Received: 9 January 2024 | Revised: 12 March 2024 | Accepted: 28 March 2024 | Published online: 29 March 2024

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2025, Vol. 3(3) 165–175
DOI: 10.47852/bonviewJDSIS42022441

Sentiment Analysis of Twitter Discourse on Omicron Vaccination in the USA Using VADER and BERT

Satvika Marrapu[1,*], William Senn[2] and Victor Prybutok[3]

[1]*College of Information, University of North Texas, USA*

[2]*College of Business, Tarleton State University, USA*

[3]*College of Business, University of North Texas, USA*

Abstract: Amid the rapidly evolving discussions surrounding the Omicron vaccination, this research leverages data from Twitter, focusing on the USA, from March 2022 to March 2023. Harnessing the capability of the snscrape Python library, a comprehensive dataset of tweets was collated and subsequently subjected to rigorous sentiment analysis techniques. Two primary methodologies were adopted: the Valence Aware Dictionary and sEntiment Reasoner (VADER) and the Bidirectional Encoder Representations from Transformers (BERT) model. The data underwent preprocessing, which included the removal of URLs, HTML tags, mentions, and stop words. Using VADER, the tweets were initially labeled, forming the foundational layer for training the BERT model. Following tokenization, data batching, and model construction, the BERT model was trained and subsequently evaluated. Results illuminated a multifaceted landscape of emotions in discussions related to the Omicron vaccination during the study period. Furthermore, a discernible relationship was identified, highlighting the emotional flux in vaccine-related Twitter dialogues throughout the Omicron period. This study provides valuable insights into public sentiment during a crucial juncture of the pandemic and underscores the potential of contemporary natural language processing tools in gauging public opinion.

Keywords: Omicron vaccination, Twitter data, sentiment analysis, VADER, BERT, NLP, public sentiment

1. Introduction

The recent rise of social media platforms, particularly Twitter, has significantly changed the public discourse domain in multiple sociopolitical and healthcare topics [1]. This transformation has been particularly pronounced in the public health crisis context, where real-time information dissemination and opinion sharing are significant. For instance, The coronavirus disease 2019 (COVID-19) pandemic resulted in high uncertainty concerning suitable treatments and public policy reactions, and the uncertainty provided a breeding ground for spreading conspiratorial anti-science narratives based on disinformation [1, 2]. The emergence of the Omicron variant of the COVID-19 virus in late 2021 presented a pressing challenge to public health authorities worldwide [3]. Since then, vaccination advocacy has been at the center of the global response to the pandemic, and the entire chain surrounding vaccination strategies, including public sentiment and attitudes, has been instrumental in shaping the course of these campaigns [4]. The purpose of this study is to employ state-of-the-art natural language processing (NLP) techniques, specifically the Valence Aware Dictionary and sEntiment Reasoner (VADER) and Bidirectional Encoder Representations from Transformers (BERT), to conduct a comprehensive sentiment analysis of Twitter discourse on Omicron vaccination within the USA.

The motivation behind conducting this study is that as considerable research has focused on the sentiment toward COVID-19 vaccinations in general, there remains a critical gap in understanding public sentiment toward specific variants, such as Omicron. This study aims to bridge this gap by delving into the nuanced differences in public sentiment toward the Omicron variant's vaccination efforts. The emergence of new variants can significantly influence public perception and vaccination strategies due to differences in transmissibility, vaccine efficacy, and severity [5–7]. By focusing on the Omicron variant, this research offers targeted insights that are vital for informing public health messaging and vaccination campaigns tailored to the evolving pandemic landscape. This novel contribution underscores the importance of this study and sets the stage for the subsequent research methodology and findings.

2. Literature Review

The vaccine sentiment analysis is essential in the ongoing COVID-19 pandemic context [3]. Public sentiment toward vaccination is a fundamental indicator of the broader population's views, concerns, and attitudes regarding vaccines [8]. Multiple social media platforms, such as Twitter, are essential in displaying

*Corresponding author: Satvika Marrapu, College of Information, University of North Texas, USA. Email: satvikamarrapu@my.unt.edu

public sentiments toward Omicron vaccination, thus enabling all the interested parties to understand the response among the target population [9]. Understanding these sentiments assumes critical importance for policymakers and healthcare professionals who aim to devise effective communication strategies and interventions to promote vaccine adoption. Numerous studies have explored the sentiment encompassing COVID-19 vaccines on social media, showcasing the practicality of sentiment analysis in monitoring public perceptions and forecasting vaccine hesitancy [3, 9]. However, the advent of novel variants, including the Omicron variant, introduced distinctive dynamics and uncertainties into the COVID-19 vaccine discussion [9]. Given the rapid dissemination of information and misinformation alike across platforms such as Twitter, the necessity of real-time assessment of public sentiment has become indispensable.

The emergence of new COVID-19 variants, including Omicron, necessitates an agile and data-driven approach to comprehend how public sentiments adapt [9]. As these variants present fresh challenges and uncertainties, determining the public's evolving attitudes toward vaccination is important in the contemporary global healthcare domain. Timely and brief insights into these shifting sentiments are essential for the appropriate adjustment of public health strategies [9]. In recent years, Twitter has emerged as a vital platform for the expression of public sentiment, especially for users targeting the global audience [10]. The succinct and real-time nature of Twitter makes it a unique platform for the rapid dissemination of opinions and information [11]. Therefore, it serves as a reflective barometer of public sentiment, offering an invaluable window into the collective psyche during health crises. However, the speed at which information and, at times, misinformation spread across social media platforms, including Twitter, is a challenge that should be addressed. From the vaccine sentiment perspective, this dynamic is especially pronounced. Misinformation leads to low vaccine adoption because of doubts and hesitancy, given that the target population for vaccination will be reluctant to consume the vaccine before their safety is assured [12]. Nonetheless, the instances of spreading positive news concerning the Omicron vaccine would result in high vaccine uptake because the information will quell all the concerns that might arise [13]. Consequently, the ability to discern and respond to these trends swiftly holds the potential to mitigate vaccine hesitancy and bolster public confidence [12]. To navigate the complexities of vaccine sentiment during the Omicron period effectively, it is essential to conduct real-time sentiment analysis. Such an approach allows for the agile adaptation of public health messaging, enabling authorities to address concerns as they arise and to counteract the influence of misinformation promptly. Moreover, real-time analysis offers a unique opportunity to engage with the public actively, responding to their evolving needs and concerns.

In the contemporary sentiment analysis domain concerning vaccine discourse, present methodologies center around the utilization of machine learning algorithms and NLP techniques to evaluate the emotional tone of tweets and other social media contents. This approach enables the extraction of valuable insights from the user-generated content areas concerning vaccines, offering a window into public sentiment and attitudes. One of the notable tools that has garnered considerable recognition in this sector is VADER [14, 15]. This sentiment analysis tool has gained prominence because of its efficacy in deciphering the emotional undercurrents embedded within short, contextually rich social media texts. VADER operates on the lexical and grammatical cues basis, allowing it to discern and quantify sentiment polarity within individual tweets. Its ability to capture emotional expressions has rendered it a valuable instrument in the modern-day public sentiment analysis surrounding vaccines.

