

# Analyzing COVID-19 Vaccination Sentiments in Nigerian Cyberspace: Insights from a Manually Annotated Twitter Dataset

Ibrahim Said Ahmad<sup>1,4</sup>[0000-0001-9514-1807], Lukman Jibril Aliyu<sup>2,4</sup>, Auwal Abubakar Khalid<sup>2,4</sup>, Saminu Muhammad Aliyu<sup>2,4</sup>, Shamsuddeen Hassan Muhammad<sup>2,4,5</sup>, Idris Abdulmumin<sup>3,4</sup>, Bala Mairiga Abduljalil<sup>4</sup>, and Bello Shehu Shehu<sup>2,4</sup>

<sup>1</sup> Institute for Experiential AI, Northeastern University, USA

<sup>2</sup> Faculty of Computing, Bayero University Kano, 700241 Kano, Nigeria

<sup>3</sup> Department of Computer Science, Ahmadu Bello University, Zaria, Nigeria

<sup>4</sup> HausaNLP, Nigeria

<sup>5</sup> Faculty of Sciences, University of Porto, Portugal  
isab7070@gmail.com

**Abstract.** Numerous successes have been achieved in combating the COVID-19 pandemic, initially using various precautionary measures like lockdowns, social distancing, and the use of face masks. More recently, various vaccinations have been developed to aid in the prevention or reduction of the severity of the COVID-19 infection. Despite the effectiveness of the precautionary measures and the vaccines, there are several controversies that are massively shared on social media platforms like Twitter. In this paper, we explore the use of state-of-the-art transformer-based language models to study people’s acceptance of vaccines in Nigeria. We developed a novel dataset by crawling multi-lingual tweets using relevant hashtags and keywords. Our analysis and visualizations revealed that most tweets expressed neutral sentiments about COVID-19 vaccines, with some individuals expressing positive views, and there was no strong preference for specific vaccine types, although Moderna received slightly more positive sentiment. We also found out that fine-tuning a pre-trained LLM with an appropriate dataset can yield competitive results, even if the LLM was not initially pre-trained on the specific language of that dataset.

**Keywords:** COVID-19 · Vaccination · Sentiment Analysis · Large language models · Fine-tune.

## 1 Introduction

Over the last decade, social media has provided abundant data that researchers often utilized to derive insights into people’s sentiments. As such, several research areas have emerged relying on social media data, including sentiment analysis, emotion detection, stance detection and social network analysis [6]. In this research, we explored the use of tweets in Nigerian cyberspace for the prediction

of people’s opinions towards the COVID-19 vaccine. This approach is practical because it allows for real-time feedback, is cost-effective, provides a large reach, offers unfiltered opinions, and enables trend analysis.

The COVID-19 pandemic is a global health crisis caused by the spread of the novel coronavirus SARS-CoV-2 [9]. It has had far-reaching and profound impacts on societies, economies, healthcare systems, and daily life worldwide. The pandemic had a significant impact on public health and society at large, necessitating the rapid development and distribution of vaccines to combat the virus. While the availability of vaccines is crucial, so too is public sentiment towards them [12, 4, 21]. In Nigeria, social media platforms such as Twitter have played a crucial role in shaping public discourse around COVID-19 [15, 1]. As such, the discussions on Twitter provide crucial data that can be used to study COVID-19 vaccination in Nigeria. Therefore, we crawl tweets in the Nigerian cyberspace using the discontinued Twitter Academic API<sup>6</sup> from 1st November 2020 to 30th June 2022 using COVID-19 vaccine keywords. A total of 5200 tweets were collected. We preprocessed the tweets by removing URLs, tweets with less than three words, and tweets not in Latin characters, so the tweets were reduced to 4320. Therefore 4,320 tweets were manually annotated by three (3) independent annotators.

