to effectively engage the population and promote vaccination acceptance.

Keywords: Social Media Analysis, Sentiment Analysis, Vaccine Acceptance, Psychological Factors.

JEL Classification codes: I18; C83; C88

1. Introduction

The COVID-19 pandemic has presented an unprecedented global challenge, requiring collective efforts to mitigate the spread of the virus and protect public health. Vaccination has emerged as a critical strategy in combating the virus and curbing the impact of the pandemic. As vaccination campaigns continue to be rolled out worldwide, understanding public sentiment and engagement towards COVID-19 vaccination is crucial for effective communication, outreach, and intervention in public health campaigns.

Social media platforms, such as Twitter, have become prominent channels for expressing opinions, sharing information, and engaging in discussions related to various topics, including COVID-19 and vaccination. These platforms provide valuable data that can be leveraged to gain insights into public sentiment, concerns, and perceptions surrounding vaccination efforts.

In this study, we focus on analyzing tweets related to COVID-19 vaccination in Saudi Arabia, a country at the forefront of early vaccine distribution efforts. Recognizing the importance of widespread immunization, Saudi Arabia swiftly implemented a comprehensive vaccination program, making it one of the first countries to provide vaccines for its citizens and residents. As such, examining public sentiment and engagement in Saudi Arabia becomes particularly relevant given the country's proactive approach to vaccination.

To achieve this, we employ state-of-the-art natural language processing techniques, with a particular focus on utilizing the BERT transformer model for sentiment analysis. BERT has demonstrated remarkable capabilities in understanding contextual information and capturing subtle nuances of language, making it an ideal choice for sentiment analysis in social media data.

Through sentiment analysis, we categorize the tweets into positive, negative, or neutral sentiments, allowing us to gauge the overall public perception towards COVID-19 vaccination in Saudi Arabia. Furthermore, we delve deeper into the emotional content of the tweets, examining the specific sentiments expressed and the underlying factors contributing to these sentiments.

Additionally, we analyze user engagement metrics, including retweets, likes, and replies, to understand the level of public interest and involvement in the vaccination discourse. By exploring the relationship between sentiment and user engagement, we aim to uncover patterns that shed light on the impact of emotional content in tweets on user behavior and engagement.

The findings of this study have the potential to provide valuable insights for public health authorities, policymakers, and communication strategists in Saudi Arabia, enabling them to tailor their messaging and interventions based on a comprehensive understanding of public sentiment and engagement. Ultimately, this research contributes to the larger goal of fostering informed decision-making and promoting effective public health campaigns.

2. Literature Review

User engagement on social media platforms has become a crucial area of research, with studies exploring various factors that influence user interactions. One such area of

interest is the impact of emotional content on user engagement. [1] explores the factors influencing the sharing of political content on social media platforms. The study investigates how emotional appeal, perceived authority, and ideological alignment impact users' decisions to share political posts. The findings reveal that emotional content, particularly those evoking positive emotions, tends to be shared more frequently.

During the COVID-19 pandemic, digital interventions emerged to address the psychological impact on healthcare workers. [6] explored the effectiveness of digital interventions in providing psychological support to frontline workers. The study highlighted the potential of technology to deliver targeted interventions and address the well-being of healthcare workers during public health crises.

Moreover, artificial intelligence has gained prominence in healthcare, particularly in the personalization of conversational agents. [7] delved into the use of AI-driven agents to enhance user engagement and improve patient outcomes. The study emphasized how personalization can create meaningful interactions in healthcare settings, paving the way for more patient-centric care.

Social media engagement has played a significant role during public health crises, such as the COVID-19 pandemic. [8] explored citizen engagement patterns across different stages of the pandemic. The research revealed variations in user engagement behaviors and interactions with government accounts, underscoring the need for tailored communication strategies during public health emergencies.

The use of emojis in brand-related user-generated content has also been a subject of interest. [9] investigated the impact of emotive visual cues on user interactions. The study highlighted how emojis can effectively capture user attention and foster engagement with brand-related content, offering practical implications for marketers to enhance brand communication strategies.

During the COVID-19 vaccination rollout, sentiment analysis played a crucial role in understanding public perceptions. [10] demonstrated how advanced analytical techniques provided insights into vaccination sentiment across different locations and time periods. This information can inform public health agencies in developing targeted communication strategies to address public concerns and increase vaccine acceptance.

Addressing vaccine hesitancy is another critical issue in public health. [11] explored social media strategies to promote vaccine acceptance. By analyzing behavior change theories and social media interventions, the study offered insights into effective communication approaches for countering hesitancy. These findings are valuable for policymakers in designing targeted interventions to bolster vaccine confidence.

