Vaccine Hesitancy Hotspots in Africa: An Insight From Geotagged Twitter Posts

Blessing Ogbuokiri⬤, Ali Ahmadi⬤, *Member, IEEE*, Zahra Movahedi Nia⬤, Bruce Mellado⬤, Jiahong Wu, James Orbinski, Ali Asgary⬤, and Jude Kong⬤

Abstract—Many social media users express concerns about vaccines and their side effects on Twitter. These concerns lead to a compromise of confidence which brings about vaccine hesitancy. In Africa, vaccine hesitancy is a major challenge faced by health policymakers in the fight against COVID-19. Given that most tweets are geotagged, clustering them according to their sentiments could help identify locations that may likely experience vaccine hesitancy for health policy and planning. In this study, we collected 70 000 geotagged vaccine-related tweets in nine African countries, from December 2020 to February 2022. The tweets were classified into three sentiment classes— positive, negative, and neutral. The quality of the classification outputs was achieved using Naíve Bayes (NB), logistic regression (LR), support vector machines (SVMs), decision tree (DT), and K-nearest neighbor (KNN) machine learning classifiers. The LR achieved the highest accuracy of 71% with an average area under the curve of 85%. The point-based location technique was used to calculate the hotspots based on the locations of the classified tweets. Locations with green, red, and gray backgrounds on the map signify a hotspot for positive, negative, and neutral sentiments. The outcome of this research shows that discussions on social media can be analyzed to identify hotspots during a disease outbreak, which could inform health policy in planning and management of vaccine hesitancy in Africa.

Index Terms—COVID-19, data mining, hotspot analysis, twitter post, vaccination, vaccine hesitancy, visualization.

I. INTRODUCTION

THE emergence of the novel COVID-19 pandemics changed the way the world does things, especially in the

Fig. 1. Countries reporting logistic data (adapted from [6]). This figure is the combination of countries with expired doses and at risk to expire doses of vaccines. It was used for comparison and validation of the results of our social media data in this article.

restrictions of movement and lockdowns of businesses in the last two years [1], [2], [3]. The impact of the lockdown of businesses on the world's economy is clearly significant in the rate of inflation in recent times [4]. As of August 2022, over 12 130 billion doses of vaccines have been administered globally [5]. Africa contributed about 5.83% of the total vaccines doses administered in the world with over 70 82 19 474 doses [5], [6]. Thirty two countries in Africa were identified as, combined, countries with expired doses and countries with doses at risk [6], see Fig. 1.

As Africa and the rest of world recovers from this obvious shock caused by the COVID-19 outbreak, vaccination against the virus has remained necessary in gaining immunity toward the management and control of the COVID-19 pandemic. The announcement of the vaccine mandate by health policymakers was graced with a lot of opposing views in Africa [7], [8], [9].

For instance, some conspiracy theorists had rumored that the COVID-19 vaccination mandate is targeted to depopulate Africa. While some religious leaders and influencers advised their followers against taking the COVID-19 vaccines [3], [10]. Others took to legal action against the compulsory vaccination of citizens, they argued that it is against the citizens' fundamental human rights [11]. In order to promote their views and pass the antivaccination message to a wider audience, most of the conveners took to social media, such as Twitter to create hashtags (topics) to drive their points [12], [13].

Information sharing on Twitter spreads very fast even if it is a rumor from an unverified source. The impact of rumors is always dangerous especially in places where users are not well

informed about the subject of discuss [14]. Antivaccination messages spread as a form of users' post, retweet, or share without any form of editorial oversight. These in a way weaken the confidence level of the public well before they are vaccinated [4], [15].

According Al-Uqdah [16], people who use social media without referencing trusted sources may be particularly vulnerable to disinformation. Similarly, vaccine-hesitant persons are more likely to be exposed to nontrusted social media sites as their only information source. In [17], discussion of the current position of social media platforms in propagating vaccine hesitancy was given full consideration. The steps on how social media may be used to improve health literacy and public trust in vaccination were also considered. Neha et al. [17] examined social media and vaccine hesitancy as a new update for the COVID-19 era and other globalized infectious diseases.

Meanwhile, Ennab et al. [18] suggested that misinformation about COVID-19 on social media could pose a significant threat to public health as it has the potential to worsen public health challenges by way of encouraging disease spread among pregnant women. The detrimental influence of vaccine misinformation on Twitter to public health cannot be over emphasized. There should be an increase in the use of social media like Twitter to support public health in the continuous struggle against vaccine hesitancy not just in the era of COVID-19 but for future outbreaks [19], [20].

Insights from Twitter posts (tweets) can help health policymakers to understand the extent of vaccination awareness from users. One of the ways to achieve this is to perform sentiment analysis on these tweets. This process in a way brings to knowledge the opinion of users' about a subject matter, such as COVID-19 vaccination. Sentiment analysis on tweets can only be achieved by, first, annotating the tweets into any of positive, negative, or neutral sentiment classes. Annotation of tweets could be done manually by human beings or automatically. The manual annotation of tweet is said to be the most reliable option. But, it is greeted with human bias and can be time consuming when working on a large amount of dataset [21]. The automatic annotation of tweets involves the use of pretrained models that recognize polarity of words in a given text. It can be implemented using Textblob, AFINN, or VADER. The use of automatic classification models is best when working on a large amount of dataset. Because it is fast and there is low rate of human bias. In the remaining part of this section, we will discuss these tools in a nutshell.

A. Sentiment Analysis With Textblob

TextBlob is a known sentiment analysis lexicon-based model available as a Python library. It provides a simplified text processing technique relevant for natural language processing (NLP). Textblob assigns score of -1 and 1 to each word based on the polarity and subjectivity of the text. In [21], Textblob was used to annotate US airline dataset containing 14 640 tweets reviews. The annotated dataset was trained on six supervised traditional machine learning models. The model developed with deep learning algorithms performed better than the model developed with the traditional machine learning algorithms when trained on Textblob annotated text dataset [22], [23]. But, with a proper parameter fine tuning, performance of the traditional machine learning algorithms could be improved [24], [25].

B. Sentiment Analysis With AFINN

AFINN is a lexicon technique that uses dictionary of words together with polarity score. It maps the corresponding polarity to every word in the text. AFFINN has been used to perform sentiment analysis in different areas, including financial forecasting [26], customer rating [27], and product reviews [28]. However, there was low performance of models trained with COVID-19 tweets dataset that was annotated with AFINN. The results showed a lower performance with traditional machine together with deep learning models. This shows that AFINN annotated tweets dataset could produce lower accuracy score than tweets annotated with Textblob or VADER [21].

C. Sentiment Analysis With VADER

VADER means Valence Aware Dictionary for sEntiment Reasoning [21]. It generates sentiment scores or intensities using a dictionary by mapping the lexicon features. VADER has been used to annotate different text-based dataset, especially tweets. Existing research shows that tweets dataset annotated by VADER produces best result when trained on traditional supervised machine learning [29], [30], [31]. The optimum performance of VADER dataset is achieved from the multiclassification of the dataset [32].

We observed that Textblob, AFINN, and VADER annotate texts differently. The sentiment classes are sometimes different depending on the tool used to annotate. This variety in annotation also influences the performance of different machine learning models. This means that, all these tools have the ability to annotate text in their own way. However, given that VADER is the most recent tool developed specifically for social media text, we choose VADER to annotate our dataset [33].

To the best of our knowledge, there is no work that used VADER to annotate geotagged COVID-19 vaccine tweets to identify hotspots. By geotagged COVID-19 vaccine tweets, we refer to the geographical identification metadata of the tweets. The tweet classification was done in nine African countries. The nine African countries are, in no particular order, Nigeria, South Africa, Zimbabwe, Botswana, Namibia, Rwanda, Mozambique, Cameroon, and Eswatini. These countries were selected because our consortium, the African-Canada AI & Data Innovation Consortium (ACADIC), has partner members in these countries. The tweets were collected using trending keywords about the COVID-19 vaccines in the selected countries.

The main contribution of this research is the application of a point-based location technique in the identification and visualization of vaccine hesitancy hotspots from labeled tweets. First, we generated a dataset containing geotagged COVID-19 vaccine tweets from the nine African countries.

The tweets were annotated according to their sentiments using VADER. The result was validated with traditional machine learning classifiers. The machine learning classifiers are Naíve Bayes (NB), logistic regression (LR), support vector machines (SVMs), decision tree (DT), and K-nearest neighbors (KNN). Then, we calculated the hotspots based on the location of the labeled tweets and visualized the same on a map. The research shows that by analyzing social media discussions, hotspots of sentiment towards vaccination during a disease outbreak can be identified. Especially, toward vaccination during a disease outbreak. We believe that this could inform health policy in planning and management of vaccine hesitancy in Africa.

The remainder of this article is organized as follows. Section II discussed different approaches used in the selection of the keywords, collection, preparation, and analysis of data. Section III presents the outcome of the analysis of the data collection, analysis, and visualization of the hotspots by countries. Section IV discusses the results in relation to other relevant existing works. Finally, Section V presents the conclusion of the manuscript and further recommendations.

II. METHODS

This section is divided into list of abbreviations, keyword selection metric, data collection, data preprocessing, tweet labeling, tweet sentiment classification, selection of model parameters, and sentiment visualization.

A. List of Abbreviations

In this section, we present the list of frequently used abbreviations in Table I.

B. Keyword Selection Metric

In selecting the keywords for the data collection, we first identified the approved and administered vaccines in the nine African countries. Table II shows the approved vaccines by country.

We identified all these approved vaccine in Table II because we used the trending topics around the vaccines to collect the data. By trending topics we refer to keywords, phrases, or topics that are commonly used or mentioned about the vaccine within a period of time in a social network or microblog such as Twitter. The trending topics, hashtags, or keywords about a vaccine are selected from the Twitter app search box. Then, we selected the Trending tab and the COVID-19 tab. This approach was used to select the following popular combined keywords and hashtags for the nine African countries. Fig. 2 shows the list of the 43 popular keywords we used for this research. These keywords were generated from the approved vaccines in the selected countries.