We conducted exploratory data analysis to derive insights from the dataset. The results of the analysis reveal that the majority of Nigerians either believe that vaccines are secure and life-saving or hold a neutral stance on the vaccines, while a smaller fraction of Nigerians have expressed concerns about the vaccine’s safety. Additionally, we found that the majority of the people in Nigeria hold a neutral stance on COVID-19 brands, notwithstanding, a small percentage of the people still hold a negative stance towards the brands.

We conducted experiments to demonstrate that a strong performance can be achieved on COVID-19 acceptance prediction using the developed dataset by both classical machine learning models and state-of-the-art pre-trained large language models. The experimental results reveal that the support vector machine achieved the best performance amongst the classical machine learning models while *xlm-roberta base* achieved the best performance amongst the large language models.

**Contributions** The main contributions of this paper are:

- [1] We developed a manually annotated COVID-19 vaccination acceptance dataset in Nigerian cyberspace.
- [2] We analyzed the data and derived novel insights from the data using multiple exploratory data analysis techniques like topic modelling.
- [3] We benchmark the dataset on COVID-19 vaccination sentiment analysis using both classical machine learning techniques and large language models.

<sup>6</sup> <https://www.cnn.com/2023/04/05/tech/academic-researchers-blast-Twitter-paywall/index.html>

The remainder of this paper is organized as follows: Section 2 summarises the related works. Section 3 provides a detailed description of the research methodology. Section 4 presents the results and discussion. Finally, Section 5 concludes the paper with recommendations and future works.

## 2 Related work

**COVID-19 Vaccine Sentiment Analysis:** People’s opinion towards the COVID-19 vaccine has been a subject of study in the research community and sentiment analysis is among the most widely used approaches for this task [2, 14, 16]. A study by [3] collected and analysed tweets on COVID-19 to understand and extract the most discussed topics. They found "Vaccination programme" to be the widely discussed topic while [17] conducted a similar research and found "vaccine hesitancy", "vaccine acceptance" and "vaccine safety" as the most discussed topic related to the COVID-19 vaccine. The study in [12] used social media data to analyse people’s opinions towards the acceptance of COVID-19 vaccines in the US. The vaccine acceptance index (VAI) was used to determine how well people accept these vaccines. The result of their experiment shows a positive increase in the acceptance of the COVID-19 vaccines in the last quarter of 2021. This result contradicts the findings of [8] that says most people are not willing to administer the vaccine. A related study by [19] analysed tweets related to people’s acceptance of the Pfizer vaccine and found that most people have negative perceptions of the vaccine.

**Fine-tuning Large Language Models:** Pre-trained language models are trained on a large amount of data which makes them very popular in NLP tasks. Researchers now prefer fine-tuning instead of training a model from scratch, especially for small datasets [18]. Fine-tuning entails changing the weights of these pre-trained models on a new task to achieve better performance [10]. Keshavarzi et al. [11] fine-tuned BERT [7] to investigate misinformation on Covid-19 vaccine from tweets. Similarly, [5] used BERT to predict COVID-19 vaccine’s side effects using data from social media. The study in [20] fine-tuned RoBERTa [13] to classify sentiments in COVID-19 vaccine tweets.

## 3 Research Methodology

### 3.1 Data collection and Annotation

We collected tweets from Twitter (now X) using the Twitter Academic API from 1st November 2020 to 30th June 2022. We collected tweets keywords relating to COVID-19 vaccination using keywords like vaccin\*, moderna, pfizer, biontech, gamaleya, janssen, astrazeneca, astrazeneca, covishield, sinopharm, jab, microchips, infertility, tracking, magnet. The collection was restricted to only Tweets in Nigeria using location ID in the query. A total of 5200 tweets were

collected. After removing URLs, tweets with less than three words, and tweets not in Latin characters, the tweets were reduced to 4320.