Alongside the prevalence of VADER, the emergence of transformer-based models, especially BERT, has resulted in a new sentiment analysis era. BERT is developing into an effective contextual analysis of textual data tool, revolutionizing the NLP field [16]. Its bidirectional architecture enables it to consider the entirety of a sentence or text, taking into account the surrounding context and thereby enhancing its ability to grasp subtle variations in sentiment and semantics [17]. Therefore, the current state of sentiment analysis within vaccine sphere underscores the importance of leveraging cutting-edge tools and techniques. The agreement among machine learning algorithms, NLP, and advanced models, such as BERT, empowers researchers and practitioners to explore the multifaceted public sentiment domain surrounding vaccines with unprecedented depth and precision. As the field continues to evolve, it holds the potential to offer increasingly nuanced insights into the intricate interplay of emotions, opinions, and information in the context of public health and vaccination campaigns.

2.1. Recent advancements in sentiment analysis for health crises

In the realm of sentiment analysis, the COVID-19 pandemic has underscored the urgent need for advanced methodologies capable of capturing and interpreting the complex landscape of public sentiment. The advent of sophisticated machine learning models, particularly transformer-based architectures, has markedly enhanced the analytical depth available to researchers.

2.2. Transformer-based architectures in sentiment analysis

A significant leap forward has been the application of transformer-based models, such as GPT-3 and RoBERTa, which have demonstrated a superior ability to comprehend context and nuance in textual data. Alghamdi et al. [18] showcased the efficacy of GPT-3 in dissecting public discourse around COVID-19 vaccination, revealing nuanced sentiments that traditional models could not capture. Similarly, Thakur [19] utilized RoBERTa to analyze sentiment trends, highlighting its robustness in understanding complex emotional expressions related to the pandemic.

2.3. Multimodal data sources

Beyond textual analysis, the integration of multimodal data sources has emerged as a pivotal advancement. Analyzing sentiments expressed through both text and visual media offers a more comprehensive view of public attitudes. Melton et al. [20] demonstrated how visual content accompanying tweets significantly influences the sentiment conveyed, providing insights that textual analysis alone might overlook. This approach underscores the multifaceted nature of public sentiment and the importance of leveraging diverse data types for a holistic analysis.

2.4. Impact of advanced sentiment analysis on public health strategies

These methodological advancements hold profound implications for public health communication strategies. By employing advanced sentiment analysis tools, health authorities can gain a nuanced understanding of public sentiment toward vaccination, including hesitancies and misinformation. This information is crucial for tailoring communication efforts to address specific concerns and boost vaccination rates amid ongoing health crises.

2.5. Research objectives

This study aims to achieve the following objectives:

1) Conduct a comprehensive sentiment analysis of Twitter discourse related to Omicron vaccination within the USA.
2) Assess the temporal evolution of sentiment and identify critical inflection points in public opinion.
3) Investigate potential correlations between sentiment trends and external events, such as vaccination milestones or regulatory announcements.
4) Compare the performance of VADER and BERT in capturing sentiment nuances within the Twitter discourse.

To achieve these objectives, the study will involve the following steps:

1) Data collection, in which there will be retrieval of a large-scale Twitter dataset containing tweets related to Omicron vaccination in the USA.
2) Preprocessing will include text normalization, tokenization, and removal of noise from the dataset.
3) Sentiment analysis will involve the application of VADER and BERT for sentiment classification.
4) Temporal analysis will include the examination of sentiment trends over time.
5) The event correlation phase will involve the identification of events influencing sentiment fluctuations.
6) Comparative analysis phase will involve the valuation of VADER and BERT's performance in sentiment analysis.

This study was designed to address a critical gap in the understanding of public sentiment regarding Omicron vaccination within the USA. By employing VADER and BERT, we aim to provide a precise and accurate depiction of sentiment dynamics in the context of a rapidly evolving public health situation. The findings of this study hold substantial implications for policymakers and health communication experts, helping them in devising targeted strategies to address vaccine hesitancy and enhance vaccination uptake rates. Moreover, the comparative analysis between VADER and BERT will contribute to the ongoing discourse on the efficacy of NLP tools in the sentiment analysis of social media data, expanding the methodological toolkit available to researchers and practitioners in this domain. This study emphasizes the importance of harnessing advanced NLP techniques to examine public sentiment effectively, thus facilitating more informed and data-driven decision-making in public health interventions. Furthermore, the insights gained from this study are transferable and can provide a valuable framework for investigating similar phenomena in the emerging global public health crisis domain.

3. Research Methodology

3.1. Data collection

The data utilized in this study encompassed 150,000 tweets from a diverse demographic spectrum across the USA, spanning from March 2022 to March 2023. The snscrape Python library was employed to conduct an extensive data retrieval process, extracting tweets through Twitter Application Programming Interface. The dataset included tweets posted during this timeframe and was collected to capture the dynamic evolution of public sentiment over the course of the Omicron variant's prevalence. Table 1 summarizes the Twitter data on US discourse regarding Omicron vaccination (March 2022–March 2023) using the snscrape Python library.

3.2. Dataset representativeness

The dataset was meticulously compiled to ensure broad representation and inclusivity. Through the use of stratified sampling, the aim was to reflect the varied demographic and geographic landscape of the USA, capturing tweets from a diverse array of locations, ages, and genders. This method was instrumental in developing a dataset that not only provides a comprehensive view of public sentiment but also delves into the nuances of different communities' responses to the Omicron vaccination. Stratified sampling increases precision and ensures

Table 1
Summary of Twitter data on US discourse regarding Omicron vaccination (March 2022–March 2023): collection and components using the snscrape Python library

Attribute	Description	Type
CreateDate	The data when the tweet was posted	Date
TweetText	Post that represents Twitter user's opinions	Text
TweetID	The unique identity of a tweet	Numeric
ReplyCount	A processed post that represents users' opinions	Text
RetweetCount	Number of repost a tweet gets	Numeric
LikeCount	Number of users interested or agreed to the tweet	Numeric
lang	Language of the tweet	Text
Hashtags	Metadata tag that is prefaced by the pound symbol	Alphanumeric
GeoCityProvince	The city and province in the tweet originated	Text
GeoCountry	The country of origin of the tweet	Text
AuthorID	A unique number that identifies a user	Text
UserName	Identification is used by a user with access to Twitter	Alphanumeric
CreatedAccountAt	Time and date the user account was created	Data and time
FollowerCount	Total number of users' followers	Numeric
FollowingCount	Total number of users a user is following	Numeric
TweetCount	Total number of tweets	Numeric
State	Name of the state	Text

the representation of all groups of interest, crucial for capturing diverse public sentiment accurately [21–23].

3.3. Size and scope of the dataset

Selecting 150,000 tweets was a strategic decision, balancing comprehensiveness with manageability. This substantial dataset size facilitates robust sentiment analysis while maintaining the feasibility of detailed manual review and annotation when necessary. Covering a full year enabled the observation and analysis of sentiment trends over time, especially in response to key events such as vaccine rollouts, policy changes, and public health announcements concerning the Omicron variant. Stratified sampling can lead to more efficient resource allocation by focusing data collection efforts on the most relevant subgroups [21].