C. Data Collection

We used the academic researcher account from the Twitter API to access the Twitter database which allows collection of up to 10 million historical tweets per month. First, we created an application to generate an access token from Twitter. The access token was used in Python version 3.6 script to

TABLE I
LIST OF ABBREVIATIONS. THIS TABLE WAS USED
FOR EASY READABILITY

Abbreviation	Meaning	Description
NLP	Natural Language Processing	Technique for processing tweets.
NLTK	Natural Language Toolkit	NLP tool for tweet processing.
VADER	Valence Aware Dictionary for Sentiment Reasoning	Pre-trained model for tweet labelling.
NB	Naive Bayes	Machine learning classifier.
LR	Logistic Regression	Machine learning classifier.
SVMs	Support Vector Machines	Machine learning classifier.
DT	Decision Tree	Machine learning classifier.
KNN	K-Nearest Neighbours	Machine learning classifier.
TP	True Positive	Tweet correctly labelled by VADER is correctly classified by a classifier.
TN	True Negative	Tweet wrongly labelled by VADER is classified wrongly by a classifier.
FP	False Positive	Tweet wrongly labelled by VADER is classified correctly by a classifier.
FN	False Nagetive	Tweet correctly labelled by VADER is wrongly classified by a classifier.
TPR	True Positive Rate	Probability of a classifier to accurately predict the properly labelled tweets.
FPR	False Positive Rate	Probability of a classifier to classify tweets correctly when they are not actually correctly labelled.
ROC	Receiver Operating Characteristic	The graph of the performance of the classifiers.
AUC	Area Under the Curve	Measures the area under the ROC curve.

TABLE II
APPROVED VACCINES BY COUNTRY [34]. THIS TABLE WAS USED TO
IDENTIFY THE POPULAR TOPICS AROUND THE VACCINES FOR DATA
COLLECTION

SN	Country	Vaccine Approved
1	Nigeria	Oxford-AstraZeneca, Moderna, Johonson & Johnson, Sputnik, PfizerBioNTech, and Sinopharm BIBP
2	South Africa	Johnson & Johnson, Pfizer-BioNTech, and Oxford-AstraZeneca
3	Zimbabwe	Sinopharm BIBP, Covaxin, Sinovac, and Sputnik V
4	Botswana	Pfizer-BioNTech, CoronaVac, Covaxin, Janssen, Moderna, and Oxford-AstraZeneca
5	Namibia	JOxford-AstraZeneca, Pfizer-BioNTech, Johonson & Johnson, Sputnik, and Sinopharm
6	Rwanda	Moderna, Pfizer-BioNTech, AstraZeneca, Johnson & Johnson, Sinopharm and Sputnik V
7	Mozambique	Oxford-AstraZeneca and Sinopharm
8	Cameroon	Pfizer-BioNTech, Sputnik, Johonson & Johnson, Oxford-AstraZeneca, and Sinopharm
9	Eswatini	Pfizer-BioNTech, Johnson & Johnson, and Oxford-AstraZeneca

authenticate and establish a connection to the Twitter database. We got historical COVID-19 vaccine-related tweets, geotagged according to the nine African countries using Python

```
Oxford-AstraZeneca,         AstraZeneca,
JohnsonJohnson,     Vaccine,     BioNTech,
anti-vaccine, jab, Vaccination, Covax, Vaccine
Rollout, Sputnik, #VaccineToSaveSouthAfrica,
#IChooseVaccination,       #TeachersVaccine,
AstraZeneca vaccine,     Pfizer,     J & J,
Johonson & Johnson, Moderna, #VaccinesWork,
#VacciNation,      #Vaccine,      Steriod,
#COVIDvaccine,    #covax,    #VaccineEquity,
#VaccineReady, #Jab, #PfizerGang, #Scamdemic,
#Plandemic,     #Scaredemic,     #COVID-19,
#coronavirus, #SARS-CoV-2, anti-vaxxers, #jab,
Pfizer, BioNTech, JJ Vaccine, JohnsonJohnson
Vaccine, Vaccine Rollout, CoronaVac
```

Fig. 2. List of popular keywords used. This list presented for open access and reusability.

TABLE III
DATASET FEATURES. THIS TABLE DESCRIBES THE CONTENT OF THE DATASET USED

Attribute	Description	Type
TweetText	Twitter post that represents users' opinions	Text
TweetID	Unique identity of a tweet	Numeric
Create Date	The date when the tweet was posted	Date
Retweet Count	Number of re-post a tweet gets	Numeric
Reply Count	A processed post that represents users' opinions	Text
Like Count	Number of users interested or agreed to the tweet	Numeric
GeoId	Unique identifier of the the region or location of the the tweet	alphanumeric
Geo City Province	The city and province the tweet originated	Text
Geo Country	The country of origin of he tweet	Text
Geo Coordinate(bbox)	The area of the tweet defined by two longitudes and two latitudes	Numeric
Hashtags	Metadata tag that is prefaced by the pound symbol	Alphanumeric

script. The Python script was used to perform a historical (an archive) search of tweets that contain COVID-19 vaccine keywords in Fig. 2. The preferred language of the tweet is English.

The total of 70 000 tweets was collected from December 2020 to February 2022 using the archive search process. Each Tweet contains most of the following features described in Table III.

All tweets collected in the search were anonymized to protect users' privacy. The distribution of the collected tweets by country is shown in Fig. 3.

D. Data Preprocessing

User tweets are normally unstructured and contain a lot of information about the data they represent that may not be useful. Cleaning the user tweets is highly needed. We collected tweets, date created, time created, and provinces from the dataset into a dataframe using Pandas version 1.2.4 [35]. The tweets were prepared for NLP by first removing the URLs, duplicate tweets, tweets with incomplete information, punctuations, special and nonalphabetical characters, non-English words, and Stopwords using the tweets-preprocessor
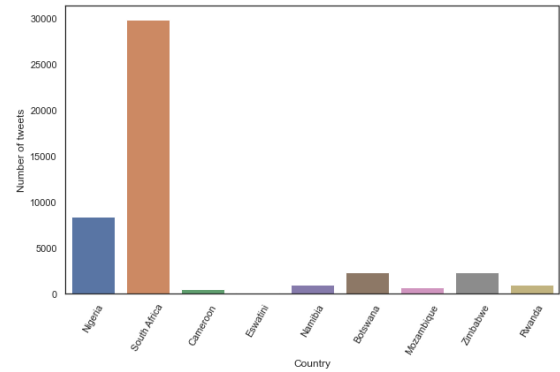

Fig. 3. Distribution of collected tweets by country. This shows the number of tweets collected from each country.

toolkit version 0.6.0 [36], Natural Language Toolkit (NLTK version 3.6.2) [37], and Spacy2 toolkit (version 3.2) [38], [39]. We also used the Spacy2 toolkit to perform tokenization of the tweets. Tokenization in this context is the process of breaking tweets into words. This process reduced the tweets in the dataset to 46 189 tweets.

E. Tweet Labeling

Given the size of the dataset, we used VADER [33], [39], a lexicon and rule-based pretrained NLP tool in the NLTK package. VADER is used to perform sentiment analysis expressed on social media text such as Twitter. It can handle the sentiment expressed in words, abbreviations, emojis, and intensity of emotions [40]. By pretrained, we mean the VADER model has been previously trained on a large dataset to recognize sentiments expressed on texts, especially social media text. When our dataset was passed into VADER model, each line of tweet was picked and a vector of sentiment scores for positive, negative, neutral, and compound polarities were produced. The sentiment score of the tweet was obtained by summing up the intensity of each word in the tweet. The sentiment scores for each of the sentiment class (positive, negative, and neutral) were normalized to be between 0 and 1.

The compound polarity score is the aggregate measure of all the sentiments, normalized to be between the range $[-1.0, +1.0]$ where -1.0 represents extreme negative and $+1.0$ represents extreme positive [40], [41]. Furthermore, the compound polarity was used to assign the sentiment such as, positive, negative, or neutral to a tweet as label. A tweet with a compound polarity ≥ 0.5 is assigned the label positive, <0 is assigned the label negative, and x, where x satisfies the inequality $0.5 > x \geq 0$ is assigned the label neutral [42].

Sample tweets and the corresponding sentiments scores are shown in Fig. 4. Table IV shows labeled tweets using VADER tool.

From the compound polarity scores in Fig. 4, the tweets are labeled as shown in Table IV.

The approach demonstrated in Section II-E was used to label all the tweets according to their sentiment class. A distribution of tweet sentiments by country labeled using VADER pretrained model is shown in Table V.

However, after labeling the tweets, the distribution of the sentiment classes in the dataset was imbalanced. Most

- Tweet 1: The vaccine jab I took is amazing
- Output: {'neg': 0.0, 'neu': 0.513, 'pos': 0.487, 'compound': 0.5859}
- Tweet 2: The new COVID-19 variant is worst
- Output: {'neg': 0.539, 'neu': 0.461,'pos': 0.0, 'compound': -0.8849}
- Tweet 3: This vaccine by ah.
- Output: {'neg': 0, 'neu': 1 ,'pos': 0.0, 'compound': 0 }

Fig. 4.   Sample tweet sentiment scores.

TABLE IV

SAMPLE LABELED TWEETS USING VADER. THIS SHOWS HOW VADER ASSIGNS SENTIMENT CLASS TO USER TWEETS

SN	Tweet	Sentiment
1	The vaccine jab I took is amazing	Positive
2	The new COVID-19 variant is worst	Negative
3	This vaccine by ah	Neutral

TABLE V

SENTIMENT CLASSIFICATION. THIS TABLE SUMMARIZES THE DISTRIBUTION OF THE TWEET SENTIMENTS BY COUNTRY AS CLASSIFIED BY VADER

SN	Country	Sentiment Class		
		Positive	Negative	Neutral
1	Botswana	498	286	1529
2	Cameroon	84	67	391
3	Eswatini	21	31	58
4	Mozambique	117	20	559
5	Namibia	323	152	543
6	Nigeria	1946	2163	4268
7	Rwanda	188	163	611
8	South Africa	5347	8881	15629
9	Zimbabwe	563	400	1365

of the tweets in the dataset were labeled neutral. The SMOTE [43] sampling technique was used to balance the dataset. We achieved a balanced distribution of 33.333% for each of the target parameters (positive, negative, and neutral) in the dataset.

