



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original Research

Dynamic assessment of the COVID-19 vaccine acceptance leveraging social media data

Lingyao Li^{a,1}, Jayan Zhou^{b,c,1}, Zihui Ma^a, Michelle T. Bensi^a, Molly A. Hall^{b,d,e}, Gregory B. Baecher^{a,*}

^a Department of Civil and Environmental Engineering, A. James Clark School of Engineering, University of Maryland, College Park, MD, USA

^b Department of Veterinary and Biomedical Sciences, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA, USA

^c Department of Statistics, Eberly College of Science, The Pennsylvania State University, University Park, PA, USA

^d The Huck Institutes of the Life Science, The Pennsylvania State University, University Park, PA, USA

^e Penn State Cancer Institute, The Pennsylvania State University, University Park, PA, USA



ARTICLE INFO

Keywords:

COVID-19 vaccine

Vaccine acceptance

Social media

Natural language processing

Text classification

ABSTRACT

Vaccination is the most effective way to provide long-lasting immunity against viral infection; thus, rapid assessment of vaccine acceptance is a pressing challenge for health authorities. Prior studies have applied survey techniques to investigate vaccine acceptance, but these may be slow and expensive. This study investigates 29 million vaccine-related tweets from August 8, 2020 to April 19, 2021 and proposes a social media-based approach that derives a vaccine acceptance index (VAI) to quantify Twitter users' opinions on COVID-19 vaccination. This index is calculated based on opinion classifications identified with the aid of natural language processing techniques and provides a quantitative metric to indicate the level of vaccine acceptance across different geographic scales in the U.S. The VAI is easily calculated from the number of positive and negative Tweets posted by a specific users and groups of users, it can be compiled for regions such as counties or states to provide geospatial information, and it can be tracked over time to assess changes in vaccine acceptance as related to trends in the media and politics. At the national level, it showed that the VAI moved from negative to positive in 2020 and maintained steady after January 2021. Through exploratory analysis of state- and county-level data, reliable assessments of VAI against subsequent vaccination rates could be made for counties with at least 30 users. The paper discusses information characteristics that enable consistent estimation of VAI. The findings support the use of social media to understand opinions and to offer a timely and cost-effective way to assess vaccine acceptance.

1. Introduction

The coronavirus disease 2019 (COVID-19) is a novel communicable disease. Along with non-pharmaceutical interventions to reduce human-to-human transmissions, the COVID-19 vaccine is the primary strategy to help contain infection, decrease morbidity, and reduce loss of life. The U.S. began its first vaccine distribution on December 14, 2020, and more than 172.2 million Americans (52.4% population) were fully vaccinated as of August 26, 2021 [1]. However, vaccine hesitancy has been an obstacle to reaching herd immunity [2].

Government agencies and health authorities need information to

assess and respond to public sentiment about vaccination, both acceptance and hesitancy. For simplicity, the term acceptance is used to encompass both acceptance and hesitancy. Prior studies have estimated vaccine hesitancy by asking respondents' willingness to receive vaccines [3–5]. Early survey-based studies used general wording that focused on certain groups to formulate an assessment for adolescent and childhood vaccine hesitancy [6,7]. With a deeper understanding of vaccine hesitancy, recent studies have integrated more aspects [4,5] and collected responses through a variety of techniques, such as online tools [8], telephone interviews [9], and systematic literature reviews [10].

Among these survey-based models, the “3C” (confidence,

* Corresponding author.

E-mail addresses: lli12316@umd.edu (L. Li), jpz5091@psu.edu (J. Zhou), zma88@terpmail.umd.edu (Z. Ma), mbensi@umd.edu (M.T. Bensi), mah546@psu.edu (M.A. Hall), gbaecher@umd.edu (G.B. Baecher).

¹ These two authors contributed equally to this work.

<https://doi.org/10.1016/j.jbi.2022.104054>

Received 26 September 2021; Received in revised form 5 March 2022; Accepted 12 March 2022

Available online 21 March 2022

1532-0464/© 2022 Elsevier Inc. All rights reserved.

complacency, convenience) model and the Working Group Determinants of Vaccine Hesitancy Matrix were widely applied, both of which were developed by the Strategic Advisory Group of Experts (SAGE) at the World Health Organization (World Health Organization (WHO) SAGE, 2014). The “3C” model focuses on measuring the “3C’s lens” and identifies barriers within the community. In contrast, the Vaccine Hesitancy Matrix emphasizes external factors (e.g., culture, society, environment, and health aspects) derived from literature reviews and expert opinions. Other commonly applied indices include, for high-income countries, the Global Vaccine Confidence Index [12] and the Vaccine Hesitancy Scale [13]; and, for the low- and middle-income countries, the Caregiver Vaccine Acceptance Scale [14].

In the context of COVID-19, recent studies have demonstrated the usefulness of survey tools to measure the acceptance of vaccination [15,16] and to understand its driving factors. In particular, the U.S. Centers for Disease Control and Prevention (CDC) applied the Household Pause Survey (HPS) to evaluate vaccine hesitancy [17]. They used survey results to estimate state-level hesitancy rates and then applied a “downscaling” method to predict county-level hesitancy rates based on the Census Bureau’s 2019 American Community Survey (ACS) [17]. The final result for each county is an estimate of the percentage of the population that may be vaccine-hesitant.

Survey-based methods have also been applied to assist the assessment of vaccine acceptance [5,18–20]. These may target specific groups (e.g., parents [7,21], healthcare providers [22]) and specific places. Therefore, these methods may provide limited insights on broad public perceptions of vaccine acceptance at different geographical scales. Developing survey-based indices for a wide application may demand a commitment of long-term investigation and a large effort to collect representative responses.

Collecting information from social media platforms, in contrast, can cover large geographical areas at comparatively low cost. Although these data may be of lower fidelity than survey questions, they can be obtained almost instantaneously. Throughout the COVID-19 pandemic, social media have been a forum for vaccine-related public discourse where opinions regarding the COVID-19 vaccines circulate rapidly [23]. Opinions expressed on social media provide data resources that can be leveraged to rapidly assess vaccine acceptance.

Existing studies using social media data to assess opinions related to vaccines mainly evaluate vaccine-related contents through semantic analysis or sentiment analysis [24–26]. They have captured online opinions (e.g., pro- or anti-vaccine campaigns) to help health organizations more readily recognize prevalent vaccination messages [27–29]. Other studies have discussed the role of social media in enhancing vaccine confidence [30,31]. For example, Rosen et al [31] found a positive correlation between social media engagement and HPV vaccine awareness. The study suggested that enhancing the connection between health organizations and the public through social media can help disseminate vaccine information strategically.

Studies that attempt to estimate vaccine acceptance using social media data are less common than survey studies. Piedrahita-Valdés et al. [32] applied sentiment analysis to estimate global vaccine hesitancy using Twitter data for all types of vaccines from 2011 to 2019. They found that the percentage of neutral tweets showed a decreasing tendency, while the percentage positive and negative tweets increased over time. Johnson et al. [33] investigated 100 million Facebook users expressing opinions regarding vaccination to study the evolution of pro- and anti-vaccine clusters from February to October 2019. To classify the clusters, they manually reviewed Facebook pages and identified whether each page was part of a pro- or anti-vaccination cluster. The study found that anti-vaccination clusters managed to become highly entangled and dominated in Facebook networks.

The broad set of studies seeking to leverage social media have demonstrated the potential of these data to investigate online messages regarding vaccine acceptance. However, most prior studies have applied sentiment and other textual analysis to indicate vaccine acceptance

qualitatively, but not to estimate acceptance quantitatively, and across different geographical regions. Further, there has been little discussion of the reliability of social media data in assessing vaccine acceptance.

To fill the research gaps, the present study explores the utility of social media data to render quick indication of COVID-19 vaccine acceptance. The primary objectives are to,

1. Explore whether social media is a reliable indicator of vaccine acceptance as validated through later vaccination rates in the context of COVID-19.
2. Discuss specific circumstances in which social media data generate a reliable estimate of vaccine acceptance (e.g., length of time to collect data, number of users to include, granularity of the analysis).

The study addresses the above two objectives via the analysis of Twitter data related to COVID-19 vaccine in the U.S. Specifically, the study proposes a vaccine acceptance index (VAI) to quantify online opinions towards vaccination. The study includes temporal and spatial analysis to validate the reliability of conclusions from social media data against vaccination rates published by official channels. The intent is to monitor vaccine acceptance across different geographical areas with rapidity, broad spatial coverage, and cost-effectiveness.

2. Material and methods

2.1. Study design

The concepts of acceptance, hesitancy, and uptake appear in a breadth of vaccine literature. While the specific definitions used in a study may differ, the present study uses the following definitions. *Vaccine acceptance* is taken as the individual or group decision to accept or refuse when presented with an opportunity to vaccinate [34]. *Vaccine hesitancy* refers to situations where people have doubts or concerns toward vaccinations, without referring to actual vaccine receipt [34]. *Vaccine uptake* refers to the proportion of a population that has received a specific vaccine. Unlike vaccine uptake which is an objective measure, vaccine acceptance and vaccine hesitancy consider people’s willingness to accept or refuse when presented the opportunity to vaccinate [35].

The research framework is illustrated in Fig. 1. It starts with data preparation, which involves collecting English-language tweets containing the keyword “vaccine” and appearing from August 9, 2020, to April 18, 2021. For consistency, this search word choice was maintained throughout the study window even if additional words and phrases began to be used later in the study. For example, the word “vaccination,” might also be contained in tweets expressing people’s opinions towards COVID-19 vaccine acceptance. However, “vaccination” is typically used to refer to receiving or administering the vaccine and the COVID-19 vaccination efforts did not begin until December 14, 2020 [1] (approximately halfway through the study window).

To develop the training set, a set of the 15,000 most frequently occurring tweets were sampled from the tweet pool. These were manually labeled (i.e., classified into one of three categories) by the authors, as documented in Section 2.2. Then, a random sample of 2,000 unique tweets (distinct from the training tweet samples) was selected and manually labeled (using the same criteria for labeling as applied to the training samples) to build the testing set. Following the creation of the training and testing sets, a text augmentation technique was used to balance the training set, generating a larger augmented dataset with an approximately balanced sample distribution between three label categories, as described in Section 2.2.4.