3.4. Representativeness and limitations

Acknowledging the constraints of Twitter as a data source, the discussion includes steps taken to mitigate potential biases and ensure the dataset's representativeness. Although Twitter users may not encompass the general population entirely, the stratified sampling method and the dataset's scale provide a significant, diverse view into public sentiment. This careful consideration highlights the commitment to presenting findings that are both insightful and reflective of a broad spectrum of opinions [23].

By detailing the data collection process and the rationale behind methodological choices, the study's foundation is reinforced, ensuring its relevance, accuracy, and contribution to ongoing discourse on public health and sentiment analysis.

3.5. Data processing

Before performing sentiment analysis, a series of preprocessing steps were executed to enhance the quality of the dataset. These steps included the removal of extraneous information, such as URLs and HTML tags, as well as the elimination of user mentions and stop words. These procedures were essential to ensure that the subsequent sentiment analysis would be based on the core textual content of the tweets, thus mitigating noise and potential biases.

3.6. Sentiment analysis

The study's sentiment analysis framework adopts a two-pronged approach, utilizing both VADER and BERT to leverage their respective advantages in analyzing social media discourse on the Omicron vaccine.

3.7. Initial sentiment classification with VADER

VADER, recognized for its efficiency in processing short texts, was used for preliminary sentiment analysis on the collected tweets. As a lexicon-based tool, it excels at swiftly categorizing texts into positive, neutral, or negative sentiments, offering an initial snapshot of public opinion toward the Omicron vaccine [24].

3.8. Enhanced analysis with BERT

To deepen the initial findings and uncover more nuanced sentiments, the BERT model was employed. With its advanced deep learning capabilities, BERT provides a more detailed analysis of the context and subtleties within tweets, going beyond the initial insights provided by VADER [25]. A subset of tweets was manually reviewed and annotated before applying BERT to train the model on accurately labeled data, thereby enhancing the depth and reliability of the sentiment analysis.

3.9. Rationale and implementation

The decision to use this hybrid analytical approach aimed to combine VADER's quick sentiment evaluation with BERT's sophisticated contextual analysis, offering a comprehensive perspective on public sentiment regarding the Omicron vaccination. This methodology was chosen to ensure the analysis encompasses a wide range of public opinions, from immediate reactions to complex sentiment expressions.

By integrating VADER and BERT in the sentiment analysis process, the goal was to conduct a balanced and in-depth exploration of public sentiment. This approach emphasizes the methodology's transparency and reproducibility, affirming its robustness and innovative nature.

The steps involved in predicting the emotions from Twitter data are as follows:

1) **Step 1:** Preprocessing the Twitter data text (it contains removing URLs, HTML tags, mentions, and stop words).
2) **Step 2:** Labeling the Twitter text data using VADER for the sentiment analyzer.
3) **Step 3:** Preparing the labeled tweet data for training the BERT model. This includes tokenizing the text and batching the data in the data loader.
4) **Step 4:** Building the BERT model.
5) **Step 5:** Training the BERT model.
6) **Step 6:** Evaluating the BERT model against the test set.

BERT being deep bidirectional, unsupervised language representation makes it significant. It is also a slow process since it has several weights to update; at the same time, it is costly to use. Since we are performing a sensitive sentimental analysis, we intend to use it because of its high accuracy with unsupervised capabilities. The framework is displayed in Figure 1, wherein we are not using pre-training but only fine-tuning because of the unsupervised characteristic of the BERT. Figure 1 shows the BERT sentiment analysis framework, emphasizing the process from preprocessing to the evaluation of sentiment predictions.

3.10. Comparative analysis between VADER and BERT

This study's sentiment analysis component required a methodology that not only provided quick, general sentiment assessments but also could delve into the nuances and complexities of public sentiment toward COVID-19 vaccination, particularly concerning the Omicron variant. In our comparative analysis, VADER's efficiency was invaluable for initial sentiment assessments across our large dataset. However, BERT's contextual analysis capabilities were crucial for a deeper understanding of the nuances in public sentiment.

While VADER excels in speed and efficiency, making it suitable for processing large volumes of data quickly, it sometimes lacks the depth of understanding that comes from considering the full context of sentences. BERT, although more computationally intensive, offers a comprehensive understanding of sentiment by analyzing the intricacies of language used in tweets.

Figure 1
Performance analysis of the model

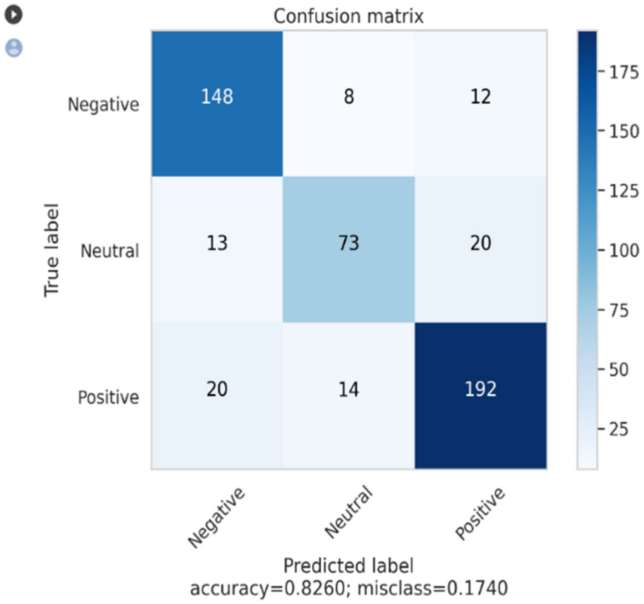
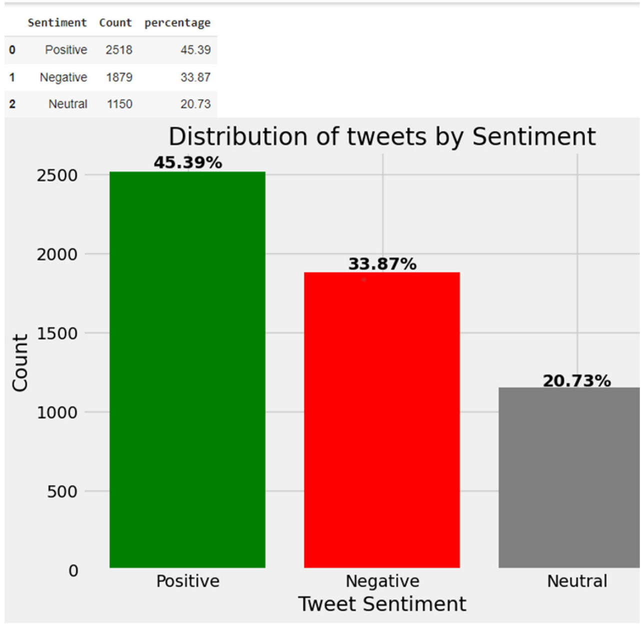


Figure 2
Identify and classify emotions in related discussions from Twitter posts during the Omicron period



3.11. Hybrid approach justification

Given these insights, this study adopts a hybrid approach, leveraging VADER for initial sentiment assessment followed by a more nuanced analysis using BERT. This combination allows us to capitalize on the strengths of both tools: VADER's efficiency and BERT's depth to achieve a comprehensive understanding of public sentiment regarding the Omicron vaccine.