F. Tweet Sentiment Classification

To validate the output of the tweet labeling done by the VADER pretrained model [44], we used five machine learning classifiers including NB [45], LR [46], [47], SVMs [47], [48], DT [49], and KNN [50]. The reason we chose these classifiers is because they have been successfully used to classify tweets in [42].

The performance evaluation metrics were used as a measure of how well the classifiers classified the tweets according to their sentiments against the VADER labeling. They include accuracy, F-measure, ROC curve, and AUC (see Table I). Accuracy shows the part of the predictions the classifier got correctly. It is the most popular performance evaluation metric. Accuracy evaluation metric was used because it focuses on the number of correctly classified tweets with respect to the total number of tweets in the dataset [51]. The accuracy formula adapted from [52] is shown in the following:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

where TP, TN, FP, and FN represent the true positive, true negative, false positive, and false negative, respectively, see Table I.

The F-measure calculates the harmonic mean between precision and recall. The F-measure score is in the range [0, 1]. It was used because it tells how precise the classifier is, that is, how many tweets it classifies correctly, as well as how robust it is (probability of making correct classification). F-measure tries to find the balance between precision and recall. The greater the F-measure, the better the performance of the classifier. F-measure also referred to as F1-score in this article, was calculated using (2) as adapted from [52]

$$\text{F-measure} = 2 \times \frac{1}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}} \tag{2}$$

where precision is the ratio of correctly labeled tweets by VADER and correctly labeled tweets predicted by the classifier. Recall is the ratio of correctly labeled tweets by VADER and relevant tweets (all tweets that should have been labeled correctly by the classifier). Precision and recall were calculated using (3) and (4) as adapted from [52]

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{4}$$

Meanwhile, the ROC curve was used because it shows the plot of sensitivity [true positive rate (TPR) or recall] against specificity (false positive rate (FPR)]. The higher the FPR, the more likely the classifier will classify the tweet sentiment wrongly. Mathematically, (5) and (6) are used to calculate the TPR and FPR as adapted from [52]

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{6}$$

The more the ROC curve aligns toward the top left corner of the plot, the better the classifier does at classifying the tweets into various sentiment classes. This, however, helps to ascertain how well the classifier classified each tweet sentiment class against the tweets labeled by VEDAR pretrained model. AUC is the area under the curve of the plot FPR against TPR at different points in the range [0, 1]. A higher value suggests that the classifier performed well. Therefore, AUC performance metric was used to ascertain how much of the plot is located under the curve.

G. Model Parameter Selection

Model performance could be optimized by tuning the user-defined parameters for better result. We present the user-defined parameters used for each classifier. Other parameters that are not stated here are on their default values in the sklearn package. Each classifier was defined from the sklearn package in Python and their parameter fine-tuned as shown in Table VI.

H. Sentiment Visualization

As Twitter no longer gives access to Latitude and Longitude of user tweets, we collected the GeoCoordinate(bbox) or

TABLE VI

DESCRIPTION OF MACHINE LEARNING HYPERPARAMETERS USED. THESE ARE USER-DEFINED PARAMETERS USED IN THE MODELS. THIS IS USEFUL FOR REUSEABILITY

| Model	Hyperparameter	Description
NB	alpha=1.0	Controls the form of the model
LR	multi_class='rbf'	Decision function for multi classification
	C=3.0	Strength of inverse regularisation
	kernel = 'rbf'	Specifies the type of kernel used.
SVM	C = 1.0	Regularization parameter.
	random_state = 42	Control random number for data shuffling.
DT	max_depth=32	Maximum depth of tree.
KNN	n_neighbors=7	Maximum data point in a neighbourhood.
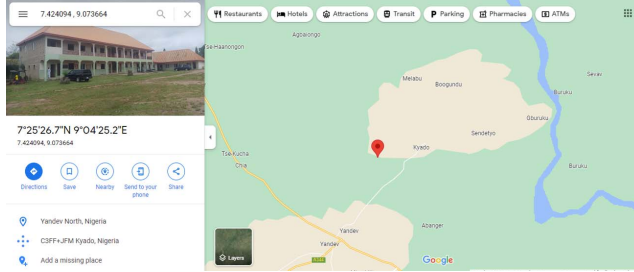
Fig. 5. Sample visualization using calculated Geocoordinate. This was used to show how tweet sentiments can be visualized on the map.

Bounding box of the tweet. It is an area defined by two longitudes and two latitudes. The latitude is a decimal number between -90.0 and 90.0 while the longitude is a decimal number between -180.0 and 180.0. For example, the set of values {7.3045544, 8.961887, 7.5436339, 9.1854407} is a bounding box for a particular location. To get the exact location, we calculate the mean of the latitude and longitude from the bounding box as follows:

$$\text{Latitude: } \frac{7.3045544 + 7.5436339}{2} = 7.424094$$

$$\text{Longitude: } \frac{8.961887 + 9.1854407}{2} = 9.073664.$$

When we searched the latitude and longitude values on Google map, it showed the location as *Yandev, North Nigeria*. See Fig. 5.

The above approach was used to calculate the latitude and longitude of all the tweets to be able to visualize the sentiment of the tweets on the map. The interception of latitude and longitude is called a point in this article. A point is made up of one or more tweets. Thus, a location can contain one or more points. We took the count of sentiments of tweets at a location to identify hotspots. The hotspots of in a location of a province or state, for all the nine countries were visualized on the map with the help of *ArcGIS Online* [53]. The `ArcGIS Online` is a web-based mapping software used to build interactive web maps. See result section for more details.

III. RESULT

In this section, we present the results in two parts. The first part deals with the output of the classification of tweets according to their corresponding sentiments using machine learning. This is to show the validation of the VADER labeled tweets. The second part deals with the output of the analysis, identification, and visualization of the sentiment hotspots on a dashboard using the `ArGIS Online`.

TABLE VII

MODEL PERFORMANCE OF TWEET SENTIMENT CLASSIFICATION

SN	Machine Learning Classifier	Accuracy (%)	Average F1-Score (%)	Average AUC (%)
1	NB	68	65	81
2	LR	78	75	90
3	SVMs	72	67	89
4	DT	65	59	70
5	KNN	58	40	63

A. Result of Tweet Sentiment Classification

In this section, we trained the classifiers (also called models) on tweets with sentiment classes labeled by VADER pretrained model. A summary of the performance of the classifiers is presented in Table VII.

While there is a clear difference in the accuracy scores of the classifiers, the LR model performed better with accuracy score of 78%, average F1-score 75%, and average AUC score of 90%, than the other machine learning classifiers used. The average AUC score of 89% for SVMs is slightly lower than the average AUC score for LR classifier. The accuracy score of 72% for SVMs classifier is slightly higher than the accuracy score of 68% for NB classifier. Similarly, the average F1-score of 67% for SVMs classifier is slightly higher than the 65% for NB classifier. Even though DT and KNN classifiers have the lowest in accuracy scores, average F1-scores, and average AUC scores, respectively, DT classifier was seen to be higher than KNN in all the performance metrics.

The above analysis is an indication that these models can classify tweets according to their sentiment classes. However, the LR classifier proved to be best fit for this type of classification problem given all indicators. One such indicator is that the 46 189 tweets generated a large feature set that was suitable for the LR classifier higher performance.

To further validate the performance of the models, we visualized the ROC metric to evaluate the quality of the multiclassification output, together with the AUC, see Fig. 6. The numbers 0, 1, and 2 represent negative, neutral, and positive sentiment classes, respectively.

From Fig. 6, we can clearly understand how well the machine learning classifiers classified the tweet sentiments. The ROC curve shows the true positive rate against false positive rate. As said in Section II-F, the more the curve aligns toward the upper left corner of the plot, the better the machine learning classifier does at classifying the tweets into various sentiment classes. As shown in Fig. 6(b), the LR classifier does well in classifying the tweets into various sentiment classes, followed by SVM model in Fig. 6(c). However, unlike the NB classifier with an average performance in the classification of the tweet sentiment [Fig. 6(a)], the DT and KNN models performed poorly in in the classification of the tweets into different sentiment classes, see Fig. 6(d) and (e), respectively. As usual, the AUC was used to ascertain how much of the plot is located under the curve. The LR classifier demonstrated to have performed better with a large feature set and multiclass classification.

Since the LR classifier performed better than the other models, understanding the features that influenced performance
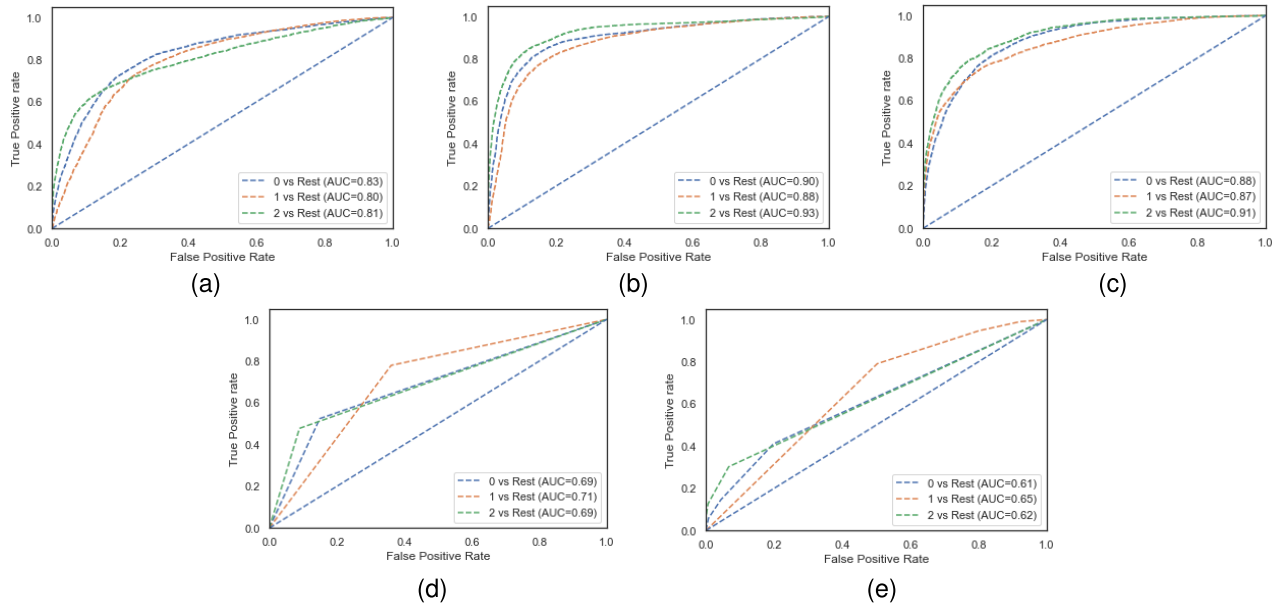
Fig. 6. ROC–AUC for different tweet sentiments classifiers. The quality of the multiclassification output. (a) NB ROC–AUC curve. (b) LR ROC–AUC curve. (c) SVMs ROC–AUC curve. (d) DT ROC–AUC curve. (e) KNN ROC–AUC curve.

of the sentiment classification of the tweets is necessary. We used ELi5 [54], an interpretable machine learning model to visualize the top 20 features in their order of importance that are responsible for the LR model performance. Table VIII shows the weight and features of the top 20 words that influenced the performance of the sentiment classes of the tweets as classified by the LR model.