Text classification pipelines were developed to classify a tweet into one of three categories of opinion toward the COVID-19 vaccine: positive, negative, or unrelated. The construction of the text classification pipelines is described in Section 2.3. The initial training set and the augmented training set were used to train nine candidate text classification pipelines (including two sentiment tools). Recall, Precision, F1-

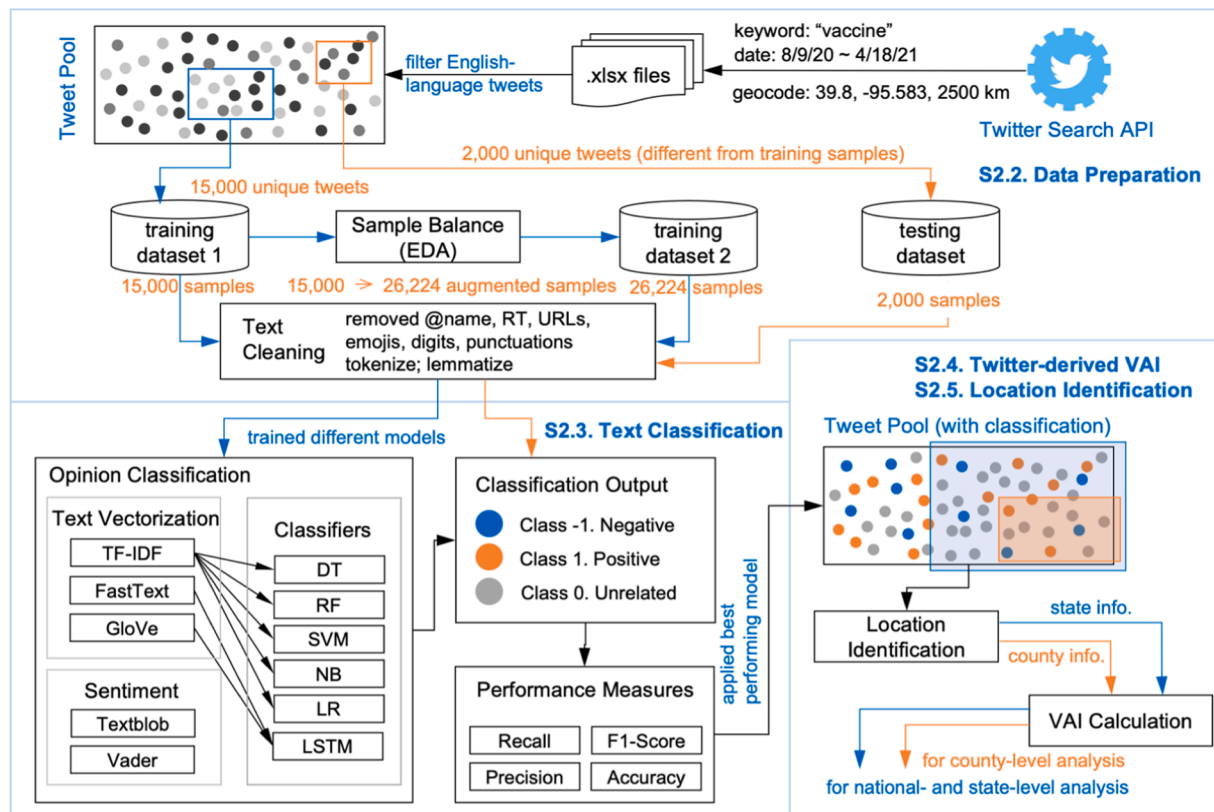


Fig. 1. The research framework for the implementation of the proposed model.

score, and Accuracy – defined in Section 2.3.4 – were used to measure the performance of each model on the testing set.

The best-performing pipeline was applied to the entire dataset to label all available tweets. Then, a vaccine acceptance index (VAI) was developed to measure the online level of vaccine acceptance in a geographic area during a particular period (Section 2.4). This VAI for a particular region requires information regarding the location of individual Twitter users. Section 2.5 explains the process of location identification.

2.2. Data preparation

2.2.1. Data collection

The Twitter Standard Search API [36] was used with the search term “vaccine” and geocode “39.8, -95.583, 2500 km” [37] to scrape vaccine-related tweets in the date range from August 9, 2020 to April 18, 2021 over the geographic area covered by the geocode. This geocode parameter value is specified by “latitude, longitude, radius,” and it returns tweets by users located within a radius of 2,500 km of the center 39.8, -95.583. The study selected this date range after the announcement of the first COVID-19 vaccine made by the Russian government on August 8, 2020, and before the U.S. mass vaccination started on April 19, 2021. The search word “vaccine” would return tweets that mentioned COVID-19 vaccine, but some fraction of the returned tweets was not COVID-19 related (e.g., HPV vaccine, flu vaccine). The latter were labeled as unrelated. The downloaded tweets included original tweets, mentions (i.e., a tweet that quotes another user’s username), replies (i.e., a comment to another user’s tweet), and retweets (i.e., a reposting of a tweet) that contained the keyword “vaccine.” The geographic search range was restricted to continental North America with the given geocode to investigate the U.S. public’s response to the COVID-19 vaccine. The resulting dataset contains 29,213,502 English language tweets.

2.2.2. Manual labeling

To classify users’ opinions towards the COVID-19 vaccine, the most frequently occurring 15,000 tweets were selected (i.e., the 15,000 unique tweets retweeted the most) and manually classified each tweet into a “positive (class 1),” “negative (class - 1),” or “unrelated (class 0)” about COVID-19 vaccine. Sentiment analysis was not used to classify the tweets because it was judged that sentiment score might not represent the user’s opinion in the research context. For example, the tweet “*Rich people did not experience the same pandemic as working-class people and now they get the vaccine first. It’s actually twisted*” is identified as negative using sentiment analysis, but it does not necessarily imply that this user was negative about the COVID-19 vaccine. A prior study also reveals the difference between sentiment and opinion mining in the text [38].

The criteria for manual labeling are based on the CDC strategy to reinforce confidence in COVID-19 vaccines [39]. Table 1 illustrates the specific labeling criteria used and examples. Tweets discussing vaccine passport (the second example in the class “unrelated” in Table 1) were labeled as unrelated (class 0) given that a Twitter user could possibly oppose to the vaccine passport while accepting the vaccine.

The manual classification process considered the 15,000 unique tweets that were retweeted the most, but subsequent analysis incorporated classifications for all retweets, given that online users who retweeted a tweet were presumed to share the same opinion as the original tweet. Each tweet in the training set was labeled by two annotators. When a tweet received the same label from both annotators, it was considered as the final label for this tweet. When there was a disagreement, the label was finalized through a short discussion by the annotators. As a result, there are 13,624 tweets (90.8% of training data) that received the same labels from two annotators, and the rate of variation between two annotators is 9.1% ((15000–13624)/15000). The manual labeling process resulted in a training set of 4,092 positive tweets (class 1), 1,783 negative tweets (class - 1), and 9,125 unrelated tweets (class 0).

Next, 2,000 unique tweets were randomly selected (different from

Table 1
Tweet classifications and examples.

Class	Label Criteria	Tweet Example
Positive (Class 1)	<ul style="list-style-type: none"> Refer to safety and effectiveness. Show positive emotion (e.g., willingness to take vaccine). Describe large distribution and administration, process and policies that lead to vaccine development, authorization, and recommendation. 	<p><i>I got the Pfizer vaccine and pso pfar pno pside effects [sic].</i></p> <p><i>Today was a good day. I have never been happier to wait in a line. If you're eligible, join me and sign up to get your vaccine. Come with me if you want to live!</i></p> <p><i>The U.S. recorded 3.1 million vaccine doses administered today, according to the CDC.</i></p> <p><i>Yesterday, in an effort to instill confidence in the vaccine, I had my shot administered in public. I have absolute confidence in the vaccine and look forward to the second shot.</i></p>
Negative (Class-1)	<ul style="list-style-type: none"> Show safety concern (e.g., illness, death, vaccine accidents, strong allergic reaction). Express negative emotion (e.g., not willing to take vaccine, protest against the vaccine). 	<p><i>The Oxford University coronavirus vaccine trial has been paused in what drug maker Astra Zeneca described as a "routine" action taken when there is a potentially unexplained illness among participants.</i></p> <p><i>23 die in Norway after receiving Pfizer COVID-19 vaccine: officials https://t.co/A56OILYmz0.</i></p> <p><i>America's Frontline Doctors are at the CDC in Atlanta to protest against forcing millions of Americans to take an experimental vaccine for Covid-19, a pathogen with a survival rate of 99.7%. We will fight against any experimental therapy being forced on anyone.</i></p>
Unrelated (Class 0)	<ul style="list-style-type: none"> Discuss topics not related to vaccine confidence (e.g., vaccine rollout, vaccine equity, vaccine passport). Describe topics not related to COVID-19 vaccine (e.g., flu vaccine, HPV vaccine). 	<p><i>Rich people did not experience the same pandemic as working class people and now they get the vaccine first. It's actually twisted.</i></p> <p><i>No vaccine passport. It doesn't get much more dystopian than being required to show your "health papers" wherever you go.</i></p> <p><i>Even with a vaccine, I'll still be wearing a mask and continuing cleaning and social distancing. I ain't had a cold or nothing since March.</i></p>

the training set), and the same label criteria was used to build the testing set. As a result, the rate of variation is 8.1% ((2000–1837)/2000), and the testing set contained 492 positive tweets (class 1), 224 negative tweets (class – 1), and 1,284 unrelated tweets (class 0). The training and testing sets were used for building and assessing the performance of a text classification model. In particular, the manually labeled samples in the testing set were considered as the “ground-truth” values and used to measure the classification performance of different candidate models.

2.2.3. Text cleaning

Before text classification, several steps were applied to clean the tweets. First, short URLs, user names (@username), the retweet prefix (RT @username), digits, emojis, and punctuations were removed. Next, each tweet was tokenized into a list of separate words and characters.

Then the tokenized words were converted to their stemming forms (lemmatization). The Natural Language Toolkit (NLTK) python package was used to complete the text cleaning [40].

2.2.4. Sample balance

Imbalance in the training set (9,125 class 0 tweets, 4,092 class 1 tweets, and 1,783 class – 1 tweets) might pose an issue for the text classification process in that the trained model could classify a random tweet as class 0 to obtain a high testing accuracy. As a result, the trained model might not produce an adequate performance for the minority class. There are two practical methods to tackle this issue, down-sampling and up-sampling. Down-sampling balances classes by training on a small subset of the majority class samples, while up-sampling balances classes by increasing the size of the minority class samples. Since down-sampling might lead to a loss of information for the model to capture “unrelated” tweet information, up-sampling was used to balance the training samples.

The easy data augmentation (EDA) technique was selected to perform up-sampling since it does not require a model pre-trained on an external dataset [41]. The EDA technique uses four operations to increase the text sample size: 1) synonym replacement (random words in a text are replaced by a synonym), 2) random insertion (random words are inserted into the text), 3) random swap (randomly swap two words in the text), and 4) random deletion (randomly delete words in the text) [41].

The coding recommendations proposed by the original article were followed and the augmentation for class 1 and class – 1 was set to one time and four times, respectively. As a result, the augmented training dataset contained 8,184 class 1 samples, 9,125 class 0 samples, and 8,915 class – 1 samples. To test the sensitivity of results to this experimental choice, both the original training set and the augmented (up-sampled) training set were used to train a classification model in the subsequent experiments.