This methodology is structured to ensure clarity in the approach and to justify the use of both VADER and BERT for sentiment analysis. It addresses the complexity of interpreting sentiment on social media platforms and underlines the importance of employing both a lexicon-based and a deep learning model to capture the full spectrum of public sentiment.

3.12. Data evaluation

The performance of the BERT model was evaluated using established metrics such as accuracy, precision, recall, and *F*1 score. The evaluation aimed to assess the model's ability to classify tweets into sentiment categories accurately. Additionally, the temporal aspects of sentiment trends were examined, considering the evolving nature of public discourse over time.

4. Results

Identify and classify emotions in related discussions from Twitter posts during the Omicron period.

The range of sentiments is categorized as follows:

1) ≤ -0.05: negative
2) > -0.05 and < 0.05: neutral
3) ≥ 0.05: positive

Figure 2 identifies and classifies emotions in related discussions from Twitter posts during the Omicron period. The tweets are distributed according to the sentiments, which are positive, negative, and neutral. From the analysis, we can interpret that the Twitter users have posted positively about the vaccination. The above analysis shows that 2518 tweets (45.39%) were classified as positive, 1879 tweets (33.87%) were classified as negative, and 1150 tweets (20.73%) were classified as neutral.

From the Daily Polarity series, we can interpret that the negativity and positivity about vaccination are similar during the starting period of the Omicron vaccination, and it took a turn later after a few months when the positivity of the vaccination increased. But, after a few months, the vaccination's negativity has increased. But from the chart, we can tell that the emotions kept changing of the people daily as their opinions kept changing according to the information in the world.

From the weekly and monthly trends, we can see that the polarity score has been changing daily, but with the sentiments of the monthly trend, we can say that the positivity of the vaccination is increasing at the start of the Omicron. But as more information about the vaccination is out, the negative impact of the vaccination has increased slowly, and to the latest, most people have become neutral about the vaccination.

4.1. BERT model

The error frequency for each class is calculated as above using the values obtained from the confusion matrix. We find that positive sentiments have the highest number of correct predictions by Table 2.

Table 2
Performance of the model

	precision	recall	f1-score	support
Negative	0.82	0.88	0.85	168
Neutral	0.77	0.69	0.73	106
Positive	0.86	0.85	0.85	226
accuracy			0.83	500
macro avg	0.81	0.81	0.81	500
weighted avg	0.83	0.83	0.82	500

The least accurate prediction is obtained for neutral. After training the dataset, the accuracy is 82%.

From the classification report above, we find that the highest number of accurate predictions of sentiments is done by the model for positive, followed by negative and neutral. The precision of positive is the highest among all sentiments. The recall of negative is highest among all sentiments; thereby, the $F1$ score is also high, although the precision is less. The implementation of the logistic model and BERT vectors for finding the sentimental analysis is done with an accuracy of 82%. Both the $F1$ score and error frequency from the confusion matrix show that Twitter users have a positive opinion of the Omicron vaccination.

4.2. Evaluation metrics and model performance

To ensure the rigor and reliability of sentiment analysis, a comprehensive evaluation framework focusing on the BERT model's performance was employed. The methodology emphasizes transparency and reproducibility and also addresses the reported 82% accuracy.

4.3. Accuracy measurement and validation

The reported accuracy of 82% is derived from the BERT model's performance on a validation dataset, meticulously segregated from the training dataset to prevent data leakage and ensure an unbiased assessment of the model's predictive capability. This separation is crucial for evaluating the model's ability to generalize across new, unseen data, thereby validating the 82% accuracy figure as a reliable indicator of the model's performance.

4.4. Additional statistical measures

Beyond accuracy, using precision, recall, and $F1$ score metrics, the model is evaluated. These additional measures provide a granular view of the model's performance across various sentiment categories (positive, neutral, and negative), allowing to address potential concerns regarding the balance between sensitivity and specificity in the model's predictions.

4.5. K-fold cross-validation

To further bolster the credibility of accuracy measurement, a k-fold cross-validation technique was implemented. This approach involved dividing the dataset into 'k' equal segments, training the model on 'k-1' segments, and testing it on the remaining segment [23, 26]. This cycle was repeated 'k' times, with each segment used as the test set once, ensuring that the accuracy figure is robust and reflective of the model's overall performance.

By detailing these steps, the aim was to clarify the methodology and reinforce the validity of the findings, providing a solid foundation for the subsequent analysis and discussion.

5. Discussion

The results of the sentiment analysis on Twitter discourse regarding Omicron vaccination in the USA reveal important insights into public sentiment during a crucial period of the pandemic. This discussion chapter interprets and contextualizes the findings under various subheadings to provide a comprehensive understanding of the emotional landscape and its evolution.

5.1. Emotion classification in Twitter discourse

A substantial portion of the analyzed tweets, comprising 45.39%, exhibited a positive sentiment. This observation highlights the presence of an optimistic and hopeful tone among Twitter users regarding Omicron vaccination [27]. The positivity likely emanates from a sense of trust in vaccination as an effective tool in countering the Omicron variant. Furthermore, it reflects the collective aspiration for a resolution to the enduring pandemic through the successful execution of widespread vaccination efforts. The introduction of COVID-19 vaccines brought a profound sense of relief to the general public worldwide [27–29]. The rapid development, testing, and distribution of vaccines marked a pivotal moment in the ongoing battle against the global pandemic. These vaccines represented a beacon of hope, offering a tangible solution to curb the spread of the virus, reduce the severity of illness, and ultimately save lives.

Approximately 33.87% of the analyzed tweets conveyed a negative sentiment. This negative sentiment can be attributed to a multifaceted set of factors, including concerns about vaccine safety, the propagation of misinformation, and the prevalence of vaccine hesitancy. People who are hesitant or refuse to vaccinate are perceived to have more self-interest, demonstrate distrust to medical specialists and authorities, exhibit greater adherence to religious beliefs, and harbor conspiratorial and suspicious beliefs [30]. The significant presence of negative sentiment highlights the tough challenges confronting public health authorities in their endeavors to address and mitigate vaccine-related apprehensions [31]. Furthermore, it underscores the pervasive influence of misinformation within vaccine discourse on social media platforms.

A noteworthy portion of the tweets, constituting 20.73%, were classified as neutral in sentiment. These neutral sentiments are representative of Twitter users who may adopt a relatively objective stance or choose not to express overtly emotional sentiments concerning Omicron vaccination [9]. This neutrality signifies the diversity of perspectives within the Twitter community and serves as a reminder that not all individuals engage in discourse through an emotional lens [32]. These individuals tend to approach the topic with a sense of objectivity and refrain from conveying strong or emotionally charged opinions and sentiments. Therefore, this neutrality portrays that involved individuals are neither overly optimistic nor excessively pessimistic about the vaccination but may be considering the topic from a more balanced and pragmatic perspective.

This sentiment analysis provides a portrayal of the emotional spectrum within the Omicron vaccination discourse on Twitter. While a substantial proportion of users expressed optimism and trust in vaccination, a significant population also harbors concerns, reflecting the complexities of public sentiment during a health crisis. The presence of neutral sentiments highlights the need for a comprehensive understanding of the diversity of perspectives in shaping public health communication strategies.