The idea behind feature importance is to know how the performance metrics (accuracy, precision, recall, and F-measure) behaved with respect to the feature existence. In Table VII, we can see that `best` has the highest weight with $+3.622$ for positive sentiment class, followed by `hell` with $+3.091$ for negative sentiment class, and `bias` with $+2.091$ for neutral sentiment class. This means that the features `best`, `hell`, and `bias` affected the model performance with a probability of $+3.622$, $+3.091$, and $+2.091$ in the classification of tweets to positive, negative, and neutral sentiment classes, respectively. Next, in Section C, we explained tweet sentiment hotspot analysis.

B. Model Performance Comparison

In this section, we used ANOVA test to record the performances of difference models on VADER labeled dataset, Textblob labeled dataset, and AFINN labeled dataset. We performed ten iterations for each model and got different accuracy scores. For each iteration, the dataset is divided into train and test sets. Such that the train and test set are different for each run. Table IX shows the accuracy scores of the models.

The corresponding boxplot showing the distribution of machine learning models (ML model) performance with respect to the annotation tools (base model) is shown in Fig. 7. From the boxplot, we can easily identify the differences between ML models and the annotation tools.

Similarly, Table X shows the outcome of the ANOVA test used to record the significant difference between the
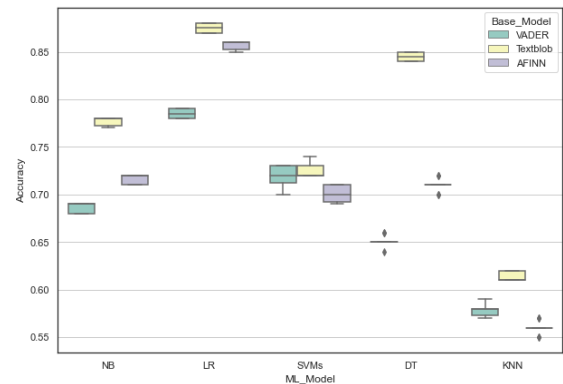

Fig. 7. Distribution of ML models with respect to accuracy.

performance of the machine learning models with respect to different annotation tools used. The ANOVA statistical test takes two hypotheses for the output as follows.

1) *Null Hypothesis H_0:* The performance of the models is equal.
2) *Alternative Hypothesis H_A:* The performance of the models is not equal.

From Table X, *p* values obtained from ANOVA analysis for ML model, Base model, and interaction are statistically significant ($p < 0.05$). This means that there is a sufficient evidence to reject the H_0 that the performance of the models is equal. We conclude that the type of annotation tool and machine learning algorithm can significantly affect the performance of the model.

Next, we visualize the interaction between the annotation tools with the machine learning model from the ANOVA analysis. This is to ascertain the interactive effect of the means of annotation tools and machine learning models. It also helps us to visualize the exact machine learning model that performs best with the annotation tools.

TABLE VIII

LR Model Feature Interpretation Using ELI5. Summarizes the Features Responsible for the Model Performance

SN	Positive Weight	Positive Feature	Negative Weight	Negative Feature	Neutral Weight	Neutral Feature
1	+3.622	best	+3.091	hell	+2.091	bias
2	+3.524	positive	+3.031	worst	-1.463	disaster
3	+3.423	love	+2.940	killed	-1.470	fraud
4	+3.210	wow	+2.874	crisis	-1.479	broke
5	+3.195	happy	+2.843	death	-1.492	flu
6	+3.180	great	+2.555	failed	-1.520	trust
7	+3.021	freedom	+2.529	cancer	-1.538	forced
8	+2.876	amazing	+2.498	killing	-1.559	confidence
9	+2.838	encouraging	+2.467	violence	-1.579	kill
10	+2.802	lmao	+2.451	disaster	-1.631	fatigue
11	+2.764	free	+2.450	poor	-1.635	amazing
12	+2.468	success	+2.418	negative	-1.637	suspended
13	+2.435	beautiful	+2.361	died	-1.663	great
14	+2.428	congratulations	+2.342	fuck	-1.666	awesome
15	+2.367	encourage	+2.328	suspended	-1.711	hell
16	+2.364	loved	+2.327	war	-1.723	best
17	+2.326	super	+2.284	dead	-1.783	congratulations
18	+2.307	ensure	+2.271	worse	-1.834	wow
19	+2.261	justice	+2.252	drop	-1.859	killing
20	+2.260	confidence	-2.259	freedom	-1.954	success

TABLE IX

Comparison of Machine Learning Model Performance on AFINN, VADER, and Textblob Labeled Dataset

VADER NB	LR	SVMs	DT	KNN	Textblob NB	LR	SVMs	DT	KNN	AFINN NB	LR	SVMs	DT	KNN
0.69	0.79	0.72	0.66	0.58	0.78	0.87	0.72	0.85	0.61	0.72	0.86	0.70	0.71	0.56
0.69	0.78	0.72	0.65	0.58	0.78	0.88	0.72	0.85	0.62	0.72	0.85	0.70	0.71	0.56
0.69	0.78	0.73	0.65	0.58	0.78	0.87	0.72	0.84	0.61	0.72	0.86	0.70	0.71	0.56
0.68	0.78	0.73	0.65	0.59	0.77	0.87	0.74	0.85	0.61	0.71	0.86	0.71	0.70	0.56
0.68	0.78	0.73	0.65	0.57	0.78	0.87	0.74	0.85	0.61	0.72	0.85	0.69	0.70	0.56
0.68	0.79	0.71	0.65	0.57	0.78	0.88	0.73	0.84	0.61	0.71	0.85	0.69	0.71	0.55
0.69	0.79	0.71	0.65	0.57	0.78	0.88	0.72	0.84	0.61	0.71	0.86	0.69	0.71	0.55
0.69	0.79	0.72	0.64	0.58	0.77	0.88	0.73	0.84	0.62	0.72	0.86	0.71	0.72	0.56
0.69	0.79	0.70	0.66	0.58	0.77	0.88	0.72	0.84	0.62	0.71	0.86	0.71	0.71	0.57
0.68	0.78	0.73	0.65	0.58	0.78	0.87	0.72	0.85	0.62	0.72	0.86	0.71	0.72	0.57

TABLE X

Summary of ANOVA Test Result

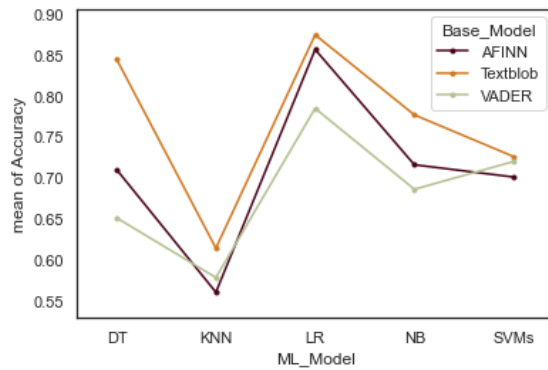|	sum_sq	df	F	p-values
C(ML_Model)	0.988529	4.0	5884.103175	2.347688e-150
C(Base_Model)	0.183409	2.0	2183.444444	1.560926e-103
C(ML_Model): C(Base_Model)	0.121291	8.0	360.984127	3.616289e-87
Residual	0.005670	135.0	NaN	NaN

Fig. 8. Interactive response of the two factors. Visualizes the interaction between the annotation tools with the machine learning model from the ANOVA analysis.

Fig. 8 shows that LR model performed best with the Textblob annotated dataset followed by AFFIN than VADER. Conversely, KNN has the least performance with AFFIN

annotated dataset. Meanwhile, VADER is developed specifically for social media text classification [33]. This was our main motivation for using VADER in this study. Moreover, VADER performed well with LR model in the multiclass classification of large feature set, see Fig. 6.

C. Tweet Sentiment Hotspot Analysis

A typical hotspot areas are concentrations of incidents within a limited geographical area that appear over time. Measuring a hotspot could be complicated. However, there are many statistical techniques designed to identify hotspots, including hierarchical technique, partitioning technique (K-means), density technique, clumping technique, risk-based technique, miscellaneous techniques, and point-based location technique [55]. There is also a two or more combinations called hybrid technique. For instance, the risk-adjusted nearest neighbor hierarchical clustering routine is a risk-based technique with a combination of clumping technique. Similarly, the grouping of partitioning and hierarchical techniques is a hybrid technique called STAC. These statistical techniques are mainly aimed at grouping incidents together into a relatively high or low coherent clusters or concentrations [56], [57].