2.3. Text classification

2.3.1. Sentiment classification

Sentiment classification, although it is different from opinion classification, was used to test the credibility of opinion classification. Sentiment classification tools use natural language processing (NLP) to analyze conversations and classify text into a positive, negative, or neutral emotional category. This study applied two popular Python-based sentiment tools, Textblob and Vader (Valence Aware Dictionary and sEntiment Reasoner). Textblob is a free Python library for processing textual data and provides a simple integrated and convenient API for diving into sentiment analysis [42]. Vader is an open-sourced and lexicon-based sentiment analysis tool under MIT license that is attuned to the sentiment expressed in social media [43].

2.3.2. Text vectorization

Text vectorization converts text into a vector or a matrix of vectors of numbers. Two text vectorization approaches were applied: Term Frequency-Inverse Document Frequency (TF-IDF) and Word Embedding methods. TF-IDF is a term weighting method used to perform text similarity, text classification, and information retrieval [44]. The key formula of TF-IDF is,

$$w(t, D) = tf(t, D) \cdot idf(t, N) = f_{t,D} \cdot \log\left(\frac{N}{n_t}\right) \tag{1}$$

in which $w(t, D)$ denotes the word t 's weight in tweet D ; $f_{t,D}$ represents the frequency of word t in tweet D ; N is the total number of tweets; and n_t is the number of tweets in which word t appears. Although TF-IDF does not capture word position or semantic similarity, it is an efficient algorithm for matching words in a query to documents [45]. Due to its simplicity and fast computation, TF-IDF is useful when dealing with a

large set of textual data.

Unlike TF-IDF, word embedding techniques can help capture semantic meanings [46], which may be used to compute text similarity and perform text classification. Word embedding methods convert each word into a vector of real numbers. Two popular pre-trained word embedding methods were applied in this study: FastText [47] and Global Vectors (GloVe) [48]. FastText applied in this research was pre-trained on Wikipedia based on 300 dimensions [47]. GloVe was trained based on Twitter content and each word was mapped using 300-dimensionals [53].

2.3.3. Multi-class classifiers

Multi-class classification involves more than two classes. In this study, the target domain contains three classes: class 1 (positive), class - 1 (negative), and class 0 (unrelated). After each word in a tweet was converted to a vector using the TF-IDF method or word embedding methods, different machine learning classifiers were applied to classify each tweet into one of the three classes.

Classifiers used in this study were Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), Logistic Regression (LR), and Long Short-Term Memory (LSTM). These are well machine learning classifiers and can be used in conjunction with the TF-IDF text vectorization method. Pipelines using these classifiers combined with the TF-IDF technique were constructed based on the Scikit-learn Python library [50]. The TF-IDF method converts a tweet into a vector of features; however, word embedding techniques convert each word in a tweet into an n-dimensional vector based on the pre-trained dataset. Thus, each tweet sample is represented as a matrix, with each column denoting the word vector. Given that the input to a classification model is a matrix, it usually requires deep learning techniques (e.g., neural networks) to deal with such classification tasks. In this study, the LSTM model was applied for text classification in conjunction with word embeddings. The implementation was constructed with the Keras Python library [51].

2.3.4. Performance measurement

Ultimately, two sentiment-based classifiers (Textblob and Vader), five vector-based machine-learning based classifiers (TF-IDF + DT, TF-IDF + RF, TF-IDF + NB, TF-IDF + SVM, and TF-IDF + LR) and two deep-learning classifiers (FastText + LSTM and GloVe + LSTM) were considered as candidate models. The nine candidate models were trained using the manual classified training data (with and without upsampling), and performance was assessed using the “unseen” 2,000 tweets contained in the testing set. Each of these machine learning classifiers contains hyperparameters that control the learning process. The grid search method was applied to tune the hyperparameters and to

Table 2
The grid search of hyperparameters for different machine learning classifiers.

Classifier	Grid search range
DT	Max depth: [10, 20, 40, 60, 80, 120, 200] Min samples leaf: [1,2,4]
RF	Max depth: [10, 20, 40, 60, 80, 120] Min samples leaf: [1,2,4] Num of estimators: [50, 100, 200, 300, 400, 500]
Multinomial NB	Alpha (smooth parameter): [0.1, 0.2, 0.5, 1, 1.5, 2]
Linear SVM	C (regularization parameter): [0.01, 0.05, 0.1, 0.5, 1, 2, 10, 100] Max iteration: [200, 500, 1000, 2000, 3000, 4000]
Multinomial LR	C (inverse of regularization strength): [0.01, 0.1, 0.5, 1, 2, 5, 10, 20, 100] Multiclass solver: ['sag', 'saga', 'lbfgs'] Max iteration: [500, 1000, 2000, 4000, 8000, 12000]
LSTM	Batch size: [16, 32, 64, 128, 256] Num of epochs: [10, 20, 30, 50, 80, 100, 150] Learning rate: [0.001, 0.01, 0.05, 0.1, 0.2] Drop rate: [0.2, 0.3, 0.4, 0.5]

search for the ideal model architecture. The parameter range of grid search for different classifiers is presented in Table 2.

Precision, Recall, and F1-score were used to assess the classification performance on the testing data. Precision measures the fraction of true positive cases over the retrieved cases a model predicts. Recall is the fraction of true positive cases over all the relevant cases that are actually identified. The F1-score is a rating of test accuracy, combining Recall and Precision [52]. They are defined as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

where TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative.

Using grid search, different hyperparameters listed in Table 2 were tuned to achieve an idea architecture for each machine learning classifier. The best hyperparameter combination for each classifier was selected mainly based on F1-scores and testing accuracy. Performance on the testing dataset of these classification models with the best tuned hyperparameters is shown in Table 3 for both the original and upsampled training sets.

The TF-IDF + RF trained on an augmented training set offered the best performance overall and therefore was applied to the whole dataset in subsequent steps. The performance of some models improved but some reduced when dealing with the augmented training set, according to the F1-scores. One possible explanation is that the text augmentation added more samples to balance the classes but also introduced more information to the training process (e.g., more words with synonym change, and more training samples with the same classification). However, with sample balance, most models showed an improvement on class 1 and class - 1. In particular, the performance of some models (built with NB and RF) for class - 1 greatly improved with sample balance, as demonstrated by a higher F1-score.

Second, pipelines using word embedding techniques were not consistently superior to those built on TF-IDF even those word-embedding methods are often considered more advanced. In applications such as those in this study, it is crucial to discriminate between words with similar meanings. However, the word embedding techniques may contain too much hidden information that is difficult to understand by the machine.

Third, the performance of sentiment classification tools was barely satisfactory, as demonstrated by lower F1-score and testing accuracy. As discussed, in applications of mining opinions from social media data, sentiment scores may not sufficiently reflect users' opinions, especially when involving complex context (e.g., irony, metaphor).

2.4. Vaccine acceptance index (VAI)

The selected classification model TF-IDF + DT trained on the upsampled training dataset was applied to 29,213,502 tweets contained in the full dataset. A classification was assigned to each tweet and used to compute a user-based vaccine acceptance measure. The following index was proposed to measure vaccine acceptance in geographic area during a particular period. Vaccine acceptance was defined for a specific Twitter user j during time-period k ($VA_{j,k}$) as:

$$VA_{j,k} = \frac{P_{jk} - N_{jk}}{P_{jk} + N_{jk}} \quad (6)$$

Table 3
Classification performance.

	Without sample balance			With sample balance		
	Class 1	Class - 1	Class 0	Class 1	Class - 1	Class 0
Precision						
Textblob	0.33	0.15	0.60			
Vader	0.33	0.19	0.58			
TF-IDF + DT	0.46	0.29	0.68	0.41	0.26	0.67
TF-IDF + RF	0.91	0.91	0.66	0.70	0.57	0.79
TF-IDF + NB	0.69	0.58	0.70	0.58	0.35	0.84
TF-IDF + SVM	0.61	0.49	0.76	0.56	0.40	0.81
TF-IDF + LR	0.68	0.58	0.75	0.58	0.41	0.81
FastText + LSTM	0.65	0.43	0.68	0.44	0.26	0.78
GloVe + LSTM	0.56	0.40	0.72	0.40	0.32	0.75
Recall						
Textblob	0.57	0.26	0.31			
Vader	0.52	0.52	0.24			
TF-IDF + DT	0.36	0.21	0.78	0.39	0.31	0.66
TF-IDF + RF	0.24	0.04	0.99	0.61	0.49	0.85
TF-IDF + NB	0.43	0.11	0.92	0.63	0.73	0.63
TF-IDF + SVM	0.56	0.38	0.82	0.62	0.66	0.67
TF-IDF + LR	0.54	0.34	0.88	0.61	0.64	0.71
FastText + LSTM	0.33	0.21	0.90	0.55	0.40	0.61
GloVe + LSTM	0.43	0.35	0.82	0.58	0.34	0.58
F1-score						
Textblob	0.41	0.19	0.41			
Vader	0.40	0.28	0.34			
TF-IDF + DT	0.40	0.25	0.72	0.40	0.28	0.66
TF-IDF + RF	0.38	0.08	0.79	0.65	0.52	0.82
TF-IDF + NB	0.53	0.19	0.79	0.61	0.47	0.72
TF-IDF + SVM	0.59	0.43	0.79	0.59	0.50	0.73
TF-IDF + LR	0.60	0.42	0.81	0.60	0.50	0.75
FastText + LSTM	0.44	0.28	0.78	0.49	0.31	0.68
GloVe + LSTM	0.50	0.37	0.76	0.47	0.33	0.65
Accuracy						
		Training	Testing		Training	Testing
Textblob			37.4%			
Vader			34.9%			
TF-IDF + DT		94.6%	59.6%	94.5%	54.3%	
TF-IDF + RF		92.8%	67.6%	99.2%	74.4%	
TF-IDF + NB		83.9%	69.1%	85.6%	64.4%	
TF-IDF + SVM		89.0%	69.7%	92.9%	65.5%	
TF-IDF + LR		85.6%	72.1%	94.6%	67.4%	
FastText + LSTM		94.2%	66.3%	86.9%	56.8%	
GloVe + LSTM		89.4%	65.9%	92.2%	55.3%	

where P_{jk} and N_{jk} denote the number of positive tweets and negative tweets, respectively, that this user posted in the given period. This Twitter-derived vaccine acceptance index reflects public opinions over COVID-19 vaccine acceptance that emerged on Twitter. This index was computed based on the unit of users rather than tweets.