In conclusion, the literature review underscores the significance of analyzing emotional content's impact on user engagement and online behavior, particularly in the context of public health campaigns and vaccination efforts. Recent studies confirm the influential role of emotional content in capturing user attention and driving engagement metrics. Leveraging advanced sentiment analysis techniques enables a deeper understanding of emotional content's nuances and informs the development of persuasive communication strategies. Nevertheless, further research is warranted to comprehensively explore the intricate interplay between emotional content, sentiment, user characteristics, and

platform dynamics to optimize the effectiveness of public health campaigns and vaccination efforts in the digital age.

3. Data Collection and Preprocessing

The dataset used in this study was collected through the TWINT API, a tool for crawling tweets from the social media platform Twitter. The dataset was specifically configured to focus on tweets related to COVID-19 vaccination in Saudi Arabia. Several settings and criteria were defined to gather relevant data.

3.1 Data Collection Settings

- **Geographic Focus:** The data collection was limited to tweets originating from Saudi Arabian cities, indicating a specific geographical scope targeting discussions within Saudi Arabia.
- **Temporal Range:** Start Date: The data collection began on the 17th of December 2020. This date corresponds to a significant event in Saudi Arabia, which is the initiation of the COVID-19 vaccination program using the Pfizer-BioNTech vaccine. This starting point was chosen to capture tweets related to the rollout of the vaccination program.
- **End Date:** The data collection concluded on the 31st of December 2022, providing an extensive timeframe to capture tweets over two years, allowing for a comprehensive analysis of evolving discussions related to COVID-19 vaccination.
- **Keyword Filters:** The dataset was filtered to include tweets containing specific keywords and phrases related to COVID-19 vaccination. These keywords encompassed various aspects of vaccination, including vaccine names (e.g., لقاح , استرازينيكا, جونسون (موديرنا, فايزر), vaccination-related terms (e.g., ضد اللقاح , 'antivaccin'), and broader COVID-19 and vaccination hashtags (e.g., 'COVID19', 'vaccine', 'COVIDvaccine').

This dataset compilation strategy aimed to create a focused and relevant collection of tweets centered around COVID-19 vaccination discussions in Saudi Arabia. It provides a valuable resource for analyzing public sentiment, perceptions, and conversations regarding vaccination in the specified geographic region and time frame.

A total of 96 files were collected, with each file containing tweets relevant to a specific Saudi city. In total, the dataset comprises 29,128 tweets, providing a comprehensive snapshot of public discourse surrounding vaccination efforts in the country during the COVID-19 pandemic.

The dataset comprises 36 columns, each representing various attributes of the tweets. These attributes include conversation_id, created_at, date, time, timezone, user_id, username, name, place, tweet, language, mentions, URLs, photos, replies_count, retweets_count, likes_count, hashtags, cashtags, link, retweet, quote_url, video, thumbnail, near, geo, source, user_rt_id, user_rt, retweet_id, reply_to, retweet_date, translate, trans_src, and trans_dest. Each column provides valuable information about the tweet, its content, and associated metadata.

The dataset, consisting of tweets in 40 different languages distributed as shown in Figure 1, offers valuable insights into the sentiments, opinions, and engagement patterns of Twitter users regarding vaccination in Saudi Arabia. Researchers interested in accessing the collected dataset for further analysis can find it at the following link¹.

¹ Dataset Repository:

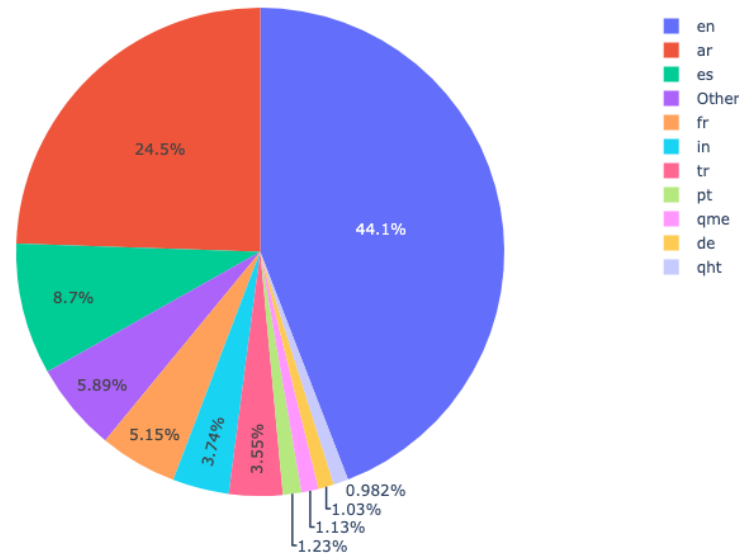


Figure 1: Language Distribution of Tweets

3.2 Data Preprocessing

In order to ensure the quality and relevance of the dataset, a series of preprocessing steps were applied to filter and clean the tweets. The initial dataset contained tweets in approximately 40 languages; however, for this analysis, only tweets in Arabic and English, the most frequent languages, were retained.