Given that we worked with latitude and longitude of all the tweets which allowed us to visualize the sources of
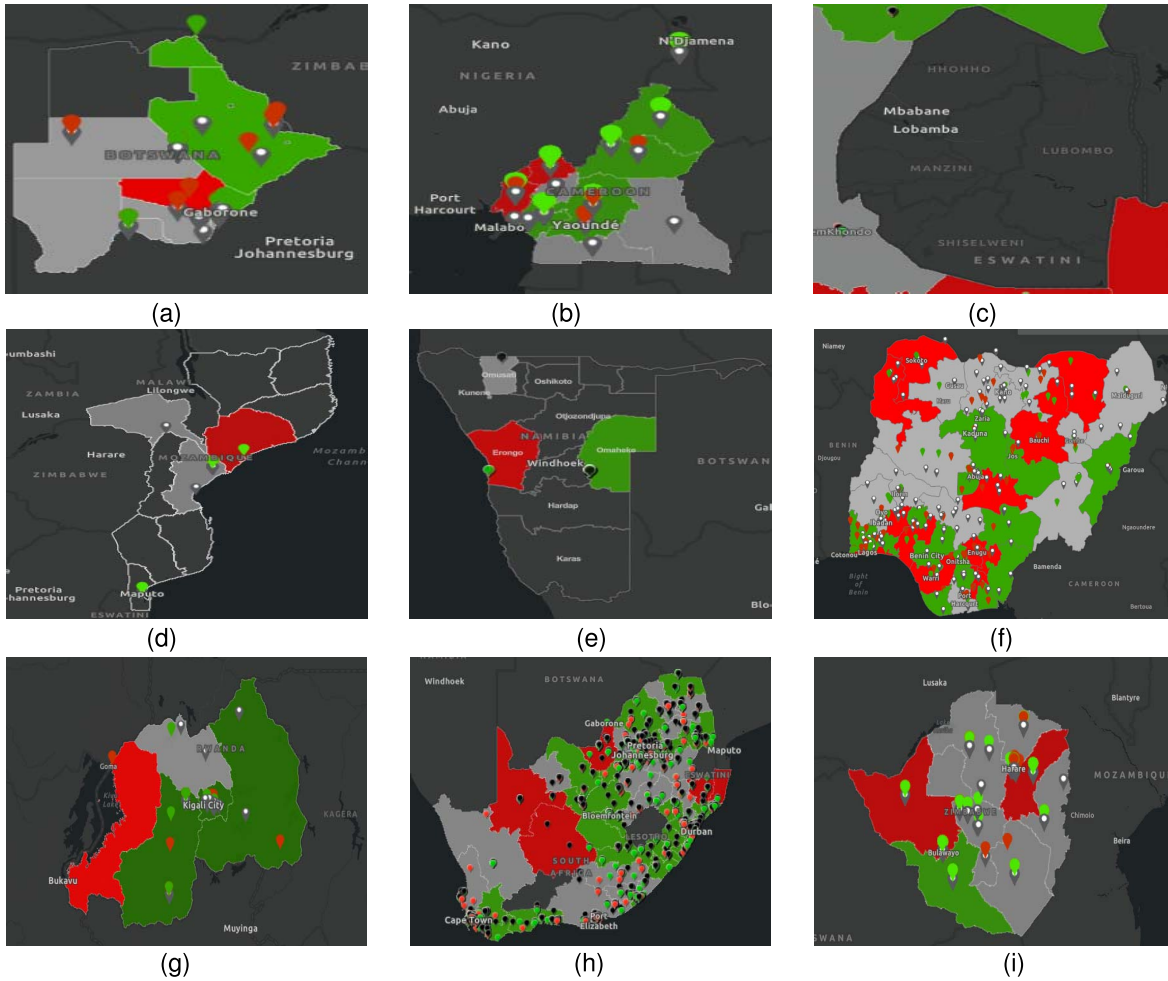
Fig. 9. Identified COVID-19 sentiment hotspots. Visualizes the map showing sentiment hotspots by countries. (a) Botswana. (b) Cameroon. (c) Eswatini. (d) Mozambique. (e) Namibia. (f) Nigeria. (g) Rwanda. (h) South Africa. (i) Zimbabwe.

the tweets as a point on the map, the point-based location technique is best fit for our analysis [58]. The point-based location technique is one of the most intuitive types of cluster involving the number of tweets coming at different locations. The locations with the most number of tweets are referred to as hotspots. The count of each tweet sentiment class in the hotspots is calculated. The sentiment class with the highest count becomes the dominant sentiment in that location, thus, it is identified with green if the sentiment class is positive, red if the sentiment class is negative, and gray if the sentiment class is neutral. To calculate the sentiment count for each sentiment class in a location (S_c), we use the formula in (7) as adapted from [58]

$$S_c = \sum_{j=1}^{n} \sum_{i=1}^{m} X_{i,j} \qquad (7)$$

where i is the count of the sentiment class in a point and m represents the total number of sentiment classes in a point. Similarly, j represents count for each point in a location and n represents the total number of points in a location. The $X_{i,j}$ represents a tweet sentiment class in a point i and location j. $\sum_{i=1}^{m} X_{i,j}$ represents the sum of sentiment classes in a point, while $\sum_{j=1}^{n} \sum_{i=1}^{m} X_{i,j}$ is the sum of each sentiment

class in a location. The outcome of the above process was visualized on a dashboard and the screenshots are shown in Fig. 9. As shown in Fig. 9, the identified sentiment hotspots are colored green, red, and gray to signify the positive, negative, and neutral sentiment dominance in that location for each of the nine countries.

D. Limitation

The Twitter data used for this research only reflects the opinion of Twitter users whose locations are in the selected nine African countries, within the period of discussion. As at November, 2022, there is about 8.46% online adults who use Twitter in Africa [59]. Therefore, this study does not fully represent the opinions of people in the identified African countries, especially regarding COVID-19 vaccines. This study only provided an insightful analysis from the Twitter data to support policy making, management and planning.

Additionally, VADER used in this study is a pretrained model for sentiment classification. It does not have the capacity to properly label figurative language, such as sarcasm, pidgin English, and vernacular. Pretrained models, such as VADER, Textblob, and AFINN cannot replace human's annotation. As the most precise annotation is the one done by

humans irrespective of human bias, subjectivity, and error-proneness. Hence, VADER annotated labels can be used as assistance for human annotators especially in the case of a large dataset.

IV. Discussion

In this section, we discussed the hotspots identified in Fig. 9 for each country. In Botswana, we identified nine locations as hotspot areas for positive, negative, and neutral sentiment classes. These locations include Chobe, Central district, and Kgatleng district, for positive sentiment hotspot. Kweneng district, for negative sentiment hotspots and Kgalagadi district, Ghanzi district, Southern district, and South-east district for neutral sentiment hotspots, respectively, see Fig. 9(a). Moreover, as of July, 2022, 64.4% of the Botswana population is said to be fully vaccinated [6]. We observed that Kweneng district with negative sentiment hotspot is predominantly Christians and the second highest populated district. This could also suggest that religion may have played a major role in influencing citizens' sentiments against vaccination, given that a lot of religious leaders were against the vaccine mandate at that time [10]. Furthermore, the combination of all the locations is identified with neutral sentiment hotspots and the Kweneng district is identified with the negative sentiment hotspot amounted to about 35% of the countries' total population. This could be a contributor as to why Botswana was listed as one of the countries with expired and at risk to expire vaccines because a large number of citizens may have been hesitant toward the vaccine at the time [6].

Similarly, in Cameroon, we identified the South, East, and West regions as neutral sentiment hotspots. The Southwest and Northwest regions were identified as negative sentiment hotspots. While the Littoral, Central, Adamawa, and North regions remain the positive sentiment hotspots, the Extreme North region of the country was not identified as a hotspot from our data, see Fig. 9(b). The suspension of all medical and humanitarian activities in South-West region of Cameroon by the Médecins Sans Frontiéres (MSF) in December 2021 may have contributed to the negative sentiment in the region. The suspension, following the arrest of two of MSF members and collaborators by the Cameroon armed police force, as such, may have played a role in influencing negative sentiments toward the vaccine at that time [60]. Moreover, Southwest and Northwest regions are mostly Christians and Muslims; hence, religion may have also played a role in influencing their sentiments toward vaccination. For instance, 70.7% and 24.4% of the countries' population practice Christianity and Islam, respectively [61]. Additionally, workers in Cameroon rejected to take the vaccine because they claim not to be well informed about the aftereffect of the vaccine; thus, Cameroon being listed as a country with expired vaccines and at risk to expire vaccines is not a surprise [62].

In Eswatini, we identified Manzini region as negative sentiment hotspot from our dataset. The Hhohho region was identified as positive sentiment hotspot. The Lubombo and Shiselweni regions were not identified as any sentiment hotspots [see Fig. 9(c)]. The Manzini region is the country's largest urban center, which is known as the *Hub* of the country.

Although 29.8% of the entire population of the country is said to have been vaccinated [63], the country was, however, listed as a country with expired and at risk to expire vaccines. This could be associated with the influence of the Manzini region being the country's largest urban center, which is identified as negative sentiment hotspot.

Out of the eleven provinces in Mozambique, including Maputo, the administrative region, we identified three regions with sentiment hotspots. These regions include Tete, Sofala, and Zambezia regions. The Tete and Sofala regions were identified as neutral sentiment hotspots. While the Zambezie was identified as negative sentiment hotspot, see Fig. 9(d). The Zambezia region is the second most populous region of Mozambique and is predominantly Muslim religion that is practiced in the region. Together with Nampula region, also in the north-central province, they account for 45% of Mozambique population. These may have influenced the sentiment of people living in the Zambezie region toward the vaccines. As such, Mozambique was listed as a country with expired and at risk to expire vaccines (see Fig. 1). Generally, the study in [64] suggested fear and lack of confidence as contributing factors to vaccine acceptance in Mozambique.

We identified three regions with sentiment hotspots in Namibia. These regions include Omusati for neutral sentiment hotspot, Omaheke for positive sentiment hotspot, and Erongo for negative sentiment hotspot, see Fig. 9(e). The Erongo region is one of the smallest of the 14 regions of Namibia, and is predominately Christian. This may be connected to the high number of negative sentiment in the region as religion has been identified as one of the contributing factors to vaccine hesitancy [65]. The result of this is the low turnout of the population of children and adolescents during vaccination against COVID-19. To date, about 3% of the Erongo Region population is said to have received the vaccine [65], [66]. Consequently, Namibia is among the countries listed to have expired and at risk to expire vaccines as shown in Fig. 1.