5.2. Emotional trends and their impact on public health policy

The analysis highlights significant emotional trends within public sentiment toward the Omicron vaccine, as revealed through Twitter discussions. Acknowledging the impact these trends may have on public health behavior and policy decisions is essential. Notably, periods characterized by positive sentiment toward vaccination often aligned with increased vaccine uptake, indicating that optimistic public discourse can play a pivotal role in supporting vaccination initiatives [33].

5.3. Potential biases

The potential biases inherent in sentiment analysis using Twitter data have been critically examined. The platform's user demographics and behavior may lead to the overrepresentation of certain viewpoints. To counter this, a diversified approach to data collection and analysis was adopted, aiming to make the findings as reflective of broader public sentiment as possible. Despite these measures, the limitations associated with social media data warrant a cautious interpretation of the results, particularly when extending these findings to the entire population. The spread of misinformation and infodemics on social media platforms can significantly affect health behaviors and attitudes toward vaccinations, further complicating the landscape [34].

5.4. Implications for public health messaging and policy

The identified sentiment trends provide valuable insights for refining public health messaging and policy. Aligning communication strategies with current public sentiment allows health authorities to more effectively combat vaccine hesitancy and misinformation. Additionally, monitoring social media sentiment may act as an early indicator of public reception to health policies, facilitating more agile and informed policy-making. The influence of social media on public health behavior and policy decisions underscores the importance of strategic communication to leverage these platforms positively [33, 34].

5.5. Emotional trends over time

The examination of temporal trends in public sentiment, conducted on Twitter discourse related to Omicron vaccination, reveals a dynamic landscape marked by evolving emotions and perceptions throughout the Omicron period [9, 35]. Figure 3 identifies the relationship between emotional changes in vaccine-related discussions on Twitter through the Omicron period. This section provides a concise overview of the key temporal trends observed, underlining the dynamic nature of public sentiment within the context of this public health discourse. Beyond the variations in subject matter within the positive and negative discussions about vaccines, examining the particular audiences engaged in these two discussions catalyzes to expose the unique traits of these groups but also facilitates comprehension of the processes involved in the formation and molding of positive and negative vaccine discourse [35]. Social media platforms, such as Twitter, offer users the opportunity to engage in interactions through actions such as following, sharing, and referencing other user accounts.

With the initiation of the Omicron vaccination campaign, a noticeable equilibrium between positive and negative sentiments emerged within the Twitter discourse [35]. This equilibrium serves as a noteworthy indication that, in the early phases of the Omicron period, the prevailing public opinion was marked by a discernible degree of ambivalence and divergence [36]. This ambivalence likely stems from the prevailing uncertainties surrounding the novel Omicron variant, coupled with the disparities in the accessibility of information at that juncture. The initial wave of positivity could be attributed to the collective hope and anticipation regarding the vaccine's efficacy in combatting the Omicron variant, which was still largely uncharted territory at the outset of the campaign.

Over the course of the Omicron period, temporal analysis indicates a gradual increase in the overall positivity regarding vaccination. This uptick in positivity suggests a potential alignment of public sentiment with the progress of vaccination campaigns and the dissemination of optimistic information regarding vaccine efficacy [36]. However, it is crucial to note that this shift in sentiment was not sustained, as the negativity surrounding vaccination also exhibited an upward trajectory in subsequent months [37]. The evolving nature of public sentiment highlights the influence of external factors, such as news events and the progression of the Omicron variant itself, in shaping Twitter discourse.

An intriguing trend emerges from the monthly sentiment analysis, indicating a notable transition toward neutrality as the Omicron period advanced [37]. Figure 4 illustrates the weekly and monthly sentiment scores. This shift toward neutrality may signify a growing sense of ambiguity or a more measured and cautious approach toward vaccination as the Omicron situation evolved [35–37]. Twitter users appeared to adopt a relatively objective stance, perhaps reflecting a collective desire to assess the situation and make informed decisions amid the evolving circumstances.

The temporal analysis of public sentiment within the Omicron vaccination discourse on Twitter underscores the fluid and adaptive nature of emotions and opinions. The initially divided sentiments,

Figure 3
Identify the relationship between emotional changes in vaccine-related discussions on Twitter through the Omicron period

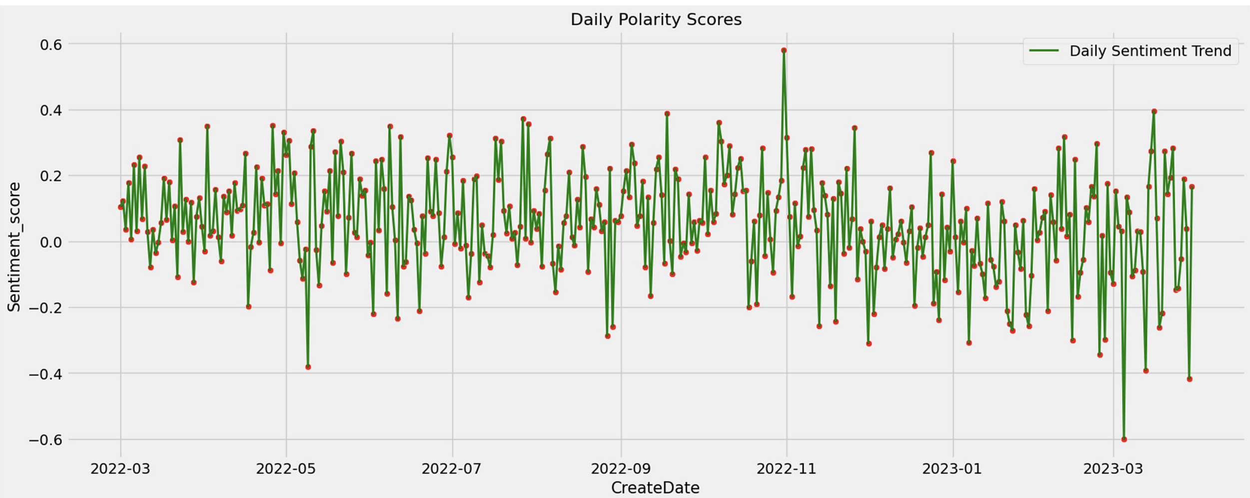
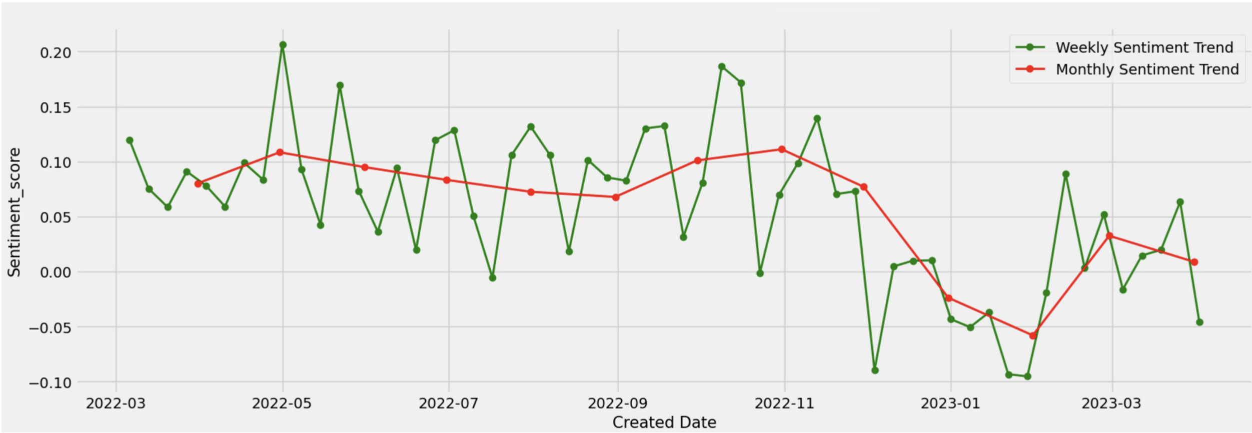


Figure 4
Weekly and monthly sentiment scores



subsequent shifts in positivity and negativity, and the eventual transition toward neutrality collectively reflect the complex interplay of external factors and the evolving dynamics of the Omicron period. These temporal trends highlight the need for a nuanced and adaptable approach to public health communication in response to evolving public sentiment during health crises.