The regional vaccine acceptance index for geographical region i during time-period k ($VAI_{i,k}$) was computed as the average of the indices computed for all $R_{i,k}$ users located in the region:

$$VAI_{i,k} = \frac{1}{R_{i,k}} \sum_{j=1}^{R_{i,k}} VA_{j,k} \tag{7}$$

For example, if a user j (located in region i) posted five tweets during time period k identified as $\{-1, 0, 1, -1, -1\}$, then the acceptance value for that user during the time period ($VA_{j,k}$) is $(1-3)/(1+3) = -0.5$. If $R_i = 6$ users from region i posted related messages during time period k , and their acceptance values are computed as $\{-0.5, 0.5, -0.5, 0.2, 0.8, 0.7\}$, then the VAI for region i and time period k ($VAI_{i,k}$) is computed as

$(-0.5 + 0.5 - 0.5 + 0.2 + 0.8 + 0.7)/6 = 0.2$. For the following analysis, this formula was used to compute VAI at the national, state, and county levels in each time period.

2.5. Location identification

The regional vaccine acceptance index requires information regarding the location of individual Twitter users. There are two types of information that can be leveraged to identify a user's location: (1) a user posting geolocation information (latitude/longitude) as part of a public tweet, and (2) a user providing registration location (i.e., home location) as part of their public profile. While geolocation tied to a tweet provides the most robust location information, only a small proportion of tweets in the dataset (18,806 out of 29,213,502 or 0.06% of tweets) were associated with latitude and longitude. This does not provide sufficient data to reveal the temporal and spatial patterns of interest. However, Twitter allows users to register home locations in their user profile. More than half (16,334,206 out of 29,213,502 or 55.9%) of tweets in the dataset had registration locations indicating either U.S. state or city. Therefore, registration location was used to ensure sufficient data for temporal and spatial analysis.

To identify state and county information from registration profiles, a simple word mapping was applied based on a list of U.S. locations (cities, counties, state names, and state codes) obtained from SimpleMaps.com [53]. For example, the following registration locations, "California, USA," "CA," and "California," indicate the State of California. Other registration locations may contain more specific location information, such as city or county. For example, the following locations, "Los Angeles, CA" "New York City," "Orlando," and "Houston, TX," contain city information. Location information was used to identify the user's profile state and, if possible, the user's profile county. If a registration location contained only city information, it was referred to its corresponding county and state. For example, if a user's registration location is "Seattle," then the county information for this location was identified as "King County" and the state is identified as "Washington." Using this approach, a Python code was developed by the authors, and 16,334,206 records were extracted containing state information for nation-level and state-level analysis and 8,885,549 records containing the county information for county-level analysis.

Some limitations of this application are noted. First, this approach cannot identify locations that were not presented in the list of U.S. locations obtained from SimpleMaps.com, such as "Times Square" or "Long Island." Second, this approach only returns the last location information when a user's profile contains multiple locations. For example, if the registration location includes two locations such as "LA and NYC," the method will only identify the last location "NYC" and refers it to the county and state information "New York, NY." Third, it cannot capture nouns of locality in the users' registration locations, such as "near," "between," "north to," and "close to." Last, this approach returns the county and state information with the most populated area when the user's input only contains the city name that may refer to multiple locations in the U.S. For example, a user's input of "Portland" can refer to a city in Oregon, Maine, Colorado, Texas, Tennessee, or Indiana. This approach generates the county and state information as "Multnomah, Oregon" given that the Portland city in Oregon has the largest population among these cities.

3. Results

3.1. National-level analysis

The volume of positive- and negative-classified tweets over time is shown in Fig. 2a. Fig. 2b shows the national VAI computed each day and on a 7-day rolling average. For the daily values (blue line in Fig. 2b), a clear rising trend in observed during the study period. This characterization appears consistent with survey results on 7,420 American adults

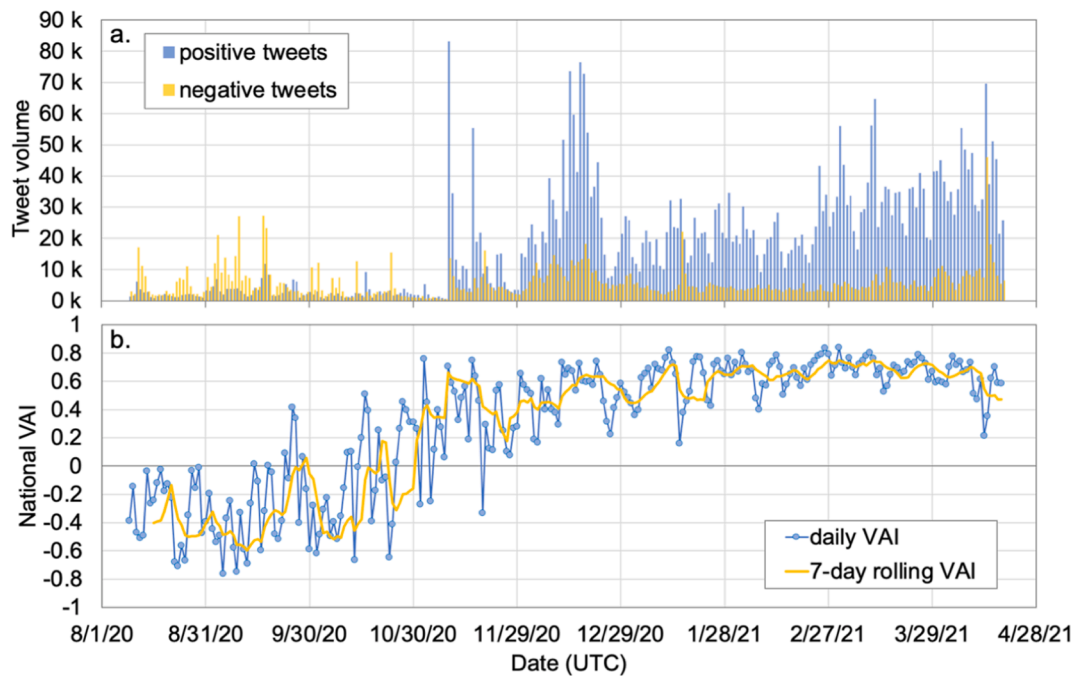


Fig. 2. Temporal results of (a) daily tweet volume of all positive and negative tweets, and (b) national VAI computed on a daily and weekly basis.

indicating that vaccine hesitancy declined by 10.8% from October 2020 to March 2021 [15]. It is also consistent with national poll results reported by Nguyen et al. [16] which showed that the intent to receive COVID-19 vaccination increased from 39.4% to 49.1% among adults and across all priority groups from September to December 2020.

In addition to providing insights on overall trends, VAI provides insights regarding public reactions to vaccine-related events. Given that Twitter users responded to news events by posting personal opinions and event observations, the index is likely to reflect events and the availability of vaccine information (e.g., clinical trial results regarding vaccine safety and efficacy). Fig. 3 shows the same daily time-series with notable events annotated.

For example, the national VAI moved from negative acceptance to the positive range, and the volume of positive tweets increased in early

November. Pfizer published its trial efficacy result on November 9, 2020. Before this date, the overall opinions on vaccine acceptance fluctuated between negative and neutral. Other notable events or opinions potentially contributing to changes in vaccine acceptance might have included: (1) scientists worried that the government put political pressure on the FDA to approve the vaccine without referencing the clinical data, and (2) the U.S. rejected joining the WHO’s COVID-19 vaccine initiative [54]. After the trial data became available from Pfizer, the VAI increased to 0.8 and remained consistently high throughout the study period (April 19, 2021). Near the end of the study period, a significant negative spike was observed in Fig. 3, which coincides temporally and thus may be associated with the pause of the Johnson & Johnson vaccine due to the report of blood clots on April 13, 2021 [55].

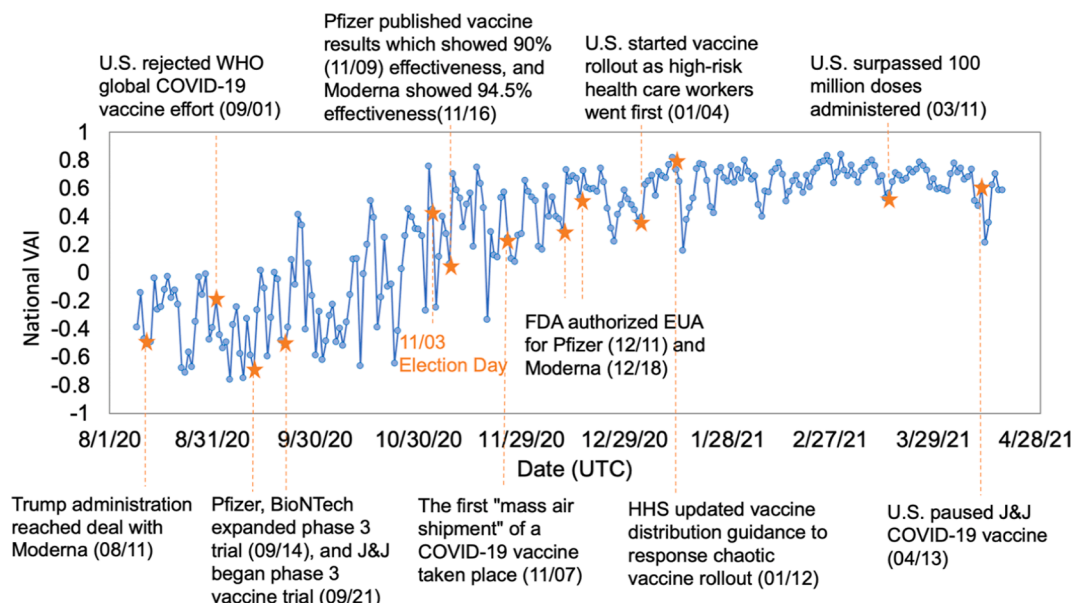


Fig. 3. Timeline of notable events relative to calculated national-level VAI.

3.2. State-level analysis

The relationship between the VAI and later vaccination rates was explored at the state level. Tweets were extracted for state-level locations at which information could be obtained. Tweets were grouped based on extracted locations to calculate the state-level index using Eq. (7). The 7-day VAI in each state was compared with the national-level trend (Fig. 4a). In each subplot, the grey dotted line is the rolling weekly national trend, and the blue line and yellow lines show state-level trend with the color denoting whether the state-level VAI is above (blue) or below (yellow) the national VAI. Detrended state-level

7-day VAI were obtained by subtracting the national VAI. This provides a clearer view of the changes in each state. In both Fig. 4a and Fig. 4b, the x-axis represents the study period, and the y-axis represents the VAI (Fig. 4a) or difference in VAI (state VAI minus national VAI) (Fig. 4b).