In Nigeria, the positive sentiment hotspots were identified in 13 states with the government capital, including Kaduna, Plateau, Adamawa, Benue, Cross River, and Anambra. Others are Imo, Akwa Ibom, Bayelsa, Edo, Ekiti, Ogun, and Abuja the national capital. The neutral sentiment hotspots were identified in 12 states including Zamfara, Katsina, Kano, Jigawa, Niger, Kwara, Oyo, Kogi, Rivers, Taraba, Gombe, and Maiduguri states. Meanwhile, the negative sentiment hotspots were identified in 12 states, namely Osun, Ondo, Delta, Enugu, Abia, Ebonyi, Nasarawa, Bauchi, Yobe, Sokoto, Kebbi, and Lagos, see Fig. 9(f). Nigeria only has 13.7% of its population that is fully vaccinated [67]. This could have led to the high amount of expired vaccines and at risk to expire vaccines in the country as shown in Fig. 1. The combination of high neutral and negative sentiment toward the vaccination may have also contributed to the expired vaccines [68]. Intuitively, one may attribute these to the lack of confidence by the citizens about the effectiveness of the vaccines at that time [69]. In addition, given the strategic nature of Lagos as the largest city in Nigeria and Africa at large, with about 15.4 million population in 2022 [70], the negative sentiment hotspot identified in the state

may have also influenced the low response toward vaccination in the country [68], [69], [71].

The following provinces were identified as sentiment hotspots in Rwanda. Kigali (Umujyi wa Kigali), Eastern (Iburasirazuba), and Northern (Amajyaruguru) for the positive sentiment hotspots. The Southern (Amajyepfo) province was identified as neutral sentiment hotspot, while the Western (Iburengerazuba) province was identified as negative sentiment hotspots, see Fig. 9(g). Despite the fact that about 81.0% of Rwanda population is fully vaccinated [44], [67], the Western province showed high negative sentiment dominance toward vaccination. The population of over 2.4 million and 95% Christians religion practiced in the Western province may have contributed to the negative sentiment dominance toward vaccination [72]. This is because religion has been said to be one of the contributing factors against vaccination in Africa. Citizens tend to express fear of the aftereffect which leads to lack of confidence in taking the vaccine [73].

Furthermore, in South Africa, we identified three provinces with partially negative sentiment dominance, including Northern Cape, Northwest, and Kwazulu-Natal. In Northern Cape province, the Siyanda and Pixley ka Seme district were identified as negative sentiment hotspots. Similarly, the Mafikeng district of the Northwest province was identified as negative sentiment hotspot. Finally, the Zululand and Umkhanyakude districts were also identified as negative sentiment dominance districts in Kwazulu-Natal province. The remaining six provinces were dominated by neutral and positive sentiments, see Fig. 9(h). Despite that South Africa only achieved 32.5% of its population to be fully vaccinated [67], [74], it was not listed among the countries that have expired and at risk to expire vaccines, see Fig. 1. This may suggest that the negative sentiment dominance combined with the neutral sentiment dominance were not influential enough to negatively impact the vaccination uptake in the country at that time [75].

Finally, in Zimbabwe, we identified Matabeleland South and Bulawayo provinces as positive sentiment hotspots. The Midland, Mashonaland West, Mashonaland Central, Masvingo, Manicaland, and Harare provinces were identified as neutral sentiment hotspots. The Matabeleland North and Matabeleland East provinces were identified as negative sentiment hotspots. The 31.6% of Zimbabwe population is fully vaccinated [67] and does not really change the fact that the country is listed as a country with expired vaccines and at risk to expire vaccines, see Fig. 1. This could be the lack of trust in the government and the uncertainty about vaccine effectiveness and safety as expressed by majority of Zimbabweans in [76].

V. Conclusion

Social media is a place where users share their opinions about a subject matter. During the lockdown as a result of the pandemic, social media became the most effective medium of communication for users to express their concerns. As the implementation of vaccination mandate began, social media platform such as Twitter became one of the tools users used to express their opinions about the vaccines and their side effects. These types of opinions generated a lot of missed feelings and concerns. These concerns could lead to a compromise of confidence toward the vaccine which brings about vaccine hesitancy. In Africa, vaccine hesitancy is a major challenge faced by health policymakers in the fight against COVID-19. Given that most tweets are geotagged, clustering them according to their sentiments could help in identifying locations that may likely experience vaccine hesitancy for health policy and planning.

In this research, we collected 70 000 COVID-19 vaccine-related geotagged Twitter posts from nine African countries. The duration is from December 2020 to February 2022. We used VADER to annotate the tweets into three sentiment classes (positive, negative, and neutral). The outputs were validated using machine learning classifiers, including, NB with an accuracy score of 66%, LR (71%), SVMs (65%), DT (61%), and KNN (56%). The average Area under the Curve score of 78%, 85%, 83%, 67%, and 63% was used to evaluate the quality of the multiclassification outputs.

The point-based location technique was used to calculate hotspots by clustering sentiment of these tweets. The green, red, and gray colors were used to visualize the dominance of positive, negative, and neutral sentiments on the map. This means that the location with a green background on the map is a hotspot for positive sentiment. The locations with red and gray background on the map are hotspots for negative and neutral sentiments, respectively. This process of visualizing the sentiment as hotspots on the map was achieved using `ArcGIS Online`.

Therefore, the main contribution of this research is the application of a point-based location technique in the identification and visualization of vaccine hesitancy hotspots from a labeled tweets sentiments dataset. This suggests that discussions on social media can be analyzed to identify hotspots, based on users' sentiments toward vaccination during a disease outbreak. We believe that this could inform health policy in planning and management of vaccine hesitancy in Africa.

Author Contributions

The authors contributed as follows: Blessing Ogbuokiri contributed to the conception, design, analysis, draft, and final revision of this article. Zahra Movahedi Nia contributed to Twitter data collection. Ali Ahmadi, Bruce Mellado, Jiahong Wu, James Orbinski, and Ali Asgary contributed to the supervision of the design analysis. Jude Kong contributed to supervision of the design analysis and article editorial revision.

Conflict of Interest

There is no conflict of interest.

DATA AVAILABILITY STATEMENT

The dataset used in this research is available on request after the necessary approvals. Meanwhile, the dashboard we used to visualize the outcome of this study is accessible at http://acadic.org/. Click Dashboards to view the drop down menu, select COVID-19 Sentiment Dashboard, then, click on any country to view the interactive dashboard.

REFERENCES

[1] H. Harapan et al., "Vaccine hesitancy among communities in ten countries in Asia, Africa, and south America during the COVID-19 pandemic," *Pathogens Global Health*, vol. 116, no. 4, pp. 236–243, 2022, doi: 10.1080/20477724.2021.2011580.

[2] C. C. D. S. Matos, B. A. Gonçalves, and M. T. Couto, "Vaccine hesitancy in the global south: Towards a critical perspective on global health," *Global Public Health*, vol. 17, no. 6, pp. 1087–1098, Jun. 2022, doi: 10.1080/17441692.2021.1912138.

[3] Y. B. Limbu, R. K. Gautam, and L. Pham, "The health belief model applied to COVID-19 vaccine hesitancy: A systematic review," *Vaccines*, vol. 10, no. 6, p. 973, Jun. 2022. [Online]. Available: https://www.mdpi.com/2076-393X/10/6/973

[4] F. Yin, X. Xia, N. Song, L. Zhu, and J. Wu, "Quantify the role of superspreaders opinion leaders on COVID-19 information propagation in the Chinese Sina-microblog," *PLoS ONE*, vol. 16, no. 6, pp. 1–20, 2020.

[5] World Health Organisation. (2022). *WHO Coronavirus (COVID-19) Dashboard*. [Online]. Available: https://covid19.who.int/

[6] World Health Organisation. (2022). *WHO Africa COVID-19 Vaccination Dashboard*. [Online]. Available: shorturl.at/BEJY7

[7] A. Seydou, "Who wants COVID-19 vaccination? in 5 west African countries, hesitancy is high, trust low," Afrobarometer, West African, Tech. Rep. 432, 2021.

[8] M. O. Afolabi et al., "Tracking the uptake and trajectory of COVID-19 vaccination coverage in 15 West African countries: An interim analysis," *BMJ Global Health*, vol. 6, no. 12, Dec. 2021, Art. no. e007518. [Online]. Available: https://gh.bmj.com/content/6/12/e007518

[9] L. Périères, V. Séror, S. Boyer, C. Sokhna, and P. Peretti-Watel, "Reasons given for non-vaccination and under-vaccination of children and adolescents in sub-Saharan Africa: A systematic review," *Hum. Vaccines Immunotherapeutics*, vol. 18, no. 5, Nov. 2022, Art. no. 2076524, doi: 10.1080/21645515.2022.2076524.

[10] S. Kabakama et al., "Commentary on COVID-19 vaccine hesitancy in sub-Saharan Africa," *Tropical Med. Infectious Disease*, vol. 7, no. 7, p. 130, Jul. 2022. [Online]. Available: https://www.mdpi.com/2414-6366/7/7/130

[11] C. S. Wiysonge et al., "Vaccine hesitancy in the era of COVID-19: Could lessons from the past help in divining the future?" *Hum. Vaccines Immunotherapeutics*, vol. 18, no. 1, pp. 1–3, Jan. 2022, doi: 10.1080/21645515.2021.1893062.

[12] S. Kwok, S. Vadde, and G. Wang, "Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis," *J. Med. Internet Res.*, vol. 23, no. 4, pp. 611–621, 2021.

[13] J. Silva, J. Bratberg, and V. Lemay, "COVID-19 and influenza vaccine hesitancy among college students," *J. Amer. Pharmacists Assoc.*, vols. 61, pp. 709–714, May 2021, doi: 10.1016/j.japh.2021.05.009.

[14] C. Udanor, S. Aneke, and B. Ogbuokiri, "Determining social media impact on the politics of developing countries using social network analytics," *Electronic Library Inf. Syst.*, vol. 50, no. 4, pp. 1–26, 2016.

[15] S. Tasnim, M. M. Hossain, and H. Mazumder, "Impact of rumors and misinformation on COVID-19 in social media," *J. Preventive Med. Public Health*, vol. 53, no. 3, pp. 171–174, 2020.