3.2.1 Language Filtering

The original dataset encompassed tweets in a multitude of languages (over 40 languages). To ensure the dataset's coherence and relevance, a meticulous language filtering process was enacted. Arabic and English, as the dominant languages, constituted approximately 70% of the dataset. Hence, particular emphasis was placed on these languages, aligning with their prevalence and significance in the context of this study.

3.2.2 Cleaning Tweets

For both Arabic and English tweets, a consistent set of preprocessing steps were applied:

- **Removal of Hashtags:** All hashtags (words or phrases preceded by '#') were removed from the tweets. Hashtags often contain non-textual characters or context-specific information that can introduce noise into the analysis.
- **Elimination of HTTP Links:** Any hyperlinks (starting with 'http' or 'https') within the tweets were removed. These links are typically not informative for sentiment analysis and can be considered as noise.

- **Special Symbol Removal:** Special symbols and characters that do not contribute to the semantic content of the tweets were stripped away. This included punctuation, emoticons, and other non-alphanumeric characters.
- **Letter Repetition:** Repeated letters within words were normalized. For instance, 'للالاااااااااا' was transformed into 'للاح'. This step helps in standardizing the text and reducing the impact of text variations.
- **Lemmatization:** A lemmatization process was applied to reduce words to their base or dictionary form. This step helps in standardizing the language and reducing dimensionality.
- **Stopword Removal:** Arabic stopwords, which are common words like articles, prepositions, and pronouns, were removed from the tweets. These words are often frequent but do not carry substantial semantic information for sentiment analysis.

By applying these preprocessing steps, the dataset was refined to include clean and relevant Arabic tweets, ready for subsequent analysis. This preprocessing ensures that the data used for sentiment analysis is of high quality and free from unnecessary noise.

4. BERT Transformer for Sentiment Analysis

4.1 Overview of Natural Language Processing Techniques for Emotional Content Analysis

In the field of natural language processing (NLP), analyzing emotional content in text has gained prominence for its applications in understanding public sentiment and emotional responses in social media data. Traditionally, sentiment analysis primarily focused on classifying text into sentiment categories such as positive, negative, or neutral. However, the evolution of NLP techniques has paved the way for more sophisticated approaches that capture the nuances of emotions and sentiments within text.

One such breakthrough in NLP is the Bidirectional Encoder Representations from Transformers (BERT) transformer model. BERT is a powerful deep learning architecture that excels in understanding the contextual meaning of words and phrases within a sentence. It leverages a bidirectional attention mechanism to consider the entire context, both left and right, when processing each word.

4.2 BERT's Application for Sentiment Analysis

BERT's ability to capture context makes it highly suitable for sentiment analysis tasks. Unlike traditional models, BERT can discern the sentiment of a word or phrase based on its surrounding context, allowing for more accurate sentiment classification. This contextual understanding enables BERT to handle negations, sarcasm, and other complex language phenomena that pose challenges for traditional sentiment analysis models.

In sentiment analysis, BERT is often fine-tuned on a specific sentiment analysis dataset, adapting its pre-trained knowledge to the nuances of the sentiment classification task at hand. This fine-tuning process optimizes BERT's performance for sentiment prediction.

4.3 Utilizing BERT for Tweets Sentiment Analysis

In this study, we harnessed the power of BERT for sentiment analysis of tweets. The dataset, comprising tweets related to COVID-19 vaccination in Saudi Arabia, presented a diverse range of language styles, including colloquial expressions, abbreviations, and emotional content. BERT's contextual understanding allowed us to effectively capture the sentiment of these tweets.

Our approach involved fine-tuning a pre-trained BERT model on a labeled sentiment dataset specific to our research domain. By fine-tuning BERT on this dataset, we tailored its sentiment analysis capabilities to the unique characteristics of tweets discussing vaccination.

Subsequently, we applied the fine-tuned BERT model to the tweets in our dataset. This enabled us to classify each tweet into sentiment categories, such as positive, negative, or neutral, based on the contextual information present in the tweet's content. **Table 1** presents some examples of tweets in the collected dataset classified using BERT into negative, positive or neutral.