The tweets were then annotated by three (3) independent annotators using an annotation guideline. The classes for the annotation are:

- (a) Positive: If the user’s tweet implies a positive attitude or opinion towards the vaccine (directly or indirectly). For example, the tweets contain positive words or some terms that imply a positive attitude towards COVID-19 vaccine like save lives, protect, appreciative, thankful, excited, or optimistic.
- (b) Negative: If the user’s tweet implies a negative attitude or opinion towards the vaccine. For example, the tweets contain negative words or some terms that imply a negative attitude towards COVID-19 vaccine like kill, harmful, tracker, chip, critical, angry, disappointed in, pessimistic, expressing sarcasm about, or mocking COVID-19 vaccine.
- (c) Neutral: If the user’s tweet does not imply any opinion.
- (d) Indeterminate: If the tweet is unintelligible.

After the annotation, *majority vote* was used to assign the final class of tweets in the following manner:

1. If all the annotators agree on the same class, then that class is automatically accepted.
2. If two annotators agree on either positive or negative and the last annotator classifies the tweet as neutral, then the tweet is assigned the majority class.
3. If two annotators classify a tweet as positive and the third annotator classifies it as negative, then the tweet is adjudicated by a new annotator. This is also the case if two annotators classify a tweet as negative and the third annotator classifies the tweet as positive.
4. If all the annotators label the tweet differently, the tweet is discarded.

At the end of the annotation, 4,320 tweets were annotated. Table 1 gives an example of tweets in the dataset and their annotated classes.

### 3.2 Exploratory Data Analysis (EDA)

Our approach is focused on analyzing public sentiment on Twitter towards COVID-19 vaccines. The initial step we explored is Exploratory Data Analysis (EDA), aimed at gaining valuable insights from the tweets. EDA entails the examination and interpretation of information and relevant data to draw conclusions regarding people’s opinions. Therefore, we visualized the most common words used in COVID-19 discussions using the word clouds. Additionally, we used bar chart to compare the values of positive, negative and neutral tweets in our dataset. We also visualize the people’s sentiments towards the most common COVID-19 vaccines.

Table 1: Example of tweets and their annotated classes

SN	Tweet	Class
1	i took my COVID vaccine jab today monday at the flag off of the fct vaccination campaign i believe it is safe praise god	Positive
2	taking a vaccine that was developed under a year goes against every damn thing we know about drug discovery and vaccine creation refer back to the vaccines are not effective if they were why are vaccinated people being asked to wear masks	Negative
3	COVID vaccine is safe and effective ive gotten mine go and get yours	Positive
4	so who didnt approve the madagascarn herb but the astrazeneca vaccine they approved is now giving issues all over the world oh sorry like this was what naija imported hmmm	Negative
5	gov makinde ssg others vaccinated as oyo govt kicks off COVID vaccination	Neutral

### 3.3 Models

For the prediction of people’s sentiment towards COVID-19 vaccine, we utilize two categories of models. First, we model our prediction based on classical machine learning algorithms. We used classical machine learning algorithms because our dataset is unlikely to yield reasonable results with deep learning models as they require a large training dataset. Additionally, classical machine learning algorithms sometimes yield competitive results compared to deep learning models despite being less computationally expensive. The machine learning algorithms used are Support Vector Machines (SVM), Logistic Regression, Naive Bayes, K-Nearest Neighbours (KNN) and Decision Trees.

Secondly, we model our prediction by fine-tuning large language models. Fine-tuning is an effective method of utilizing pre-trained large language models for downstream NLP tasks using available datasets. The large language models we fine-tuned are Bert, Afro-XLMR, and XLM-roBERTa.