Other events potentially contributing to the evolution of vaccine acceptance include positive results from vaccine clinical trials, FDA's approval of the emergency use authorization [56], vaccinations of front-line healthcare workers [57], and quick vaccine distribution across the U.S. [58]. These interactions between the Twitter-derived national VAI and the events suggest that social media could serve as an effective tool

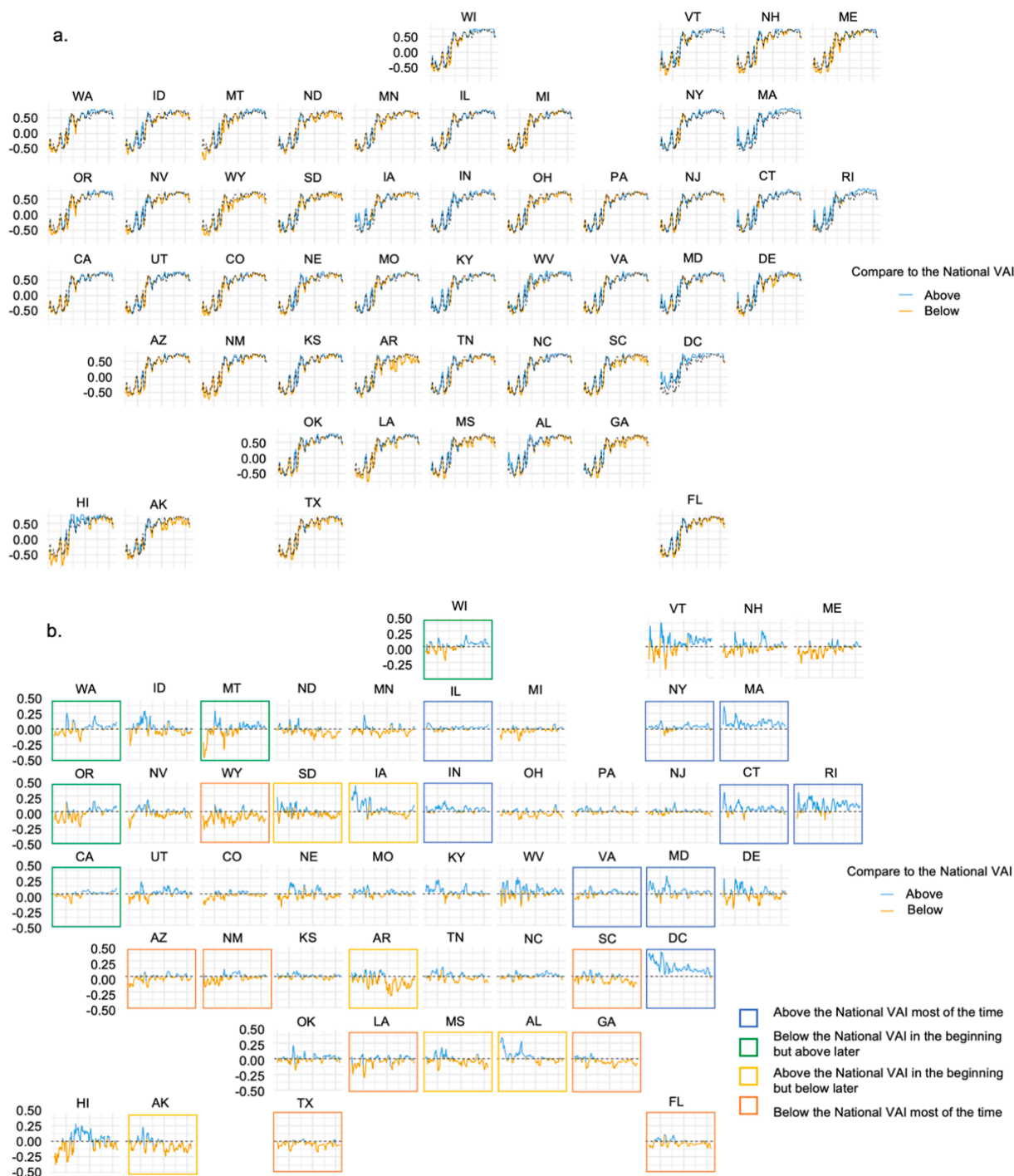


Fig. 4. 7-day rolling average of the state-level VAI (a) compared with the national VAI (y-axis is the VAI value and x-axis is the date across the full study period), and (b) with detrended from the national VAI (y-axis is the difference between state-level VAI and national VAI, and x-axis is the date across the full study period).

to reflect online opinions and support near real-time tracking of vaccine acceptance.

Fig. 4a shows that the general trend of the index increased in all states. However, several patterns were observed in different states. These patterns were summarized based on manual classification and were only used as illustrative to show changes of the trends of state-level VAI. Overall, the East Coast of the U.S. (e.g., NY, MA, CT, RI, MD, VA, and DC) and selected Midwest states (i.e., IL and IN) tended to exceed the national index most of the time, while states in the south (e.g., AZ, NM, TX, LA, SC, and FL) were lower than the national level. Some states (e.g., SD, IA, AR, TN, MS, AL, and GA) had a higher acceptance index at the beginning but dropped afterward. Other states (e.g., WA, OR, CA, WI, and MT) were more negative in the first few months but later turned positive.

Correlations were calculated between the Twitter-derived VAI and the cumulative vaccination rate [1] at the state level. This aimed to: (1) measure the potential value of using the social media-derived index to understand future vaccination rates at the state level, and (2) determine conditions when the Twitter-derived VAI shows a strong correlation to vaccination rate.

Fig. 5a-c shows the correlation coefficient calculated between the daily (Fig. 5a), 7-day rolling average (Fig. 5b), and 30-day rolling average (Fig. 5c) state-level VAI and the corresponding state's vaccination rates on June 22, 2021. The data for vaccination rates were obtained from CDC COVID-19 Data Tracker [1]. The regression line between time and the correlation coefficient is shown by the dotted trend line (with associated R^2 values applying to the whole regression line annotated on the graphs).

As expected, the VAI calculated using more recent data shows a stronger positive correlation with the June vaccination rates. It was further observed that calculating the VAI over a more extended time (monthly rolling average) generates a more stable relationship. The correlation is higher after the first vaccine distribution, which occurred on December 14, 2020, as indicated by the yellow boxes on Fig. 5c. This suggests that, at the state-level, the online opinions on vaccine acceptance expressed after the initial vaccine rollout are more strongly related

to later vaccination rates.

To further explore the temporal change in correlation between vaccination rates and the 30-day VAI, the correlation analysis was repeated using vaccination rates on three dates: April 22, May 22, and June 22, 2021. Fig. 5d illustrates the correlation results. A similar level of moderate-to-strong correlation (between 0.3 and 0.7) was observed between the social-media-derived VAI and vaccination rates on May 22 and June 22, 2021. A lower correlation was observed with vaccination rates on April 22. The difference in strength of correlation could reflect that earlier vaccination rates were affected by the distribution plan rather than people's willingness to receive the vaccine whereas later vaccination rates began to reflect preferences and opinions. This may also reflect the evolution from a time when vaccination rates reflect vaccination availability to a time when vaccination rates are driven by preferences.

3.3. County-level analysis

County-level VAI was computed as the average of the acceptance levels of all users from that county in a given period, where the county assigned to a particular user was determined from profile information. However, increased spatial discretization means there were fewer data available for each spatial unit (county). Fig. 6a presents the distribution of Twitter users by county. Fig. 6b shows county-level VAI computed across the period after the first vaccine distribution (December 14, 2020, to April 18, 2021) for counties with more than 30 users.

Fig. 6a illustrates that most Twitter users in the dataset considered were from urbanized areas, such as New York, Los Angeles, Washington DC, and San Francisco. This observation is consistent with prior studies that social media data are likely to be biased, especially when considering urban and rural distributions [59,60]. It also reflects on a key question: in what circumstances can social media data generate a reliable estimate for county-level vaccine acceptance?

To address this, the correlation between the Twitter-derived county-level VAI and county-level vaccination rate changes was calculated as a function of the number of users in that county. The analysis considered

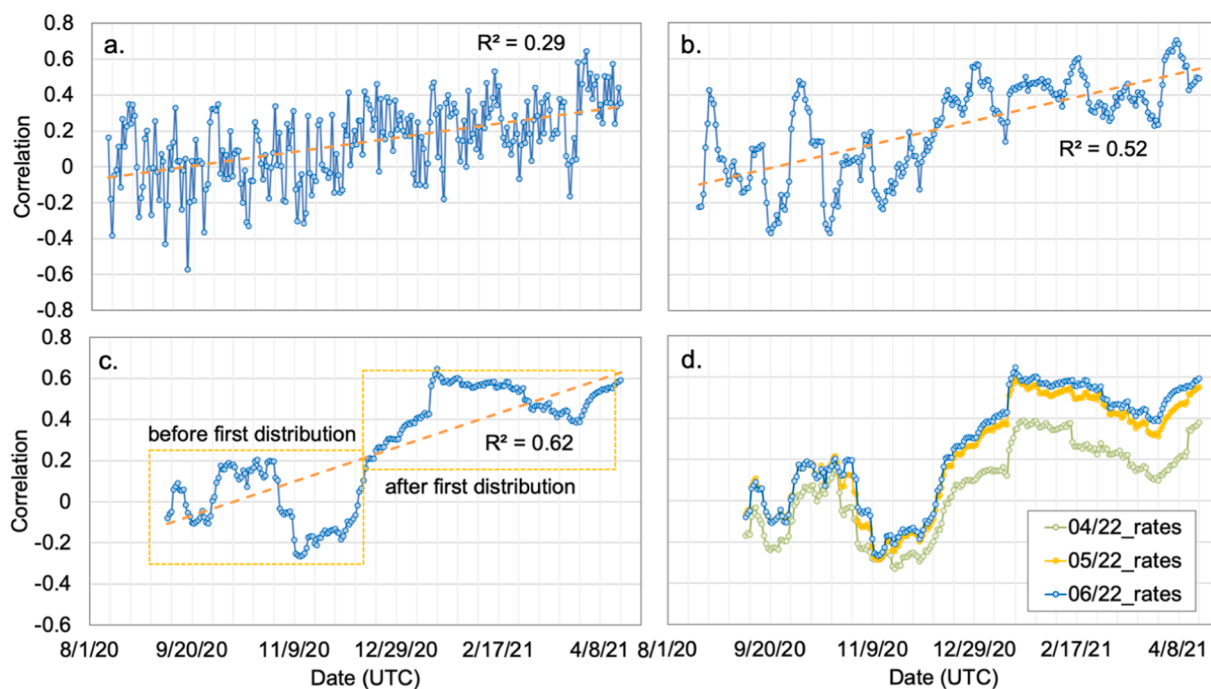


Fig. 5. The correlations of the state-level VAI with accumulative state-level vaccination rates reported on June 22, 2021, on a (a) daily, (b) rolling weekly, (c) rolling 30-day basis, and against accumulative state-level vaccination rates (d) on April 22, May 22, and June 22, 2021, (the legend “04/22_rates” represents the correlation of state-level VAI with vaccination rates reported on April 22).