Table 1: Examples of tweets from each sentiment category

Sentiment Label	Example of tweets
Negative	<ul style="list-style-type: none"> End the year right: Mask up and stay safe. For your loved ones, friends and community. Are you (or someone you know) on the fence about getting vaccinated or getting your booster shot? It's no secret where @..... is... he's in an ICU on a ventilator after catching #COVID19! #VaccinesSaveLives #MasksWork
Neutral	<ul style="list-style-type: none"> First dose: Pfizer. Second dose: AstraZeneca. Third dose: Pfizer According to what I read, the third dose is only Pfizer or Moderna
Positive	<ul style="list-style-type: none"> So thankful my kids are vaccinated, masking, distancing, and washing their hands! The best way to end the year! I'm so thankful to have been able to receive my COVID-19 booster shot today! Biggest thank you to our incredible healthcare workers for making this happen.

By employing BERT for sentiment analysis, we aimed to provide a nuanced understanding of the emotional content within tweets, allowing us to uncover insights into public sentiment surrounding COVID-19 vaccination efforts in Saudi Arabia.

5. Results and Analysis

5.1 Sentiment Distribution

In order to gain insights into the sentiment and engagement patterns of tweets related to COVID-19 vaccination in Saudi Arabia, a descriptive analysis was conducted. The sentiment distribution of the collected tweets, presented in **Figure 2**, revealed that the majority of tweets were categorized as negative (12,139 tweets), followed by positive (4,501 tweets) and neutral (3,340 tweets). Furthermore, engagement statistics were analyzed, including likes, replies, and retweets. **Table 2** provides an overview of the engagement metrics, showing the count, mean, standard deviation, and other key measures. It can be observed that the average number of likes per tweet was 5.85, with a standard deviation of 36.78, while the average number of replies and retweets were

0.81 and 1.05, respectively. The maximum values for likes, replies, and retweets were 1,019, 77, and 341, respectively, indicating a wide range of engagement levels across the dataset.

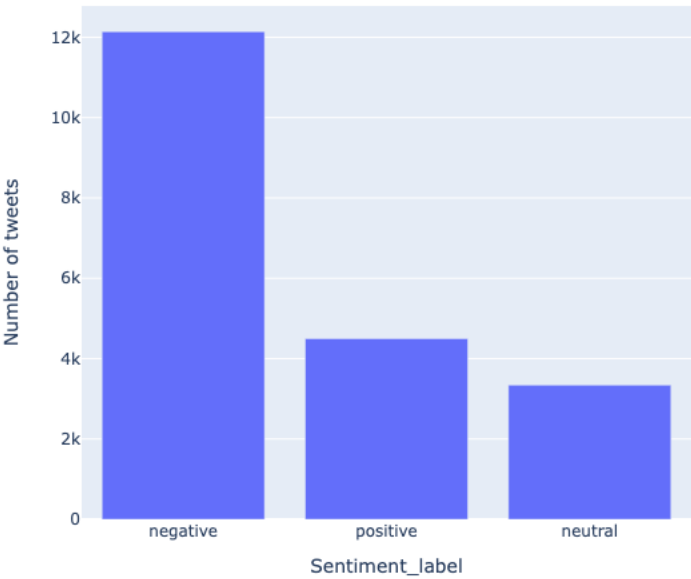


Figure 2: Sentiment Distribution

Table 2: Engagement Statistics

Measure	Likes	Replies	Retweets
Count	19,980	19,980	19,980
Mean	5.85	0.81	1.05
Std. Deviation	36.78	2.67	7.49
Minimum	0	0	0
25th Percentile	0	0	0
Median	0	0	0
75th Percentile	2	1	0
Maximum	1,019	77	341

5.2 Relationship Between User Engagement and Sentiments

The analysis of the relationship between sentiment labels and user engagement metrics revealed interesting findings (Table 3). The average number of retweet varied across different sentiment labels, with negative sentiment tweets receiving the highest average retweets (1.39), followed by positive sentiment tweets (0.54), and neutral sentiment tweets (0.52). However, when it comes to likes and replies, the pattern was slightly different. Positive sentiment tweets had the highest average number of likes (6.66) and replies (0.89), followed by negative sentiment tweets having an average number of likes (6.54) and replies (0.8). Neutral sentiment tweets showed a lower average number of likes (2.24) and relatively lower average number of replies (0.75). The observation that tweets expressing neutral sentiment are associated with lower levels of engagement raises intriguing questions about the underlying dynamics of user

interaction on social media platforms. While tweets conveying strong positive or negative emotions often garner more attention and interaction, the relatively muted response to neutral content suggests a potential gap in user engagement strategies. Further exploration into the reasons behind this phenomenon could provide valuable insights for communication practitioners. Possible factors contributing to the lower engagement with neutral sentiment tweets may include a lack of emotional resonance, diminished user interest, or the perceived informational value of the content.