[16] L. Al-Uqdah, F. A. Franklin, C.-C. Chiu, and B. N. Boyd, "Associations between social media engagement and vaccine hesitancy," *J. Community Health*, vol. 47, no. 4, pp. 577–587, Aug. 2022, doi: 10.1007/s10900-022-01081-9.

[17] N. Puri, E. A. Coomes, H. Haghbayan, and K. Gunaratne, "Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases," *Hum. Vaccines Immunotherapeutics*, vol. 16, no. 11, pp. 2586–2593, Nov. 2020.

[18] F. Ennab et al., "Implications of social media misinformation on COVID-19 vaccine confidence among pregnant women in Africa," *Clin. Epidemiology Global Health*, vol. 14, Mar. 2022, Art. no. 100981. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213398422000215

[19] S. E. Clark, M. C. Bledsoe, and C. J. Harrison, "The role of social media in promoting vaccine hesitancy," *Current Opinion Pediatrics*, vol. 34, no. 2, pp. 156–162, 2022, doi: 10.1097/MOP.0000000000001111.

[20] S. A. Thorakkattil, S. Abdulsalim, M. S. Karattuthodi, M. K. Unnikrishnan, M. Rashid, and G. Thunga, "COVID-19 vaccine hesitancy: The perils of peddling science by social media and the lay press," *Vaccines*, vol. 10, no. 7, p. 1059, Jun. 2022. [Online]. Available: https://www.mdpi.com/2076-393X/10/7/1059

[21] W. Aljedaani et al., "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry," *Knowl.-Based Syst.*, vol. 255, Nov. 2022, Art. no. 109780. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S09507051 22009017

[22] G. Chandrasekaran and J. Hemanth, "Deep learning and TextBlob based sentiment analysis for coronavirus (COVID-19) using Twitter data," *Int. J. Artif. Intell. Tools*, vol. 31, no. 1, Feb. 2022, Art. no. 2250011. [Online]. Available: https://www.worldscientific.com/doi/epdf/10.1142/S0218213022500117

[23] N. Mansouri, M. Soui, I. Alhassan, and M. Abed, "TextBlob and BiLSTM for sentiment analysis toward COVID-19 vaccines," in *Proc. 7th Int. Conf. Data Sci. Mach. Learn. Appl. (CDMA)*, Mar. 2022, pp. 73–78.

[24] M. Gautam, N. Sahai, A. S. Yadav, and M. Tomar, "Sentiment analysis about COVID-19 vaccine on Twitter data: Understanding public opinion," in *Proc. 6th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2022, pp. 1487–1493.

[25] F. H. Rachman and B. S. Rintyarna, "Sentiment analysis of Madura tourism in new normal era using text blob and KNN with hyperparameter tuning," in *Proc. Int. Seminar Mach. Learn., Optim., Data Sci. (ISMODE)*, Jan. 2022, pp. 23–27.

[26] E. Christodoulaki, M. Kampouridis, and P. Kanellopoulos, "Technical and sentiment analysis in financial forecasting with genetic programming," in *Proc. IEEE Symp. Comput. Intell. Financial Eng. Econ. (CIFER)*, May 2022, pp. 1–8.

[27] M. S. Hossain and M. F. Rahman, "Customer sentiment analysis and prediction of insurance products reviews using machine learning approaches," *FIIB Bus. Rev.*, Aug. 2022, doi: 10.1177/23197145 221115793.

[28] S. Hossain, K. Uddin, K. H. Szechenyi, and F. Rahman, "User sentiment analysis and review rating prediction for the blended learning platform app," in *Applying Data Science and Learning Analytics Throughout a Learner's Lifespan*. Hershey, PA, USA: IGI Global, 2022, pp. 113–132.

[29] A. Mardjo and C. Choksuchat, "HyVADRF: Hybrid VADER–random forest and GWO for bitcoin tweet sentiment analysis," *IEEE Access*, vol. 10, pp. 101889–101897, 2022.

[30] J. Al-Garaady and M. Mahyoob, "Public sentiment analysis in social media on the SARS-COV-2 vaccination using VADER lexicon polarity," *Humanities Educ. Sci. J.*, pp. 609–1591, Jul. 2022. [Online]. Available: https://ssrn.com/abstract=3647564

[31] S. Biswas and S. Ghosh, "Drug usage analysis by VADER sentiment analysis on leading countries," *Mapana J. Sci.*, vol. 21, no. 3, pp. 609–1591, 2022.

[32] P. S. K. Chaitanya, K. Kasoju, S. Bhutada, and B. N. U. Chandrika, "Ranking companies based on VADER sentiment analysis," in *Proc. 1st Int. Conf. Artif. Intell. Trends Pattern Recognit. (ICAITPR)*, Mar. 2022, pp. 1–4.

[33] A. M. Alsharhan, H. R. Almansoori, S. Salloum, and K. Shaalan, "Three Mars missions from three countries: Multilingual sentiment analysis using VADER," in *Proc. 8th Int. Conf. Adv. Mach. Learn. Technol. Appl.*, A. E. Hassanien, R. Y. Rizk, V. Snasel, and R. F. Abdel-Kader, Eds. Cham, Switzerland: Springer, 2022, pp. 371–387.

[34] VIPER Group. (2022). *COVID-19 Vaccine Tracker*. [Online]. Available: https://covid19.trackvaccines.org/vaccines/approved/

[35] F. Li et al. Accessed: Jun. 1, 2022. *What's New in Pandas 1.2.4*. [Online]. Available: https://pandas.pydata.org/pandas-docs/stable/whatsnew/v1.2.4.html

[36] *Tweet-preprocessor 0.6.0*. Accessed: Jun. 1, 2022. [Online]. Available: https://pypi.org/project/tweet-preprocessor/

[37] *Natural Language Toolkit*. Accessed: Jun. 1, 2022. [Online]. Available: https://www.nltk.org/

[38] M. Honnibal, "spaCY 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing," *Sentometrics Res.*, vol. 1, no. 1, 2017, Art. no. 25862593, Accessed: Feb. 28, 2022. [Online]. Available: https://sentometrics-research.com/publication/72/

[39] B. Aditya. *Industrial Strenght Natural Language Processing in Python*. Accessed: Jun. 1, 2022. [Online]. Available: https://spacy.io/

[40] T. Pano and R. Kashef, "A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19," *Big Data Cognit. Comput.*, vol. 4, no. 4, p. 33, Nov. 2020. [Online]. Available: https://www.mdpi.com/2504-2289/4/4/33

[41] A. Borg and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Exp. Syst. Appl.*, vol. 162, Dec. 2020, Art. no. 113746. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420305704

[42] B. Ogbuokiri et al., "Public sentiments towards COVID-19 vaccines in South African cities: An analysis of Twitter posts," *Frontiers Public Health*, vol. 10, Feb. 2022, Art. no. 987376. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpubh.2022.987376/full

[43] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002. [Online]. Available: https://www.jair.org/index.php/jair/article/view/10302

[44] A. Binagwaho, K. Mathewos, and S. Davis, "Equitable and effective distribution of the COVID-19 vaccines—A scientific and moral obligation," *Int. J. Health Policy Manag.*, vol. 11, pp. 100–102, Apr. 2021. [Online]. Available: https://www.ijhpm.com/article_4038.html

[45] N. Kewsuwun and S. Kajornkasirat, "A sentiment analysis model of agritech startup on Facebook comments using naive Bayes classifier," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 2829–2838, 2022.

[46] K. Gulati, S. S. Kumar, R. S. K. Boddu, K. Sarvakar, D. K. Sharma, and M. Z. M. Nomani, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic," *Mater. Today, Proc.*, vol. 51, pp. 38–41, Jan. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214785321032843

[47] T. H. J. Hidayat, Y. Ruldeviyani, A. R. Aditama, G. R. Madya, A. W. Nugraha, and M. W. Adisaputra, "Sentiment analysis of Twitter data related to Rinca island development using Doc2Vec and SVM and logistic regression as classifier," *Proc. Comput. Sci.*, vol. 197, pp. 660–667, Jan. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187705092102411X

[48] R. Vasista. (2018). *Sentiment Analysis Using SVM*. Accessed: Jun. 11, 2022. [Online]. Available: https://medium.com/vasista/sentiment-analysis-using-SVM-338d418e3ff1

[49] J. Ritanshi, B. Seema, and S. Seemu, "Sentiment analysis of COVID-19 tweets by machine learning and deep learning classifiers," in *Advances in Data and Information Sciences* (Lecture Notes in Networks and Systems), vol. 318. Berlin, Germany: Springer, 2022, pp. 329–339.

[50] P. Bachchu, G. Anchita, D. Tanushree, D. A. Debashri, and B. Somnath, "A comparative study on sentiment analysis influencing word embedding using SVM and KNN," in *Cyber Intelligence and Information Retrieval* (Lecture Notes in Networks and Systems), vol. 291. Berlin, Germany: Springer, 2022, pp. 199–211.

[51] E. Zuccarelli. (2020). *Performance Metrics in Machine Learning—Part 1: Classification*. Accessed: Jul. 23, 2022. [Online]. Available: https://towardsdatascience.com/performance-metrics-in-machine-learning-part-1-classification-6c6b8d8a8c92

[52] G. Obaido et al., "An interpretable machine learning approach for hepatitis B diagnosis," *Appl. Sci.*, vol. 12, no. 21, p. 11127, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/21/11127

[53] U. Hague, C. Follett, J. Schwieterman, and C. S. Smith, "The development of COVID-19 data dashboards in illinois: Interactive tools for municipalities and public health professionals," *Illinois Municipal Policy J.*, vol. 6, no. 1, pp. 1–22, 2021.

[54] A. Fan, Y. Jernite, E. Perez, D. Grangier, J. Weston, and M. Auli, "ELI5: Long form question answering," 2019, *arXiv:1907.09190*.

[55] H. Wang, H. Zou, Y. Yue, and Q. Li, "Visualizing hot spot analysis result based on mashup," in *Proc. Int. Workshop Location Based Social Netw.*, Nov. 2009, pp. 45–58.

[56] K. P. Vadrevu et al., "Hotspot analysis of vegetation fires and intensity in the Indian region," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 1, pp. 224–238, Feb. 2013.