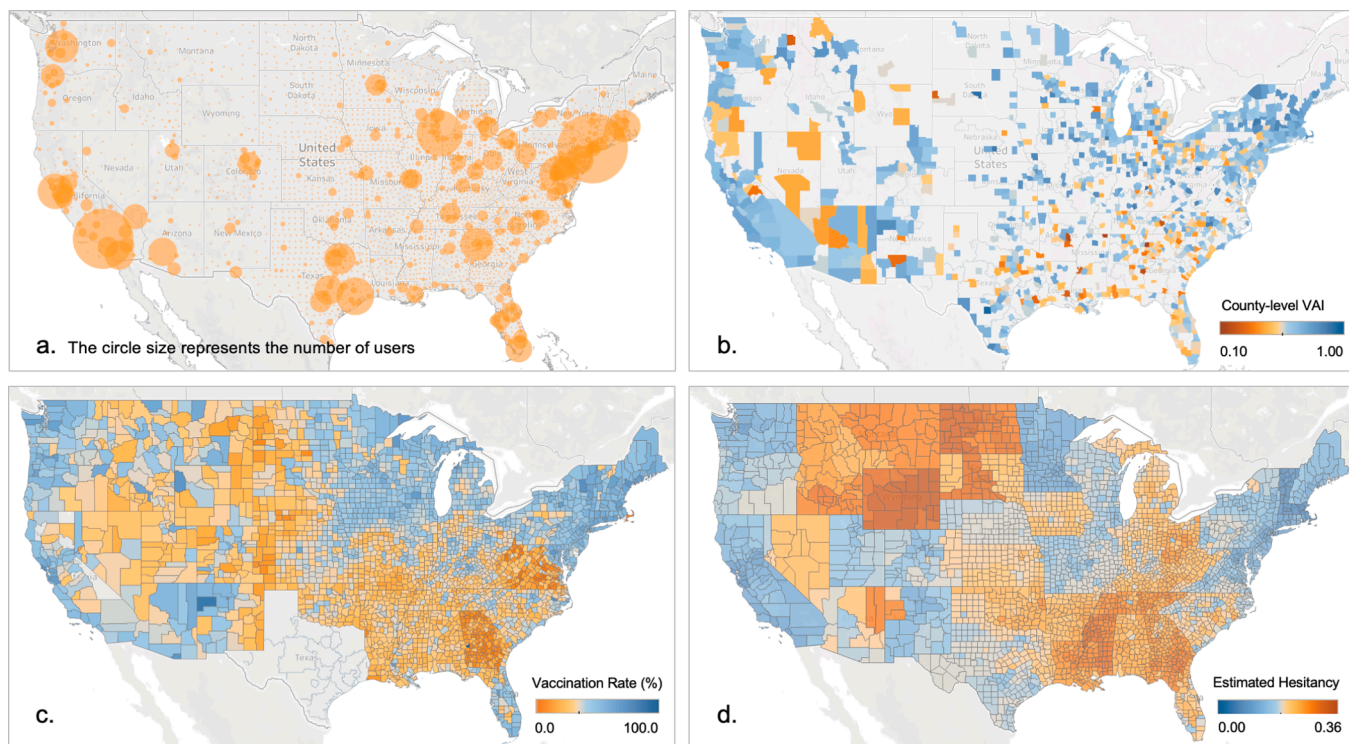


Fig. 6. (a) The number of Twitter users in each county during full study period, (b) County-level VAI computed for the period (December 14, 2020 to April 18, 2021) for counties with the number of users larger than 30, (c) County-level vaccination rate based on the data published on June 22, 2021, and (d) CDC estimated vaccine hesitancy based on the HPS from March 3 to 15, 2021.

vaccinations published by CDC on April 22, May 22, and June 22, 2021 [1]. County-level vaccination rates on June 22 are presented in Fig. 6c. County-level vaccination rates for Texas and selected counties in California were not available. The change in correlation with number of users in a county was based on two scenarios: (1) using all the data in the study period, and (2) using data after the first vaccine distribution. The results are presented in Fig. 7a and Fig. 7b for scenarios 1 and 2, respectively.

The correlation sharply increases when the number of users increases from 0 to 10. The correlation continues to rise, but at a slower rate, until leveling off between 20 and 30. Similar to the state-level analysis, data after the first vaccine distribution (scenario 2) generated a higher correlation than all the data did (scenario 1). This observation again demonstrates that the pattern of online messages after the vaccine distribution had a positive correlation with people’s inclination to receive the vaccination.

The low correlation observed when data are only available from a small number of users demonstrates that those data are likely not

representative of the actual county-level vaccine acceptance. For example, when a county only has one user posting related opinions, the county-level estimation is entirely up to this user’s attitude, which could bias the analysis. However, the correlation shows general stability when counties are included in the correlation analysis only when the number of users is larger than about 30. It is further noted that the increase in correlation with the number of users may also pick up on the effects of unequal geographic sampling. For example, the correlation between vaccine acceptance index and vaccination rates may be higher in metropolitan areas than other regions.

Finally, county-level VAI was compared with the CDC estimated vaccine hesitancy. The resulting percentages are presented in Fig. 6d (based on the data collected by the CDC from March 3 to 15, 2021) [17]. The result of the state-to-county downscaling approach is that it shows clear boundaries between states even when two counties are close to each other (e.g., note the border between Idaho and Washington). This shows negative correlation between county-level vaccine hesitancy (as estimated by the down-scaled survey) and vaccination rates (more

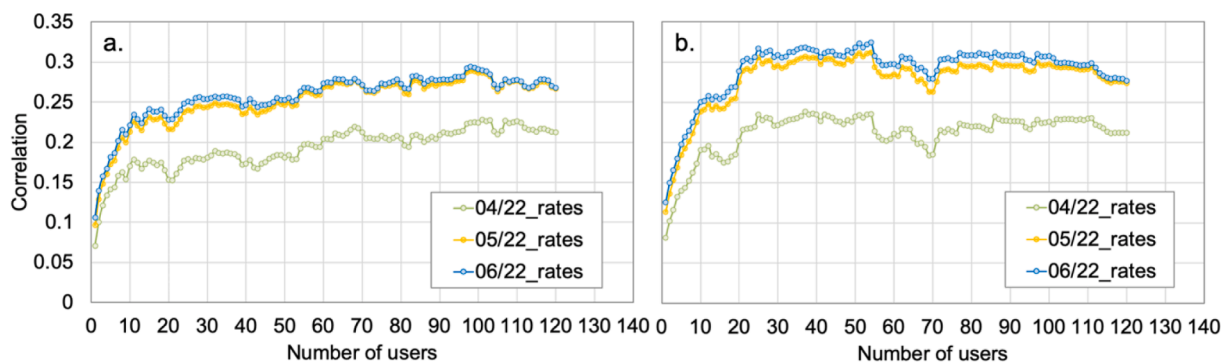


Fig. 7. The correlations of the county-level VAI (based on different numbers of users) with vaccination rates using tweets obtained (a) from August 9, 2020, to April 18, 2021, and (b) from December 14, 2020, to April 18, 2021 (after the first distribution).

hesitancy is associated with lower vaccination rate) on April 22 of -0.23 (p-value < 0.001), May 22 of -0.37 (p-value < 0.001), and June 22 of -0.41 (p-value < 0.001). Compared to VAI, the survey-based down-sampling method shows higher correlation. Still, the survey-based approach requires a larger commitment of time and resources and does not offer the same ability to track vaccine hesitancy over a long period or to see effects such as the impact of newsworthy events on opinions.

4. Discussion

While vaccination is the most successful public health intervention to contain infectious diseases, strategies are needed to rapidly assess vaccine acceptance. Survey-based methods have been applied to estimate vaccine hesitancy. In the COVID-19 context, studies have demonstrated the usefulness of survey tools to measure the intent for COVID-19 vaccination and understand the driving factors [18–20]. However, two challenges remain with survey-based methods. First, collecting survey responses requires time and effort. Second, since information about demographic factors is difficult to collect over large areas, most surveys target one place or several places that may provide limited insights on assessing vaccine acceptance at a variety of geographical scales.

Building on the existing body of knowledge related to the utility of social media in vaccine and other health studies, this study demonstrates that information can be drawn from Twitter users' posts about attitudes toward vaccination. These posted messages combined with textual analysis can be leveraged to provide a fast and cost-effective assessment. Existing studies have utilized social media data with semantic network or topic modeling to analyze pro- or anti-vaccine opinions. However, there has been less focus on quantitative assessment of vaccine acceptance using social media data. While a few studies have leveraged sentiment analysis to estimate vaccine hesitancy, sentiment scores may not show appropriate representation of opinions. The present study fills this gap and extends previous studies by introducing a Twitter-derived vaccine acceptance index (VAI) derived using text classification. The result reveals a rising trend of this index, which appears consistent with survey studies that indicate vaccine hesitancy in the U.S. declined from October 2020 to March 2021 [15]. The study further explores the correlation of VAI with vaccination rates at the state, and county levels.

4.1. Theoretical and practical contribution

From a theoretical perspective, the present study contributes to research focusing on information extraction from social media data. First, it provides insights into natural language processing tools to perform multi-class classification tasks on social media data. In this context, the TF-IDF method may be as effective as more advanced word embedding methods, possibly because word embedding methods may not discriminate minor distinctions among tweet data or bring too much hidden information to the training process (e.g., because each word is converted into a multi-dimensional vector).

Second, this study indicates differences between sentiment classification and opinion mining. A few studies have used sentiment analysis to assess vaccine hesitancy because sentiment could be a good indicator to reveal the overall emotions about vaccine confidence [32]. However, sentiment classification may result in a bias in opinion extraction in applications such as this study. This statement is supported by the comparatively low testing accuracy of sentiment classification compared to text classification models.

Third, the research proposes a vaccine acceptance index based on tweeters' opinions, which is a quantitative approach to measure online opinions. While this study is presented as a COVID-19 vaccine case study, the approach is applicable to similar issues and can be generalized as a framework with medicine and health relevance. For example, during the COVID-19 pandemic, online opinions on wearing masks and implementation of lockdown can be investigated using a similar

research framework. It can also be generalized to study the vaccine acceptance of the flu vaccine and HPV vaccine or opinions on other important healthcare events or policies.

From a practical perspective, this study provides an approach to the rapid assessment of vaccine acceptance. At the national level, this Twitter-based index provides an instrument to track online opinions on vaccine acceptance over time. The national-level analysis also reflects that the index is subject to the dynamics of notable events and the availability of vaccine information (e.g., clinical trial results regarding vaccine safety and efficacy). At the state level, through the exploration of correlation with later vaccination rates, the study has shown the potential of social media to assess vaccine acceptance. Observations are made regarding conditions under which correlations are high. For example, data obtained after the first vaccination rollout yields higher positive correlations, and the use of a longer period of data collection (e.g., one month) for computing the VAI shows a higher correlation with later state-level vaccination rates. At the county level, since social media data may be biased towards higher population (urbanized) areas, the study explores the effect of sample size on the computed correlation between the index and county-level vaccination rates. The correlation analysis suggests that correlations stabilized when only considering counties for which data was available from at least 30 users.