Sentiment Label	Average Likes	Average Replies	Average Retweets
Negative	6.54	0.80	1.39
Neutral	2.24	0.75	0.52
Positive	6.66	0.89	0.54

Table 3: Average User Engagement Metrics by Sentiment Label

To determine the statistical significance of these differences, an ANOVA test was conducted for the 'likes-count' metric. The results of the test revealed a highly significant difference in average likes across the sentiment labels (ANOVA p-value < 0.001), indicating that the variation in likes is not due to chance alone. This suggests that the sentiment expressed in tweets plays a role in attracting engagement from users, particularly in terms of likes.

These findings highlight the importance of considering sentiment in understanding user engagement on social media platforms, such as Twitter. The results suggest that positive sentiment tweets tend to receive more likes and replies, while negative sentiment tweets attract more retweets. These insights contribute to our understanding of how sentiment influences user engagement and can inform strategies for optimizing engagement levels in the context of COVID-19 vaccination discussions on social media platforms.

5.3 User Engagement Trends

The graphs depicting the number of likes, retweets, and replies over time provides valuable insights into user engagement with the subject matter (**Figures 1 - 4**). Upon analyzing these graphs, it is evident that user engagement has shown an increasing trend over time. This finding suggests a growing interest and involvement of Twitter users in the topic of vaccination in Saudi Arabia.

Additionally, the **Figure 1** reveals an interesting pattern regarding the distribution of user interactions during the initial days of vaccination distribution. At the outset, the numbers of replies, retweets, and likes were relatively close in magnitude, indicating a balanced level of engagement across these metrics. However, as time progressed, a notable gap started to emerge between the metrics. This observation implies a potential shift in user behavior and preferences as the vaccination campaign progressed.

The widening gap between the numbers of replies, retweets, and likes may signify a changing dynamic in how Twitter users engage with vaccination-related content. It could indicate a shift towards more passive engagement, where users are more likely to retweet or like the content rather than actively participate in discussions through replies. This trend could be influenced by various factors, such as the evolution of the public

discourse on vaccination, the availability of reliable information, or changing attitudes towards the topic.

Overall, the analysis of these graphs highlights the increasing user engagement over time and the evolving patterns of user interactions. These insights can be valuable for public health campaigns and vaccination efforts in Saudi Arabia, providing a deeper understanding of how user engagement evolves and indicating areas where targeted interventions or strategies may be needed to maintain and enhance user engagement throughout the vaccination campaign.