## 4 Results and Discussion

### 4.1 Findings from Exploratory Data Analysis

fig. 1 illustrates the findings from the EDA. Figure 1a shows the sentiment distribution of people towards COVID-19 vaccines. This information helps gauge whether Nigerians have embraced or rejected the safety of the vaccines. The



Table 2: Performance of Classical Machine Learning Models

Model	Accuracy	Precision	Recall	F1- Score
<b>SVM</b>	0.66	0.65	0.66	<b>0.65</b>
Logistic Regression	0.61	0.63	0.61	0.65
Naive Bayes	0.57	0.65	0.57	0.45
K Nearest Neighbours	0.62	0.61	0.62	0.61
Decision Trees	0.53	0.53	0.53	0.53

Table 3: Performance of Fine-tuned Large Language Models

Model	Accuracy	Precision	Recall	F1- Score
Bert base	0.70	0.70	0.70	0.70
Afro-xmlr base	0.71	0.70	0.71	0.69
<b>xlm-roberta base</b>	0.71	0.71	0.71	<b>0.71</b>

where individuals hold negative sentiments, they use words like *risk*, *kill*, *pain*, *problem*, and *wrong* to convey their concerns.

#### 4.2 Classical Machine Learning Models

Table 2 shows the performance of classical machine learning algorithms on COVID-19 vaccine sentiment classification. SVM and Logistic Regression achieved the best performance with an F1-score of 65%. However, SVM outperformed the decision tree in terms of accuracy. Naive Bayes perform poorly with an F1-score of 45% and accuracy of 57%. Similarly, the Decision tree achieved poor results with F1-score of 53% and an accuracy of 53%.

#### 4.3 Fine-tuned Large Language Models

Table 3 shows the evaluation results of fine-tuned LLMs. The models exhibited better performance, which aligns with expectations. XLM-Roberta and Afro-XLMR achieved the best with and F1-score of 71% and 69% respectively, and an accuracy of 71%. These two algorithms have been trained on African languages and as such were expected to attain superior performance due to the prevalence of Nigerian English code-mixed with other Nigerian languages in the dataset. Notwithstanding, the bert model also performed competitively with F1-score and accuracy of 70%. The overall effectiveness of LLMs can be attributed to their extensive training using diverse natural language texts. From our analysis of the results, it is evident that fine-tuning a pre-trained LLM with an appropriate dataset can yield competitive results, even if the LLM was not initially pre-trained on the specific language of that dataset.

## 5 Conclusion and Recommendation

In this work, we developed and publicly released a manually annotated tweet dataset for COVID-19 vaccination exploration in Nigerian cyberspace. The dataset consists of 4320 annotated tweets. In addition, we employed multiple exploratory data analysis techniques to derive insights from the dataset. We found that the majority of the tweets indicate people generally have a neutral sentiment towards the COVID-19 vaccine. Notwithstanding, a good number of people also hold positive sentiments towards COVID-19 vaccines. We also found out that the majority of the people in Nigeria have no significant preference for the common vaccine types, even though the moderna vaccine has slightly more positive sentiments compared to the other vaccines.

In Addition to the dataset, we also investigate the efficiency of fine-tuning LLM for COVID-19 vaccine acceptance prediction in contrast to classical machine learning algorithms using a small dataset. We found out that all the machine learning algorithms performed below the fine-tuned LLMs. This finding can be attributed to the fundamental differences between these two model categories. While LLMs, predominantly based on transformer architectures have been trained to grasp contextual information, classical machine learning models primarily focus on predicting labels without considering contextual nuances. Among the classical machine learning algorithms, SVM outperformed the other algorithms in terms of macro average F1-score. This superiority can be elucidated by the algorithm’s ability to capture non-linear relationships within the data points.

Future work should dig further to fine-tune the hyperparameters and pre-processing steps to obtain the best model performance. Moreover, the limitation of our study is the relatively small dataset, comprising fewer than four thousand samples, coupled with label imbalance. These factors may constrain the performance of the fine-tuned models.