The social media approach can benefit government agencies and health officials in the following ways. First, the study reveals that on the use of social media data can contribute to additional insights on public reactions to governments' decisions on the vaccine. Such insights can usually be obtained via social media channels. Government agencies can also refer to social media analysis to assist decision-making related to the mechanisms for releasing information. Second, the study indicates conditions for using social media data for a granular assessment of vaccine acceptance. For example, calculating the state-level and county-level index with data after the first vaccine distribution can generate a more reliable analysis. The county-level assessment stabilizes when considering counties with more than 30 users. Third, this approach can help assess vaccine acceptance and help decision-makers to provide vaccine education for places where people feel hesitant to receive the vaccine.

4.2. Limitations and future work

Biases associated with this social media approach are worthy of note. First, relying on social media data can help eliminate the bias of those choosing to participate in a survey study, but it still has the bias of those choosing to leave a public comment about the vaccine. As described in the county-level analysis, previous studies have demonstrated that social media data are likely to be biased in a more granular level estimation (e.g., social media users are younger and from more urbanized areas). These intrinsic characteristics of social media users might affect the estimates, and this limitation can be difficult to overcome in gathering social media data [59,60].

Two processes during the model development could also introduce model bias. First, this study selected the keyword "vaccine" to scrape tweets, mainly to ensure consistency throughout the study period. However, other keywords, such as "vaccination" or common vaccine-related misspellings and slang (e.g., "vax"), may also contain helpful information about vaccine acceptance. Second, manual classifications of tweets (including the "ground-truth" samples in the testing set) were based on the authors' interpretations of the COVID-19 vaccine acceptance expressed in the tweet. However, the authors might incorrectly identify a user's opinion described by a tweet.

When this social media-based index is applied to investigate vaccine acceptance, several limitations need to be addressed. First, this study used a general list of U.S. states and cities to match users' registration locations. Given that a Twitter user's input for the registration location could differ from the tweet location, there is no guarantee that the extracted information reflects a user's actual location. Some location

descriptors were not considered in our analysis. For example, a user may input the location “Times Square,” which suggests a place in New York City, but this situation is not considered in this study. Future work will consider better location descriptors and integrate them with the current pipeline to identify location (e.g., city and county).

Although the “simpler” algorithm TF-IDF technique obtained reasonable accuracy in this study, it may not be suitable for situations where tweet data differ from the training tweet data, especially when the size of the training set is small. TF-IDF is a bag-of-words method that converts each word into a number (a word count with the IDF score). If the training set does not include many words that appear in the whole dataset, the TF-IDF technique may not generate a satisfactory performance. However, word embedding may also fail to generate competent performance given that it is not pre-trained based on health-related databases or cannot capture minor discriminations of words between tweets. Future work will focus on the application of neural network architectures with text vectorization techniques to improve performance, such as Bidirectional Encoder Representations from Transformers (BERT).

The analysis completed in this study uses English tweets and English-based NLP techniques. When the VAI is applied to estimate vaccine acceptance in non-English speaking countries, tweets written in other languages need to be collected to train the model, and correspondingly, text classification models need to be developed using NLP techniques developed in other languages. Moreover, as previously mentioned, social media data may overrepresent the young, educated, and urbanized population, so the VAI may not be suitable for applications in those areas with few people using social media to communicate. Therefore, future work will also consider using NLP and machine learning tools trained in other languages and integrating data from surveys or models (e.g., CDC HPS model [17]) to complement the social media estimation.

Ongoing and future work could build off the insights and approaches described in this study. For example, an extension of this study could be used to evaluate the impacts of campaigns or interventions by state or federal public and private sectors by tracking the dissemination patterns (e.g., retweets, mentions) of tweets from state-, federal-, and certain private-affiliated accounts. This can help to inform how those vaccine campaigns affect the online community. As another example, the proposed VAI can be investigated hours to days after the interventions applied to assess their real-time impacts on vaccine acceptance. Another extension of this study could consider integrating data from other social media platforms and comparing the results of VAI computed based on different social media data. This can help reveal the characteristics of different social media users when focusing upon policies or events with medicine and health relevance.

Other future work could explore the relations of the online vaccine acceptance with socioeconomic factors (e.g., education level, gender, age, employment rate) and other COVID-19 factors (e.g., death rate, hospitalization rate). This social media approach may provide valuable insights into the geographical dynamics of vaccine acceptance and aid health professionals in conducting vaccine interventions, such as vaccine education. Last, future work could consider the generalizability of this research framework by applying it to investigate the vaccine acceptance on other types of vaccine and in a broader geographical scale, such as the vaccine acceptance over HPV vaccine and flu vaccine in different countries.

5. Conclusions

The study investigated 29 million vaccine-related tweets from August 8, 2020 to April 19, 2021, and proposes a vaccine acceptance index (VAI) to assess COVID-19 vaccine acceptance. To classify the opinions regarding vaccine acceptance, 17,000 sampled tweets were manually labeled, and different pipelines were developed with the aid of NLP techniques. Significant findings include: (1) the national VAI moved from negative to positive in 2020 and maintained constant after January

2021, consistent with survey results, (2) the East Coast of the U.S. and selected Midwest states exceeded the national index most of the time, while states in the south were lower than the national level, (3) calculating VAI over a more extended time and based on data after the first vaccine distribution generated a more reliable analysis, (4) calculating the county-level VAI when only considering counties for which data were available from at least 30 users ensured a more consistent analysis.

These findings have demonstrated the usefulness of social media data for the dynamic assessment of vaccine acceptance and provided insights into the reliability of using social media data. This proposed social media approach provides potentials for fast and cost-effective assessment of vaccine acceptance across large geographical scales.

CRedit authorship contribution statement

Lingyao Li: Conceptualization, Data curation, Methodology, Formal analysis, Visualization, Writing – review & editing. **Jiayan Zhou:** Conceptualization, Data curation, Formal analysis, Visualization, Writing – review & editing. **Zihui Ma:** Data curation, Visualization, Writing – review & editing. **Michelle T. Bensi:** Conceptualization, Project administration, Supervision, Writing – review & editing. **Molly A. Hall:** Conceptualization, Supervision, Writing – review & editing. **Gregory B. Baecher:** Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank Sanggyu Lee from the Department of Civil and Environmental Engineering at the University of Maryland for his assistance in tweet labeling. The authors also thank two anonymous reviewers whose suggestions helped improve and clarify this paper.

This work was additionally supported by the USDA National Institute of Food and Agriculture and Hatch Appropriations under Project #PEN04275 and Accession #1018544, startup funds from the College of Agricultural Sciences, Pennsylvania State University (<https://agsci.psu.edu/>), and the Dr. Frances Keesler Graham Early Career Professorship from the Social Science Research Institute, Pennsylvania State University (<https://ssri.psu.edu/>) to MAH.

References

- [1] CDC, COVID Data Tracker, Centers for Disease Control and Prevention, 2021. <https://covid.cdc.gov/covid-data-tracker> (accessed Jul. 19, 2021).
- [2] S.G. Sheryl and Karni, Nation Faces ‘Hand-to-Hand Combat’ to Get Reluctant Americans Vaccinated – The New York Times, Apr. 21, 2021. <https://www.nytimes.com/2021/04/21/us/politics/coronavirus-vaccine-rates.html> (accessed Jul. 19, 2021).
- [3] K.A. Fisher, S.J. Bloomstone, J. Walder, S. Crawford, H. Fouayzi, K.M. Mazor, Attitudes Toward a Potential SARS-CoV-2 Vaccine, *Ann. Intern. Med.* 173 (12) (2020) 964–973, <https://doi.org/10.7326/M20-3569>.
- [4] H.J. Larson, et al., Measuring vaccine hesitancy: The development of a survey tool, *Vaccine* 33 (34) (2015) 4165–4175, <https://doi.org/10.1016/j.vaccine.2015.04.037>.
- [5] D. Sarathchandra, M.C. Navin, M.A. Largent, A.M. McCright, A survey instrument for measuring vaccine acceptance, *Prev. Med.* 109 (2018) 1–7, <https://doi.org/10.1016/j.ypmed.2018.01.006>.
- [6] M.B. Gilkey, B.E. Magnus, P.L. Reiter, A.-L. McRee, A.F. Dempsey, N.T. Brewer, The Vaccination Confidence Scale: A brief measure of parents’ vaccination beliefs, *Vaccine* 32 (47) (2014) 6259–6265, <https://doi.org/10.1016/j.vaccine.2014.09.007>.
- [7] D.J. Opel, et al., Development of a survey to identify vaccine-hesitant parents, *Hum. Vaccin.* 7 (4) (2011) 419–425, <https://doi.org/10.4161/hv.7.4.14120>.
- [8] P.G. Szilagyi, et al., Likelihood of COVID-19 vaccination by subgroups across the US: post-election trends and disparities, *Hum. Vaccines Immunother.* (2021) 1–6, <https://doi.org/10.1080/21645515.2021.1929695>.