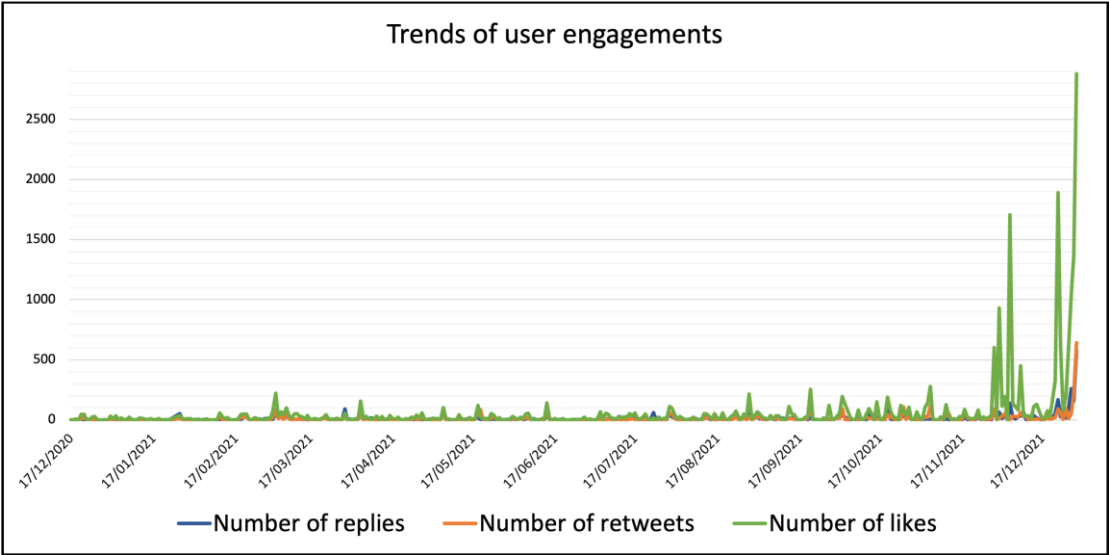


Figure 2: Trends of user engagements

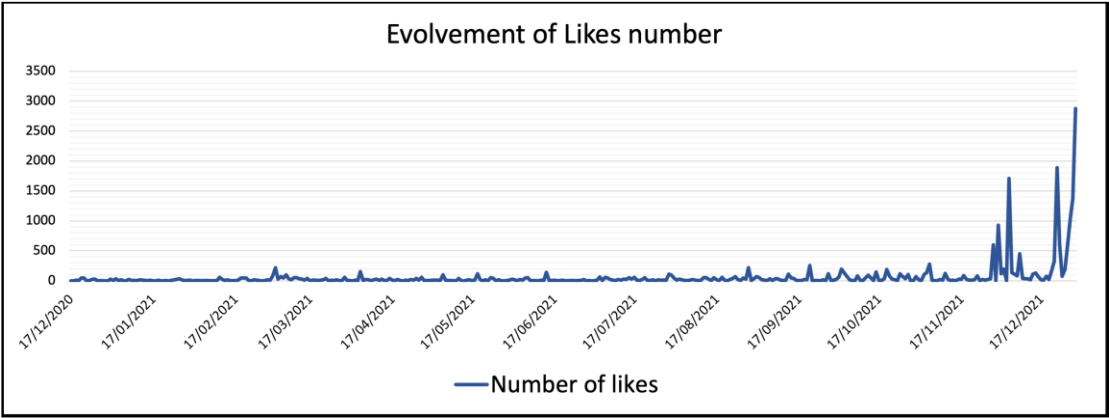


Figure 3: Evolution of "Likes" number

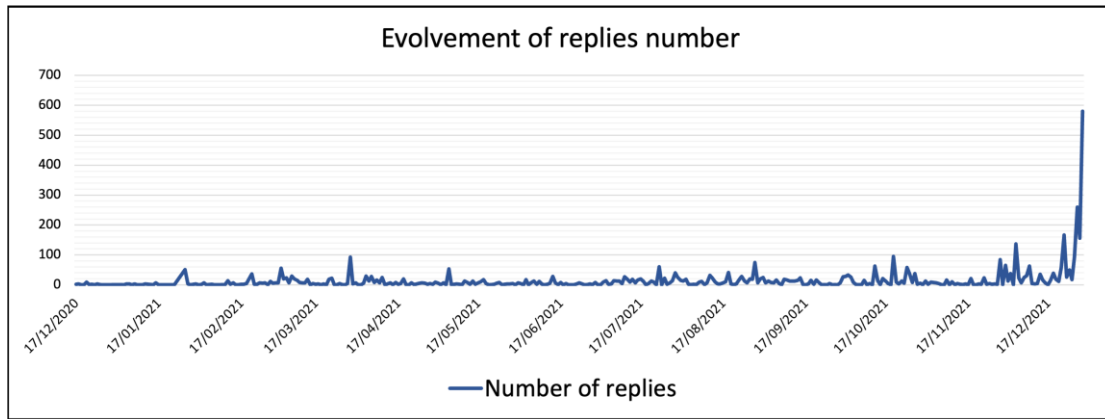


Figure 4: Evolution of "Replies" number

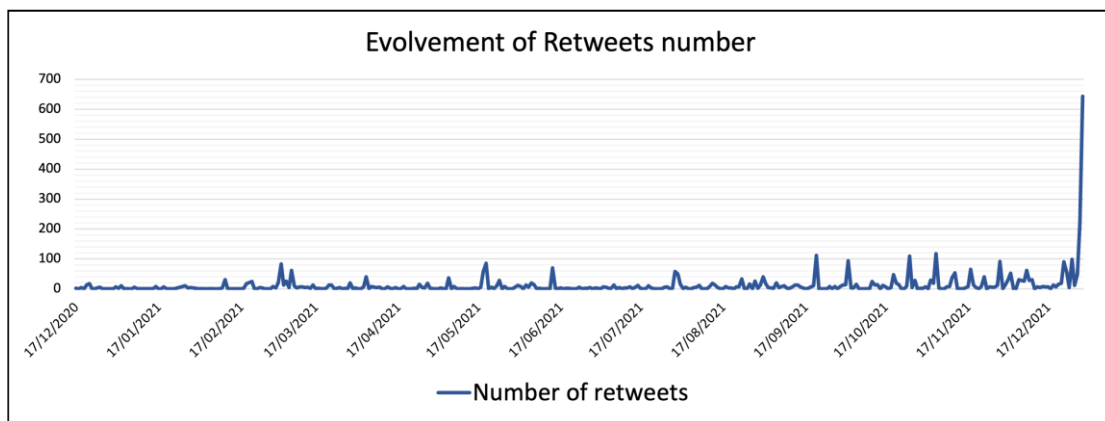


Figure 5: Evolution of "Retweets" number

5.4 Interpretation

The findings of this analysis have significant implications for public health campaigns and vaccination efforts in Saudi Arabia. The following interpretations can be drawn from the results:

1. **Understanding sentiment and engagement:** The analysis provides insights into the relationship between sentiment and user engagement metrics, such as likes, replies, and retweets. This understanding is crucial for shaping effective public health campaigns. Positive sentiment tends to be associated with higher engagement, indicating that creating a positive narrative around vaccination can significantly impact participation rates.
2. **Tailored messaging:** The results suggest the need for tailored messaging strategies based on sentiment categories. Different sentiment categories (positive, negative, neutral) can elicit varying levels of engagement. Public health campaigns should address concerns associated with negative sentiment, provide accurate information, and focus on amplifying positive experiences to foster a supportive and encouraging environment.
3. **Monitoring sentiment:** Continuous sentiment analysis can serve as a valuable monitoring tool for public health campaigns. By regularly assessing sentiment trends, organizations can evaluate the effectiveness of their messaging and adapt

their strategies accordingly. This proactive approach allows for timely adjustments and targeted interventions to address emerging sentiments and concerns.

4. **Collaboration and communication:** The correlation between sentiment and engagement highlights the importance of collaboration between public health authorities, healthcare providers, and influential individuals or organizations. Engaging with community influencers who can help spread positive sentiment and encourage engagement can significantly enhance the success of vaccination campaigns [12]. Social media platforms offer valuable tools for public health campaigns, providing a framework for effective communication and collaboration [13]. By leveraging best practices and practical approaches for public health campaigns, stakeholders can maximize their impact and reach a wider audience.
5. **Implications for public health initiatives:** The findings suggest the significance of incorporating sentiment analysis into public health strategies. By understanding the sentiments expressed by the target audience, public health campaigns can craft messages that resonate with individuals, address their concerns, and foster a positive attitude towards vaccination.