[57] Y.-P. Lin, H.-J. Chu, C.-F. Wu, T.-K. Chang, and C.-Y. Chen, "Hotspot analysis of spatial environmental pollutants using kernel density estimation and geostatistical techniques," *Int. J. Environ. Res. Public Health*, vol. 8, no. 1, pp. 75–88, Dec. 2010. [Online]. Available: https://www.mdpi.com/1660-4601/8/1/75

[58] Y. Liao, "Hot spot analysis of tourist attractions based on stay point spatial clustering," *J. Inf. Process. Syst.*, vol. 16, no. 14, pp. 750–759, 2020. [Online]. Available: https://koreascience.kr/article/JAKO202026964745203.page

[59] S. Ghai, L. Magis-Weinberg, M. Stoilova, S. Livingstone, and A. Orben, "Social media and adolescent well-being in the global south," *Current Opinion Psychol.*, vol. 46, Aug. 2022, Art. no. 101318. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352250X22000288

[60] A. P. Marín and R. B. Ali, "Guilt by association: Restricting humanitarian assistance in the name of counterterrorism," *Int. Rev. Red Cross*, vol. 103, nos. 916–917, pp. 539–554, Apr. 2021.

[61] J. N. Dinga, L. K. Sinda, and V. P. K. Titanji, "Assessment of vaccine hesitancy to a COVID-19 vaccine in cameroonian adults and its global implication," *Vaccines*, vol. 9, no. 2, p. 175, Feb. 2021. [Online]. Available: https://www.mdpi.com/2076-393X/9/2/175

[62] A. Amani et al., "The first 30 days of COVID-19 vaccination in cameroon: Achievements, challenges, and lessons learned," *Pan Afr. Med. J.*, vol. 14, no. 12, pp. 1–14, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9146659/

[63] S. Padidar, S.-M. Liao, S. Magagula, T. A. M. Mahlaba, N. M. Nhlabatsi, and S. Lukas, "Assessment of early COVID-19 compliance to and challenges with public health and social prevention measures in the Kingdom of Eswatini, using an online survey," *PLoS ONE*, vol. 16, no. 6, Jun. 2021, Art. no. e0253954, doi: 10.1371/journal.pone.0253954.

[64] J. Dula et al., "COVID-19 vaccine acceptability and its determinants in mozambique: An online survey," *Vaccines*, vol. 9, no. 8, p. 828, Jul. 2021. [Online]. Available: https://www.mdpi.com/2076-393X/9/8/828

[65] L. Shiweva and T. H. Musa, "COVID-19 in Southern Africa: Current scenario in Namibia," *South Afr. J. Public Health*, vol. 4, no. 3, p. 69, 2020. [Online]. Available: http://www.shsjournal.org/index.php/shsj/article/view/126

[66] A. Mare and A. Munoriyarwa, "Guardians of truth? Fact-checking the 'disinfodemic' in Southern Africa during the COVID-19 pandemic," *J. Afr. Media Stud.*, vol. 14, no. 1, pp. 63–79, Mar. 2022, doi: 10.1386/jams_00065_1.

[67] E. Mathieu et al., "A global database of COVID-19 vaccinations," *Nature Hum. Behav.*, vol. 5, no. 7, pp. 947–953, May 2021. [Online]. Available: https://www.nature.com/articles/s41562-021-01122-8#citeas

[68] O. Olu-Abiodun, O. Abiodun, and N. Okafor, "COVID-19 vaccination in Nigeria: A rapid review of vaccine acceptance rate and the associated factors," *PLoS ONE*, vol. 17, no. 5, May 2022, Art. no. e0267691. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0267691

[69] H. Adedeji-Adenola, O. A. Olugbake, and S. A. Adeosun, "Factors influencing COVID-19 vaccine uptake among adults in Nigeria," *PLoS ONE*, vol. 17, no. 2, Feb. 2022, Art. no. e0264371. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0264371

[70] D. D. Sasu. *Population of Lagos, Nigeria 2000–2035*. Accessed: Aug. 19, 2022. [Online]. Available: https://www.statista.com/statistics/1308467/population-of-lagos-nigeria/#:~:text=In%202022%2C%20roughly%2015.4%20million,the%20largest%20city%20in%20Africa

[71] C. Wonodi et al., "Conspiracy theories and misinformation about COVID-19 in Nigeria: Implications for vaccine demand generation communications," *Vaccine*, vol. 40, no. 13, pp. 2114–2121, Mar. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264410X22001268

[72] World Population Review. *Rwanda Population 2022 (Live)*. Accessed: Aug. 19, 2022. [Online]. Available: https://worldpopulationreview.com/countries/rwanda-population

[73] A. Binagwaho and K. Mathewos, "Rwanda's success in rolling out its COVID-19 vaccination campaign is a lesson to us all," *BMJ*, vol. 377, Apr. 2022. [Online]. Available: https://www.bmj.com/content/377/bmj.o881.full

[74] S. A. Madhi et al., "Population immunity and COVID-19 severity with omicron variant in South Africa," *New England J. Med.*, vol. 386, no. 14, pp. 1314–1326, 2022, doi: 10.1056/NEJMoa2119658.

[75] S. Cooper, H. Van Rooyen, and C. S. Wiysonge, "COVID-19 vaccine hesitancy in South Africa: How can we maximize uptake of COVID-19 vaccines?" *Exp. Rev. Vaccines*, vol. 20, no. 8, pp. 921–933, Aug. 2021, doi: 10.1080/14760584.2021.1949291.

[76] P. T. Mundagowa, S. N. Tozivepi, E. T. Chiyaka, F. Mukora-Mutseyekwa, and R. Makurumidze, "Assessment of COVID-19 vaccine hesitancy among Zimbabweans: A rapid national survey," *PLoS ONE*, vol. 17, no. 4, pp. 1–17, Apr. 2022, doi: 10.1371/journal.pone.0266724.

Blessing Ogbuokiri received the Ph.D. degree in computer science from the University of the Witwatersrand, Johannesburg, South Africa, in 2020.

He is currently a Post-Doctoral Researcher with the Africa Canada Artificial Intelligence and Data Innovation Consortium (ACADIC) Laboratory, Department of Mathematics and Statistics, York University, Toronto, ON, Canada. His recent research interests include machine learning for health, NLP, data science for social good and social media computing.

Dr. Ogbuokiri has received several academic awards and research grants. He is a recipient of the Dahdaleh Institute Seed Grant for Critical Social Science Perspectives in Global Health Research.

Ali Ahmadi (Member, IEEE) is currently with the Faculty of Computer Engineering, K. N. Toosi University, Tehran, Iran. His research interests include distributed computing, semantic data mining, big data, and artificial intelligence. His current project is Improvement of National Image Search Engine Kavosh.
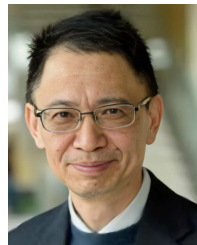
Zahra Movahedi Nia received the Ph.D. degree in computer engineering from the University of Isfahan, Tehran, Iran, in 2020.

She is currently a Post-Doctoral Researcher with the Africa Canada Artificial Intelligence and Data Innovation Consortium (ACADIC) Laboratory, York University, Toronto, ON, Canada. Her research interests include machine learning, deep learning, and data analytics.

Bruce Mellado is currently the Co-President of the Africa Canada Artificial Intelligence and Data Innovation Consortium (ACADIC), York University, Toronto, ON, Canada, and a member of the Gauteng Premier's COVID-19 Advisory Committee, where he leads work on predicting and forecasting the dynamics of COVID-19. He is an Internationally Acclaimed, B1-Rated Researcher of the National Research Foundation of South Africa. He is an Expert in artificial intelligence.

Mr. Mellado was a recipient of several awards and fellowships.

Jiahong Wu was a CRC Partner of the IDRC-CRC International Research Chair, and has led bilateral projects with Scientists in (Health Informatics), Brazil, (Vector-Borne Diseases Modeling), India, (Zika Modeling), South and Latin American countries. He is currently a Tier I CRC and NSERC/Sanofi Industrial Research Chair, an expert in mathematical modeling and data analytics, and the Director of the Laboratory for Industrial and Applied Mathematics (LIAM), the Laboratory of Mathematics for Public Health (MfPH), and the NSERC EIDM MfPH Network.

James Orbinski has led multiple humanitarian responses to Complex Humanitarian Emergencies and was an International President of Doctors Without Borders (MSF) when it was awarded the 1999 Nobel Peace Prize. He is the Director of the Dahdaleh Institute for Global Health Research (DIGHR), York University, Toronto, ON, Canada. He is a leading scholar in acute global health and humanitarian emergencies, disaster early warning systems, and modeling global-warming related disasters and health impacts.

Ali Asgary is the Executive Director of York University's Advanced Disaster, Emergency and Rapid-response Simulation (ADERSIM) and the Director of CIFAL York (created by UNITAR, York University, Toronto, ON, Canada, and York Region for leaders, executives, and civil society training). He brings world-leading expertise in disaster risk assessment and modeling, agent-based simulation, integrated AI and spatio-temporal analysis, and urban and industrial risks at national and international levels.

Jude Kong is the Director of the Africa–Canada Artificial Intelligence and Data Innovation Consortium (ACADIC), York University, Toronto, ON, Canada. He also leads two other networks of researchers that are designing early warning frameworks for emerging infectious disease outbreaks: Mathematics for Public Health (MfPH) and One Health Modeling Network for Emerging Infectious Diseases (OMNI). He is an expert in mathematical modeling, artificial intelligence, data science, infectious disease modeling and mathematics education. His principal research program focuses on the use of quantitative methods to improve decision-making for epidemic and pandemic prevention, preparedness and response.

Mr. Kong won the York Research Leader Award in 2020. He was spotlighted among Canadian Innovation Research Leaders 2021(https://researchinfosource.com/pdf/CIL2021.pdf) for his work with ACADIC in 2021. In 2022, he was spotlighted as a Change Maker by People of York University for his work in helping others learn mathematical concepts and encouraging them to find their passion and achieve more than they thought was possible (https://www.yorku.ca/peopleofyu/2022/02/18/jude-kong-faculty/).