5.6. BERT model performance evaluation

The utilization of the BERT model for sentiment analysis within the Twitter discourse on Omicron vaccination has yielded promising results, signifying its effectiveness in capturing nuanced sentiment expressions. This section provides a concise overview of the vital performance metrics and insights drawn from the BERT model evaluation.

The BERT model achieved an overall accuracy of 82% in its sentiment classification task. This commendable accuracy rate indicates the model's proficiency in correctly categorizing tweets into their respective sentiment categories, thereby demonstrating its capability to discern and capture sentiment nuances embedded within the Twitter discourse [38].

A detailed examination of the model's performance across distinct sentiment categories reveals varying levels of accuracy. The BERT model exhibited the highest accuracy in correctly identifying positive sentiments among the three categories. This observation highlights the model's robustness in accurately recognizing and categorizing tweets expressing optimism and positive outlooks regarding Omicron vaccination [39]. In contrast, the accuracy for neutral sentiments was relatively lower [36]. This variance in accuracy underscores the challenge associated with discerning more nuanced and less emotionally charged expressions within the Twitter discourse. Neutral sentiments often embody a wide spectrum of opinions, making their classification inherently more intricate.

Beyond accuracy, precision and recall metrics provide additional insights into the BERT model's performance. The precision of the BERT model in classifying positive sentiments was notably high. This signifies that when the model categorized a tweet as positive, it was indeed positive, resulting in a low rate of false positives [40]. The model demonstrated the highest recall for negative sentiments. This indicates the model's proficiency in identifying and correctly classifying tweets with negative sentiment [32]. The relatively high $F1$ score for negative sentiments, despite lower precision, suggests that the model excelled in identifying negative tweets, even though it may have occasionally categorized some neutral or positive tweets as negative.

The BERT model's performance evaluation underscores its effectiveness in capturing sentiment nuances within the Omicron vaccination discourse on Twitter. Its overall accuracy rate of 82% indicates its ability to categorize tweets accurately, while sentiment-specific analysis reveals strengths in identifying positive and negative sentiments. The performance metrics provide valuable insights into the model's strengths and areas where further refinement may be beneficial. Overall, the BERT model serves as a robust tool for discerning public sentiment in complex and dynamic contexts such as health crises.

5.7. Overall assessment of public opinion

The sentiment analysis conducted through the BERT model and logistic regression demonstrates that Twitter users predominantly expressed positive sentiments regarding Omicron vaccination during the study period. However, it is crucial to acknowledge the dynamic nature of sentiment, with shifting trends over time and a growing sense of neutrality. The findings emphasize the importance of monitoring and understanding public sentiment on social media platforms, as it can impact public health interventions and communication strategies. Public health authorities should consider the evolving nature of sentiment and the potential influence of external events when devising vaccination campaigns and efforts to combat vaccine hesitancy. This study contributes to the broader understanding of sentiment analysis and highlights the potential of NLP techniques in gauging public opinion. Future research in this area should continue to explore the intricacies of sentiment in the context of emerging health crises, incorporating real-time data analysis and adaptive strategies to address evolving public sentiments effectively.

6. Limitations

The study's reliance on Twitter data may introduce bias, as Twitter users do not represent a fully random sample of the population. Users who engage in vaccine-related discussions on Twitter may have distinct demographics and attitudes compared with the general public. Furthermore, the study's focus on English-language tweets may overlook sentiments expressed in other languages. While rigorous preprocessing was applied to enhance data quality, the removal of certain elements, such as URLs and mentions, may have impacted the contextual richness of the tweets. Additionally, the subjective nature of sentiment analysis could introduce errors in classification. The study's temporal analysis relied on sentiment fluctuations over time, but the underlying reasons for these shifts remain unexplored. Further investigation into the specific events or factors driving sentiment changes could provide deeper insights. Sentiment analysis alone may not capture the context and underlying reasons for sentiment expressions. Understanding the aspects of public sentiment would benefit from qualitative research methods, such as thematic analysis or in-depth interviews. Lastly, the study's findings are specific to the Omicron vaccination period in the USA and may not be directly applicable to other timeframes or geographical regions. Public sentiment can vary significantly based on local factors and cultural contexts.

7. Conclusion

This study provides a comprehensive analysis of public sentiment regarding Omicron vaccination in the USA, as expressed on Twitter from March 2022 to March 2023. The findings of this research elaborate on the evolving emotional landscape during a pivotal phase of the pandemic and highlight the potential of advanced NLP techniques in gauging public opinion. The analysis reveals a notable prevalence of positive sentiment (45.39%) among Twitter users regarding Omicron vaccination, suggesting a general optimism about vaccination as a means to combat the variant. However, it is crucial to note that public sentiment is dynamic, as evidenced by shifts over time. While initial positivity was observed, it was followed by a surge in negativity and a subsequent transition toward neutrality, underscoring the complex interplay of external events and evolving information. The performance of the sentiment analysis model, based on the BERT model and logistic regression, demonstrates its effectiveness in capturing sentiment nuances within the Twitter discourse, with an overall accuracy of 82%. This highlights the potential of NLP tools in monitoring and understanding public sentiment in real time. In the context of further research, it is essential to address the identified limitations, including data bias, the need for contextual analysis, and the consideration of ethical aspects in data collection. Future studies could expand the scope to include multilingual analyses, fine-grained sentiment categories, and real-time monitoring to enhance our understanding of sentiment dynamics in the context of emerging health crises. This study contributes valuable insights into the complex and evolving landscape of public sentiment during a critical period of the COVID-19 pandemic. The findings can inform evidence-based decision-making for public health authorities and underscore the importance of adaptive communication strategies in addressing vaccine hesitancy and promoting vaccination uptake.