In conclusion, these interpretations emphasize the importance of considering sentiment and user engagement in public health campaigns and vaccination efforts. By tailoring messaging, addressing concerns, and fostering positive sentiment, public health initiatives in Saudi Arabia can effectively engage the population and encourage broader participation in vaccination programs.

5.5 Limitations

While this study offers valuable insights into the relationship between sentiment and user engagement metrics in the context of COVID-19 vaccination discussions on Twitter, several limitations should be noted:

- **Data Source and Representativeness:** The dataset used in this study comprises tweets collected from Twitter, which may not fully represent the sentiments and opinions of the entire population. Moreover, the analysis focused primarily on tweets from Saudi Arabia, limiting the generalizability of the findings to other geographical regions or demographic groups.
- **Language and Cultural Variations:** The analysis was conducted primarily on tweets in Arabic and English, overlooking sentiments expressed in other languages prevalent in the region. Furthermore, variations in language nuances, colloquial expressions, and cultural contexts were not fully accounted for, potentially influencing sentiment analysis results.
- **Scope of User Engagement Metrics:** While this study examined user engagement metrics such as likes, retweets, and replies, other important metrics such as shares or mentions were not included due to limitations in data availability. Future research could benefit from incorporating a broader range of engagement metrics to provide a more comprehensive understanding of user behavior on social media platforms.
- **Inherent Biases in Social Media Data:** Social media data inherently contain biases due to factors such as user demographics, platform algorithms, and the nature of

online discourse. As a result, the sentiments expressed on Twitter may not always accurately reflect real-world attitudes and behaviors towards COVID-19 vaccination.

- **Interpretation of Sentiment Analysis Results:** While sentiment analysis provides valuable insights, it may not capture the full complexity of human emotions and intentions expressed in textual data. The interpretation of sentiment labels (positive, negative, neutral) is subjective and context-dependent, requiring careful consideration and validation.

Acknowledging these limitations is crucial for interpreting the study findings accurately and guiding future research efforts to address gaps and refine methodologies for analyzing sentiment and user engagement on social media platforms.