## References

- [1] Hassan Adamu et al. “Framing twitter public sentiment on Nigerian government COVID-19 palliatives distribution using machine learning”. In: *Sustainability* 13.6 (2021), p. 3497.
- [2] Irfan Aygün, Buket Kaya, and Mehmet Kaya. “Aspect based twitter sentiment analysis on vaccination and vaccine types in covid-19 pandemic with deep learning”. In: *IEEE Journal of Biomedical and Health Informatics* 26.5 (2021), pp. 2360–2369.
- [3] Trisha Baldha et al. “Covid-19 Vaccine Tweets Sentiment Analysis and Topic Modelling for Public Opinion Mining”. In: *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*. IEEE, 2021, pp. 1–6.
- [4] Guillermo Blanco and Anália Lourenço. “Optimism and pessimism analysis using deep learning on COVID-19 related twitter conversations”. In: *Information processing & management* 59.3 (2022), p. 102918.



- [5] Saeyeon Cheon, Thanin Methiyothin, and Insung Ahn. “Analysis of COVID-19 vaccine adverse event using language model and unsupervised machine learning”. In: *Plos one* 18.2 (2023), e0282119.
- [6] Jingfeng Cui et al. “Survey on sentiment analysis: evolution of research methods and topics”. In: *Artificial Intelligence Review* (2023), pp. 1–42.
- [7] Jacob Devlin et al. “Bert: Pre-training of deep bidirectional transformers for language understanding”. In: *arXiv preprint arXiv:1810.04805* (2018).
- [8] Ariel Fridman, Rachel Gershon, and Ayelet Gneezy. “COVID-19 and vaccine hesitancy: A longitudinal study”. In: *PloS one* 16.4 (2021), e0250123.
- [9] Abdullah Al-Hashedi et al. “Ensemble classifiers for arabic sentiment analysis of social network (twitter data) towards covid-19-related conspiracy theories”. In: *Applied Computational Intelligence and Soft Computing* 2022 (2022), pp. 1–10.
- [10] Haoming Jiang et al. “Smart: Robust and efficient fine-tuning for pre-trained natural language models through principled regularized optimization”. In: *arXiv preprint arXiv:1911.03437* (2019).
- [11] Arash Keshavarzi Arshadi et al. “Artificial intelligence for COVID-19 drug discovery and vaccine development”. In: *Frontiers in Artificial Intelligence* (2020), p. 65.
- [12] Lingyao Li et al. “Dynamic assessment of the COVID-19 vaccine acceptance leveraging social media data”. In: *Journal of Biomedical Informatics* 129 (2022), p. 104054.
- [13] Yinhan Liu et al. “Roberta: A robustly optimized bert pretraining approach”. In: *arXiv preprint arXiv:1907.11692* (2019).
- [14] Chad A Melton et al. “Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence”. In: *Journal of Infection and Public Health* 14.10 (2021), pp. 1505–1512.
- [15] Ngozika A Obi-Ani, Chinenye Anikwenze, and Mathias Chukwudi Isiani. “Social media and the Covid-19 pandemic: Observations from Nigeria”. In: *Cogent arts & humanities* 7.1 (2020), p. 1799483.
- [16] Annisa Ristya Rahmanti et al. “Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout”. In: *Computer Methods and Programs in Biomedicine* 221 (2022), p. 106838.
- [17] Akila Sarirete. “A bibliometric analysis of COVID-19 vaccines and sentiment analysis”. In: *Procedia Computer Science* 194 (2021), pp. 280–287.
- [18] Annika Marie Schoene et al. “An Example of (Too Much) Hyper-Parameter Tuning In Suicide Ideation Detection”. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 17. 2023, pp. 1158–1162.
- [19] Milind Shah et al. “Sentiment Analysis of Twitter Tweets for COVID-19 Pfizer Vaccines”. In: *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE. 2023, pp. 929–937.
- [20] Jason Wei et al. “Finetuned language models are zero-shot learners”. In: *arXiv preprint arXiv:2109.01652* (2021).

- [21] Hui Yin et al. “Sentiment analysis and topic modeling for COVID-19 vaccine discussions”. In: *World Wide Web* 25.3 (2022), pp. 1067–1083.