Recommendations

Expanding the analysis to include tweets in multiple languages could provide a more comprehensive understanding of global sentiment regarding Omicron vaccination. Comparative analyses between languages and regions could reveal cross-cultural differences in public opinion. Future research could also explore comprehensively sentiment analysis by considering more fine-grained sentiment categories, such as fear, anger, trust, and disgust. This could yield a more nuanced understanding of the emotional dimensions of vaccine discourse. To elucidate the reasons behind sentiment shifts, it is essential to conduct qualitative investigations, such as content analysis or interviews, to gain insights into the context and factors influencing public sentiment. Furthermore, examining sentiment variations at the regional level within the USA or across different countries could uncover geographical disparities in vaccine sentiment, helping tailor interventions to specific regions. Moreover, given the dynamic nature of sentiment on social media, real-time monitoring of sentiment trends could provide timely information for public health authorities to adapt their strategies and messaging. Finally, in order to maximize the breadth and depth of observations, sentiment analyses could be complemented with other novel methods detecting exact sentiment features [41] or advanced analytical techniques, such as association rule mining [42], feature clustering [43], and topological data analysis [44]. While this study offers valuable insights into public sentiment on Omicron vaccination in the USA, it is a stepping stone in a broader research landscape. Addressing the aforementioned limitations and pursuing further research directions can contribute to a more comprehensive understanding of public sentiment in the context of public health crises and vaccination campaigns.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Satvika Marrapu: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **William Senn:** Conceptualization, Validation, Writing – review & editing, Supervision, Project administration. **Victor Prybutok:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision.

References

[1] Darius, P., & Urquhart, M. (2021). Disinformed social movements: A large-scale mapping of conspiracy narratives as online harms during the COVID-19 pandemic. *Online Social Networks and Media*, *26*, 100174. https://doi.org/10.1016/j.osnem.2021.100174

[2] Nguyen, A., & Catalan-Matamoros, D. (2020). Digital mis/disinformation and public engagement with health and science controversies: Fresh perspectives from Covid-19. *Media and Communication*, *8*(2), 323–328. https://doi.org/10.17645/mac.v8i2.3352

[3] Khandia, R., Singhal, S., Alqahtani, T., Kamal, M. A., El-Shall, N. A., Nainu, F., . . . , & Dhama, K. (2022). Emergence of SARS-CoV-2 Omicron (B.1.1.529) variant, salient features, high global health concerns and strategies to counter it amid ongoing COVID-19 pandemic. *Environmental Research*, *209*, 112816. https://doi.org/10.1016/j.envres.2022.112816

[4] Nan, X., Iles, I. A., Yang, B., & Ma, Z. (2022). Public health messaging during the COVID-19 pandemic and beyond: Lessons from communication science. *Health Communication*, *37*(1), 1–19. https://doi.org/10.1080/10410236.2021.1994910

[5] Adeyinka, D. A., Neudorf, C., Camillo, C. A., Marks, W. N., & Muhajarine, N. (2022). COVID-19 vaccination and public health countermeasures on variants of concern in Canada: Evidence from a spatial hierarchical cluster analysis. *JMIR Public Health and Surveillance*, *8*(5), e31968. https://doi.org/10.2196/31968

[6] Nadesalingam, A., Cantoni, D., Wells, D. A., Aguinam, E. T., Ferrari, M., Smith, P., . . . , & Heeney, J. L. (2021). Paucity and discordance of neutralising antibody responses to SARS-CoV-2 VOCs in vaccinated immunodeficient patients and health-care workers in the UK. *The Lancet Microbe*, *2*(9), e416–e418. https://doi.org/10.1016/S2666-5247(21)00157-9

[7] Zhuang, C., Liu, X., Chen, Q., Sun, Y., Su, Y., Huang, S., . . . , & Xia, N. (2022). Protection duration of COVID-19 vaccines: Waning effectiveness and future perspective. *Frontiers in Microbiology*, *13*, 828806. https://doi.org/10.3389/fmicb.2022.828806

[8] Cadeddu, C., Sapienza, M., Castagna, C., Regazzi, L., Paladini, A., Ricciardi, W., & Rosano, A. (2021). Vaccine hesitancy and trust in the scientific community in Italy: Comparative analysis from two recent surveys. *Vaccines*, *9*(10), 1206. https://doi.org/10.3390/VACCINES9101206

[9] Mahyoob, M., Algaraady, J., Alrahaili, M., & Alblwi, A. (2022). Sentiment analysis of public tweets towards the emergence of SARS-CoV-2 Omicron variant: A social media analytics framework. *Engineering, Technology and Applied Science Research*, *12*(3), 8525–8531. https://doi.org/10.48084/etasr.4865

[10] Gunter, B., Koteyko, N., & Atanasova, D. (2014). Sentiment analysis: A market-relevant and reliable measure of public feeling? *International Journal of Market Research*, *56*(2), 231–247. https://doi.org/10.2501/ijmr-2014-014

[11] Elbagir, S., & Yang, J. (2019). Twitter sentiment analysis using natural language toolkit and Vader sentiment. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 1–5.

[12] Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M. D., & Paterson, P. (2014). Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: A systematic review of published literature, 2007–2012. *Vaccine*, *32*(19), 2150–2159. https://doi.org/10.1016/j.vaccine.2014.01.081

[13] Shelby, A., & Ernst, K. (2013). Story and science: How providers and parents can utilize storytelling to combat anti-vaccine misinformation. *Human Vaccines and Immunotherapeutics*, *9*(8), 1795–1801. https://doi.org/10.4161/hv.24828

[14] Dubé, È., Ward, J. K., Verger, P., & Macdonald, N. E. (2021). Vaccine hesitancy, acceptance, and anti-vaccination: Trends and future prospects for public health. *Annual Review of Public Health*, *42*(1), 175–191. https://doi.org/10.1146/annurev-publhealth-090419-102240

[15] William, P., Shrivastava, A., Chauhan, P. S., Raja, M., Ojha, S. B., Kumar, K. (2023). Natural language processing implementation for sentiment analysis on tweets. In N. Marriwala, C. C. Tripathi, S. Jain, & D. Kumar (Eds.), *Mobile radio communications and 5G networks: Proceedings of third MRCN 2022* (pp. 317–327). Springer. https://doi.org/10.1007/978-981-19-7982-8_26

[16] Swarnkar, S. K. (2023). Enhancing sentiment understanding in social media content with BERT algorithm. *Journal of Namibian Studies: History Politics Culture*, *36*, 1–10. https://doi.org/10.59670/jns.v36i.4666

[17] Li, Z., Zhou, L., Yang, X., Jia, H., Li, W., & Zhang, J. (2023). User sentiment analysis of COVID-19 via adversarial training based on the BERT-FGM-BiGRU model. *Systems*, *11*(3), 129. https://doi.org/10.3390/systems11030129

[18] Alghamdi, A. M., Pileggi, S. F., & Sohaib, O. (2023). Social media analysis to enhance sustainable knowledge management: A concise literature review. *Sustainability*, *15*(13), 9957. https://doi.org/10.3390/su15139957

[19] Thakur, N. (2023). Sentiment analysis and text analysis of the public discourse on Twitter about COVID-19 and MPox. *Big Data and Cognitive Computing*, *7*(2), 116. https://doi.org/10.3390/bdcc7020116

[20] Melton, C. A., Olusanya, O. A., Ammar, N., & Shaban-Nejad, A. (2021). Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health*, *14*(10), 1505–1512. https://doi.org/10.1016/j.jiph.2021.08.010

[21] Henry, M. (n.d.). *Stratified sampling: The benefits & how it works.* Retrieved from: https://inmoment.com/blog/stratified-sampling/