- [9] M.C.S. Wong, et al., Acceptance of the COVID-19 vaccine based on the health belief model: A population-based survey in Hong Kong, *Vaccine* 39 (7) (2021) 1148–1156, <https://doi.org/10.1016/j.vaccine.2020.12.083>.
- [10] Y. Wang, M. McKee, A. Torbica, D. Stuckler, Systematic Literature Review on the Spread of Health-related Misinformation on Social Media, *Soc. Sci. Med.* 240 (2019) 112552, <https://doi.org/10.1016/j.socscimed.2019.112552>.
- [12] H.J. Larson, et al., The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey, *EBioMedicine* 12 (2016) 295–301, <https://doi.org/10.1016/j.ebiom.2016.08.042>.
- [13] G.K. Shapiro, et al., The vaccine hesitancy scale: Psychometric properties and validation, *Vaccine* 36 (5) (2018) 660–667, <https://doi.org/10.1016/j.vaccine.2017.12.043>.
- [14] A.S. Wallace, et al., Development of a valid and reliable scale to assess parents' beliefs and attitudes about childhood vaccines and their association with vaccination uptake and delay in Ghana, *Vaccine* 37 (6) (2019) 848–856, <https://doi.org/10.1016/j.vaccine.2018.12.055>.
- [15] M. Daly, A. Jones, E. Robinson, Public Trust and Willingness to Vaccinate Against COVID-19 in the US From October 14, 2020, to March 29, 2021, *JAMA* 325 (23) (2021) 2397, <https://doi.org/10.1001/jama.2021.8246>.
- [16] K. H. Nguyen et al., COVID-19 vaccination intent, perceptions, and reasons for not vaccinating among groups prioritized for early vaccination, United States, September 2020, Jan. 27, 2021. <https://www.cdc.gov/vaccines/imz-managers/coverage/adultvaxview/pubs-resources/COVID-online-report2020.html> (accessed Jul. 17, 2021).
- [17] CDC, Vaccine Hesitancy for COVID-19, 2021. <https://data.cdc.gov/stories/s/Vaccine-Hesitancy-for-COVID-19/cnd2-a6zw/> (accessed Jul. 21, 2021).
- [18] M. Daly, E. Robinson, Willingness to vaccinate against COVID-19 in the US: Longitudinal evidence from a nationally representative sample of adults from April-October 2020, *MedRxiv Prepr. Serv. Health Sci.*, p. 2020.11.27.20239970, Nov. 2020, <https://doi.org/10.1101/2020.11.27.20239970>.
- [19] M.B. Gilkey, P.L. Reiter, B.E. Magnus, A.-L. McRee, A.F. Dempsey, N.T. Brewer, Validation of the Vaccination Confidence Scale: A Brief Measure to Identify Parents at Risk for Refusing Adolescent Vaccines, *Acad. Pediatr.* 16 (1) (2016) 42–49, <https://doi.org/10.1016/j.acap.2015.06.007>.
- [20] A.A. Malik, S.M. McFadden, J. Elharake, S.B. Omer, Determinants of COVID-19 vaccine acceptance in the US, *EclinicalMedicine* 26 (2020) 100495, <https://doi.org/10.1016/j.eclinm.2020.100495>.
- [21] C. Giambi, et al., Parental vaccine hesitancy in Italy – Results from a national survey, *Vaccine* 36 (6) (2018) 779–787, <https://doi.org/10.1016/j.vaccine.2017.12.074>.
- [22] I. Elharazi, S. Al-Hamad, S. Alfalasi, R. Aldhaheri, E. Dubé, A.R. Alsuwaidi, Exploring vaccine hesitancy among healthcare providers in the United Arab Emirates: a qualitative study, *Hum. Vaccines Immunother.* 17 (7) (2021) 2018–2025, <https://doi.org/10.1080/21645515.2020.1855953>.
- [23] H. Lyu et al., Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination, *Intell. Med.*, p. S266710262100036X, Aug. 2021. <https://doi.org/10.1016/j.imed.2021.08.001>.
- [24] M. pui S. Chan, K.H. Jamieson, D. Albarracín, Prospective associations of regional social media messages with attitudes and actual vaccination: A big data and survey study of the influenza vaccine in the United States, *Vaccine* 38 (40) (2020) 6236–6247, <https://doi.org/10.1016/j.vaccine.2020.07.054>.
- [25] A.E. Leader, A. Burke-Garcia, P.M. Massey, J.B. Roark, Understanding the messages and motivation of vaccine hesitant or refusing social media influencers, *Vaccine* 39 (2) (2021) 350–356, <https://doi.org/10.1016/j.vaccine.2020.11.058>.
- [26] A.B. Wiyeh, S. Cooper, A. Jaca, E. Mavundza, D. Ndwandwe, C.S. Wiysonge, Social media and HPV vaccination: Unsolicited public comments on a Facebook post by the Western Cape Department of Health provide insights into determinants of vaccine hesitancy in South Africa, *Vaccine* 37 (43) (2019) 6317–6323, <https://doi.org/10.1016/j.vaccine.2019.09.019>.
- [27] D. Hu, C. Martin, M. Dredze, D.A. Broniatowski, Chinese social media suggest decreased vaccine acceptance in China: An observational study on Weibo following the 2018 Changchun Changsheng vaccine incident, *Vaccine* 38 (13) (2020) 2764–2770, <https://doi.org/10.1016/j.vaccine.2020.02.027>.
- [28] G.J. Kang, et al., Semantic network analysis of vaccine sentiment in online social media, *Vaccine* 35 (29) (2017) 3621–3638, <https://doi.org/10.1016/j.vaccine.2017.05.052>.
- [29] S. Martin, et al., 'Vaccines for pregnant women...?! Absurd' - Mapping maternal vaccination discourse and stance on social media over six months, *Vaccine* 38 (42) (2020) 6627–6637, <https://doi.org/10.1016/j.vaccine.2020.07.072>.
- [30] N. Ahmed, S.C. Quinn, G.R. Hancock, V.S. Freimuth, A. Jamison, Social media use and influenza vaccine uptake among White and African American adults, *Vaccine* 36 (49) (2018) 7556–7561, <https://doi.org/10.1016/j.vaccine.2018.10.049>.
- [31] B.L. Rosen, C. Wheldon, E.L. Thompson, S. Maness, P.M. Massey, Social media engagement association with human papillomavirus and vaccine awareness and perceptions: Results from the 2017 US Health Information National Trends Survey, *Prev. Med.* 138 (2020) 106151, <https://doi.org/10.1016/j.jpmed.2020.106151>.
- [32] H. Piedrahita-Valdés, et al., Vaccine Hesitancy on Social Media: Sentiment Analysis from June 2011 to April 2019, *Vaccines* 9 (1) (2021) 28, <https://doi.org/10.3390/vaccines9010028>.
- [33] N.F. Johnson, et al., The online competition between pro- and anti-vaccination views, *Nature* 582 (7811) (2020) 230–233, <https://doi.org/10.1038/s41586-020-2281-1>.
- [34] M.Z. Dudley, L. Privor-Dumm, È. Dubé, N.E. MacDonald, Words matter: Vaccine hesitancy, vaccine demand, vaccine confidence, herd immunity and mandatory vaccination, *Vaccine* 38 (4) (2020) 709–711, <https://doi.org/10.1016/j.vaccine.2019.11.056>.
- [35] È. Dubé, J.K. Ward, P. Verger, N.E. MacDonald, Vaccine Hesitancy, Acceptance, and Anti-Vaccination: Trends and Future Prospects for Public Health, *Annu. Rev. Public Health* 42 (1) (2021) 175–191, <https://doi.org/10.1146/annurev-publhealth-090419-102240>.
- [36] Twitter Developer, Standard search API | Docs | Twitter Developer Platform, 2021. <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets> (accessed Aug. 04, 2021).
- [37] S. Smys, J.M.R.S. Tavares, V.E. Balas, A.M. Ilyasu, Computational Vision and Bio-Inspired Computing ICCVBI 2019, in: ICCVBI: International Conference On Computational Vision and Bio-Inspired Computing, 2020. Accessed: Aug. 09, 2021. [Online]. Available: <https://doi.org/10.1007/978-3-030-37218-7>.
- [38] M. Munero, C.S. Montero, E. Sutinen, J. Pajunen, Are they different? affect, feeling, emotion, sentiment, and opinion detection in text, *IEEE Trans. Affect. Comput.* 5 (2) (2014) 101–111, <https://doi.org/10.1109/TAFFC.2014.2317187>.
- [39] CDC, Building Confidence in COVID-19 Vaccines | CDC, Jul. 08, 2021. <https://www.cdc.gov/vaccines/covid-19/vaccinate-with-confidence.html> (accessed Jul. 21, 2021).
- [40] S. Bird, E. Loper, E. Klein, *Natural Language Processing with Python*, O'Reilly Media Inc., 2009.
- [41] J. Wei, K. Zou, EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, pages 6382–6388, Hong Kong, November 3–7, 2019.
- [42] S. Loria, TextBlob: Simplified Text Processing, 2018. <https://textblob.readthedocs.io/en/dev/>.
- [43] C.J. Hutto, E. Gilbert, VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text, in: Eighth International AAAI Conference on Weblogs and Social Media, vol. 8, no. 1, 2014, pp. 10.
- [44] A. Rajaraman, J.D. Ullman, Mining of Massive Datasets, vol. 9781107015. 2011. <https://doi.org/10.1017/CBO9781139058452>.
- [45] K. Kowsari, et al., Text classification algorithms: a survey, *Inf. Switz.* 10 (2019), <https://doi.org/10.3390/info10040150>.
- [46] K. Kowsari, K.J. Meimandi, M. Heidarysafa, S. Mendu, L.E. Barnes, D.E. Brown, Text classification algorithms: a survey, *Inf. Switz.* 10 (4) (2019), <https://doi.org/10.3390/info10040150>.
- [47] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriching Word Vectors with Subword Information, *Trans. Assoc. Comput. Linguist.* 5 (2016) 135–146.
- [48] J. Pennington, R. Socher, C. Manning, Glove: Global Vectors for Word Representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 2014, pp. 1532–1543. <https://doi.org/10.3115/v1/D14-1162>.
- [50] F. Pedregosa et al., Scikit-learn: Machine Learning in Python, *Mach. Learn. PYTHON*, 2011, pp. 2825–2830.
- [51] F. Chollet, keras, GitHub., 2015. <https://github.com/fchollet/keras>.
- [52] J. Lever, M. Krzywinski, N. Altman, Classification evaluation, *Nat. Methods* 13 (8) (2016) 603–604, <https://doi.org/10.1038/nmeth.3945>.
- [53] Pareto Software, US Cities Database | Simplemaps.com, 2021. <https://simplemaps.com/data/us-cities> (accessed Aug. 04, 2021).
- [54] J. Byrnes, US won't join global coronavirus vaccine initiative, *TheHill*, Sep. 01, 2020. <https://thehill.com/policy/healthcare/514657-us-wont-join-global-coronavirus-vaccine-initiative> (accessed Jun. 20, 2021).
- [55] Johnson & Johnson Janssen COVID-19 Vaccine Pause, *Johns Hopkins Medicine*, 2021. <https://www.hopkinsmedicine.org/coronavirus/covid-19-vaccine/jj-vaccine-pause.html> (accessed Jun. 20, 2021).
- [56] A. Inserro, FDA Agrees to EUA for COVID-19 Vaccine from Pfizer, *BioNTech*, *AJMC*, Sep. 11, 2020. <https://www.ajmc.com/view/fda-agrees-to-eua-for-covid-19-vaccine-from-pfizer-biontech> (accessed Jun. 20, 2021).
- [57] I. Kwai, U.S. Starts Vaccine Rollout as High-Risk Health Care Workers Go First, *The New York Times*, Dec. 14, 2020. Accessed: Jun. 20, 2021. [Online]. Available: <https://www.nytimes.com/live/2020/12/14/world/covid-19-coronavirus>.
- [58] A. Sider, P. Loftus, United Begins Flying Pfizer's Covid-19 Vaccine, *Wall Street J.* 27 (2020). <https://www.wsj.com/articles/united-begins-flying-pfizers-covid-19-vaccine-11606512293>.
- [59] P. Barberá, G. Rivero, Understanding the Political Representativeness of Twitter Users, *Soc. Sci. Comput. Rev.* 33 (6) (2015) 712–729, <https://doi.org/10.1177/0894439314558836>.
- [60] J. Mellon, C. Prosser, Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users, *205316801772000*, *Res. Polit.* 4 (3) (2017), <https://doi.org/10.1177/2053168017720008>.