6. Conclusion

In conclusion, this study utilized advanced natural language processing techniques, specifically the BERT transformer model, to examine sentiments within tweets concerning COVID-19 vaccination in Saudi Arabia. Our analysis revealed a diverse emotional landscape, with a significant prevalence of negative sentiments. Surprisingly, negative tweets elicited higher user engagement, underscoring the nuanced nature of public discourse. Moreover, we observed a growing trend in user engagement over time, emphasizing the evolving dynamics of public sentiment and the importance of timely communication strategies.

From the study's findings, several key actionable recommendations emerge. Firstly, tailored public health communication strategies should address prevalent negative sentiments surrounding COVID-19 vaccination, emphasizing accurate information dissemination and addressing concerns effectively. Secondly, collaborative efforts involving public health authorities, healthcare providers, and influential individuals or organizations are essential for amplifying positive sentiment and encouraging engagement, as highlighted by our findings. Moreover, continuous sentiment monitoring using natural language processing techniques can serve as a valuable tool for public health campaigns, enabling timely adjustments and targeted interventions to address emerging sentiments and concerns.

Looking ahead, future research endeavors should explore the drivers behind user engagement with negative sentiments, providing valuable insights for tailored communication strategies. Additionally, longitudinal studies tracking sentiment and engagement trends over extended periods could yield a deeper understanding of evolving public perceptions. Comparative analyses across different regions, vaccination campaigns, or public health crises could further elucidate variations in sentiment and engagement dynamics.

References

- [1] Shahbaznezhad, H., Dolan, R. & Rashidirad, M. The role of social media content format and platform in users' engagement behavior. *Journal Of Interactive Marketing*. 53, 47-65 (2021)
- [2] Ye, J., Hai, J., Wang, Z., Wei, C. & Song, J. Leveraging natural language processing and geospatial time series model to analyze COVID-19 vaccination sentiment dynamics on Tweets. *JAMIA Open*. 6, ooad023 (2023)
- [3] Li, L., Wood, C. & Kostkova, P. Vaccine hesitancy and behavior change theory-based social media interventions: a systematic review. *Translational Behavioral Medicine*. 12, 243-272 (2022)
- [4] Ko, E., Kim, D. & Kim, G. Influence of emojis on user engagement in brand-related user generated content. *Computers In Human Behavior*. 136 pp. 107387 (2022)
- [5] Weismueller, J., Harrigan, P., Coussement, K. & Tessitore, T. What makes people share political content on social media? The role of emotion, authority and ideology. *Computers In Human Behavior*. 129 pp. 107150 (2022)
- [6] Blake, H., Bermingham, F., Johnson, G. & Tabner, A. Mitigating the psychological impact of COVID-19 on healthcare workers: a digital learning package. *International Journal Of Environmental Research And Public Health*. 17, 2997 (2020)
- [7] Kocaballi, A., Berkovsky, S., Quiroz, J., Laranjo, L., Tong, H., Rezazadegan, D., Briatore, A. & Coiera, E. The personalization of conversational agents 12 in health care: systematic review. *Journal Of Medical Internet Research*. 21, e15360 (2019)
- [8] Zhang, W., Yuan, H., Zhu, C., Chen, Q. & Evans, R. Does citizen engagement with government social media accounts differ during the different stages of public health crises? An empirical examination of the COVID-19 pandemic. *Frontiers In Public Health*. 10 pp. 807459 (2022)
- [9] Ko, E., Kim, D. & Kim, G. Influence of emojis on user engagement in brand-related user generated content. *Computers In Human Behavior*. 136 pp. 107387 (2022)
- [10] Ye, J., Hai, J., Wang, Z., Wei, C. & Song, J. Leveraging natural language processing and geospatial time series model to analyze COVID-19 vaccination sentiment dynamics on Tweets. *JAMIA Open*. 6, ooad023 (2023)
- [11] Li, L., Wood, C. & Kostkova, P. Vaccine hesitancy and behavior change theory-based social media interventions: a systematic review. *Translational Behavioral Medicine*. 12, 243-272 (2022)
- [12] Center for Tobacco Products. (2024, February 5). Influencers 101: Best Practices for Public Health Campaigns. Retrieved from <https://digitalmedia.hhs.gov/tobacco/hosted/CTP-Influencers-101-JUN23.pdf>.
- [13] De Vere Hunt I, Linos E. Social Media for Public Health: Framework for Social Media-Based Public Health Campaigns. *J Med Internet Res*. 2022. Dec 14;24(12): e42179. (2022)