[22] Frost, J. (n.d.). *Stratified sampling: Definition, advantages & examples*. Retrieved from: https://statisticsbyjim.com/basics/stratified-sampling/

[23] Roma, P., Monaro, M., Muzi, L., Colasanti, M., Ricci, E., Biondi, S., . . . , & Mazza, C. (2020). How to improve compliance with protective health measures during the COVID-19 outbreak: Testing a moderated mediation model and machine learning algorithms. *International Journal of Environmental Research and Public Health*, *17*(19), 7252. https://doi.org/10.3390/ijerph17197252

[24] Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

[25] Pipalia, K., Bhadja, R., & Shukla, M. (2020). Comparative analysis of different transformer based architectures used in sentiment analysis. In *2020 9th International Conference System Modeling and Advancement in Research Trends,* 411–415. https://doi.org/10.1109/SMART50582.2020.9337081

[26] Guo, P., Liu, T., Zhang, Q., Wang, L., Xiao, J., Zhang, Q., . . . , & Ma, W. (2017). Developing a dengue forecast model using machine learning: A case study in China. *PLoS Neglected Tropical Diseases*, *11*(10), e0005973. https://doi.org/10.1371/journal.pntd.0005973

[27] Qorib, M., Oladunni, T., Denis, M., Ososanya, E., & Cotae, P. (2023). Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset. *Expert Systems with Applications*, *212*, 118715. https://doi.org/10.1016/j.eswa.2022.118715

[28] Albahli, S., & Nawaz, M. (2023). TSM-CV: Twitter sentiment analysis for COVID-19 vaccines using deep learning. *Electronics*, *12*(15), 3372. https://doi.org/10.3390/electronics12153372

[29] Mamdouh Farghaly, H., Gomaa, M. M., Elgeldawi, E., Askr, H., Elshaier, Y. A. M. M., Aboul Ella, H., . . . , & Hassanien, A. E. (2023). A deep learning predictive model for public health concerns and hesitancy toward the COVID-19 vaccines. *Scientific Reports*, *13*(1), 9171. https://doi.org/10.1038/s41598-023-36319-6

[30] Romate, J., Rajkumar, E., Gopi, A., Abraham, J., Rages, J., Lakshmi, R., . . . , & Bhogle, S. (2022). What contributes to COVID-19 vaccine hesitancy? A systematic review of the psychological factors associated with COVID-19 vaccine hesitancy. *Vaccines*, *10*(11), 1777. https://doi.org/10.3390/vaccines10111777

[31] Reñosa, M. D. C., Wachinger, J., Bärnighausen, K., Endoma, V., Landicho-Guevarra, J., Landicho, J., . . . , & McMahon, S. A. (2022). Misinformation, infighting, backlash, and an 'endless' recovery; policymakers recount challenges and mitigating measures after a vaccine scare in the Philippines. *Global Health Action*, *15*(1), 2077536. https://doi.org/10.1080/16549716.2022.2077536

[32] Eom, G., Yun, S., & Byeon, H. (2022). Predicting the sentiment of South Korean Twitter users toward vaccination after the emergence of COVID-19 Omicron variant using deep learning-based natural language processing. *Frontiers in Medicine*, *9*, 948917. https://doi.org/10.3389/fmed.2022.948917

[33] Ghahramani, A., de Courten, M., & Prokofieva, M. (2022). The potential of social media in health promotion beyond creating awareness: An integrative review. *BMC Public Health*, *22*(1), 2402. https://doi.org/10.1186/s12889-022-14885-0

[34] World Health Organization. (2022). *Infodemics and misinformation negatively affect people's health behaviours, new WHO review finds* [Press release]. https://www.who.int/europe/news/item/01-09-2022-infodemics-and-misinformation-negatively-affect-people-s-health-behaviours–new-who-review-finds

[35] Hwang, J., Su, M. H., Jiang, X., Lian, R., Tveleneva, A., & Shah, D. (2022). Vaccine discourse during the onset of the COVID-19 pandemic: Topical structure and source patterns informing efforts to combat vaccine hesitancy. *PLoS ONE*, *17*(7), e0271394. https://doi.org/10.1371/journal.pone.0271394

[36] Catalan-Matamoros, D., Prieto-Sanchez, I., & Langbecker, A. (2023). Crisis communication during COVID-19: English, French, Portuguese, and Spanish discourse of AstraZeneca vaccine and Omicron variant on social media. *Vaccines*, *11*(6), 1100. https://doi.org/10.3390/vaccines11061100

[37] Cotfas, L. A., Crăciun, L., Delcea, C., Florescu, M. S., Kovacs, E. R., Molănescu, A. G., & Orzan, M. (2023). Unveiling vaccine hesitancy on Twitter: Analyzing trends and reasons during the emergence of COVID-19 Delta and Omicron variants. *Vaccines*, *11*(8), 1381. https://doi.org/10.3390/vaccines11081381

[38] 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

[39] Saha, S., Showrov, M. I. H., Rahman, M. M., & Majumder, M. Z. H. (2023). VADER vs. BERT: A comparative performance analysis for sentiment on coronavirus outbreak. In *Machine Intelligence and Emerging Technologies: First International Conference*, 371–385. https://doi.org/10.1007/978-3-031-34619-4_30

[40] Agrawal, S., Jain, S. K., Sharma, S., & Khatri, A. (2023). COVID-19 public opinion: A Twitter healthcare data processing using machine learning methodologies. *International Journal of Environmental Research and Public Health*, *20*(1), 432. https://doi.org/10.3390/ijerph20010432

[41] Agbesi, V. K., Chen, W., Ukwuoma, C. C., Kuadey, N. A., Agbesi, C. C. M., Ejiyi, C. J., . . . , & Kuadey, A. M. (2023). Multichannel 2D-CNN attention-based BiLSTM method for low-resource Ewe sentiment analysis. *Journal of Data Science and Intelligent Systems, 3*(1), 67–77. https://doi.org/10.47852/bonviewJDSIS32021512

[42] Buathong, W., & Jarupunphol, P. (2023). ARM for analyzing factors influencing vaccinations during the COVID-19 outbreak. *Journal of Data Science and Intelligent Systems*, *1*(1), 45–54. https://doi.org/10.47852/bonviewJDSIS3202669

[43] Islam, A., Seth, S., Bhadra, T., Mallik, S., Roy, A., Li, A., & Sarkar, M. (2024). Feature selection, clustering and IoMT on biomedical engineering for COVID-19 pandemic: A comprehensive review. *Journal of Data Science and Intelligent Systems, 2*(4), 191–204. https://doi.org/10.47852/bonviewJDSIS3202916

[44] Onyango, A., Okelo, B., & Omollo, R. (2023). Topological data analysis of COVID-19 using artificial intelligence and machine learning techniques in big datasets of Hausdorff spaces. *Journal of Data Science and Intelligent Systems*, *1*(1), 55–64. https://doi.org/10.47852/bonviewJDSIS3202701

How to Cite: Marrapu, S., Senn, W., & Prybutok, V. (2025). Sentiment Analysis of Twitter Discourse on Omicron Vaccination in the USA Using VADER and BERT. *Journal of Data Science and Intelligent Systems*, *3*(3), 165–175. https://doi.org/10.47852/bonviewJDSIS42022441