2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC) | 978-1-6654-2463-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/COMPSAC51774.2021.00098

Unmasking the Mask Debate on Social Media

Luca Cerbin, Jason DeJesus, Julia Warnken, Swapna S. Gokhale
Dept. of Computer Science & Engg.
Univ. of Connecticut, Storrs, CT 06269
{swapna.gokhale}@uconn.edu

Abstract

Masks are believed to slow the spread of Covid-19, and can prevent many deaths, yet this inexpensive, common sense public health measure has ignited a fierce debate in the U.S.. Opponents of masks or anti-maskers have resorted to measures such as organizing protests and marches to make their views public. They have also taken to social media platforms to vigorously argue against the use of masks, and spread misinformation, lies, and myths regarding their use. Even with the advent of vaccines, masks are still likely to be recommended for a long time. It is therefore necessary to identify those tweets that spread falsehoods regarding the use and effectiveness of masks in order to limit their appeal and damage. This paper proposes a classification framework to detect anti-mask tweets from social media dialogue shared on Twitter during the months of July and August 2020. The framework relies on popular machine learning models trained using a combination of linguistic, auxiliary, psycholinguistic and sentiment features for detection. The proposed classification framework can detect anti-mask tweets with excellent accuracy of over 90%, and hence, it can be used to tag tweets that sow misinformation about masks before they spread through the ether and influence people.

Keywords

Masks, Anti-mask, Pro-mask, Twitter, Classification, Machine learning.

I. Introduction & Motivation

The coronavirus pandemic has upended every single tenet and ritual of our modern society. Discussion and practice of measures such as masks, social and physical distancing, vaccines, hand hygiene, and disinfectants have now become a part of our daily routines. Of these, one of the most contentious issue that has bitterly divided the U.S. society is the wearing of masks. A seemingly simple act of wearing a facial covering that covers both the mouth and the nose serves as a stark reminder of the pandemic, and has also been the topic of a fierce debate. Proponents of masks point to several studies that recommend their use to slow the spread of Covid-19 [14]. Opponents, however, contend that most of the studies have looked at the use of face masks in health care, and not community settings. They further claim that these studies were observational, not the gold standard of science, which is randomized controlled trials. It does not help that early in the pandemic public health officials in the U.S. discouraged the use of masks by the general public. At the time "mass masking" was not recommended either by the CDC or the WHO, perhaps to conserve them for healthcare and other front line workers [8]. Later, however, they backtracked from this initial position and vigorously advocated the use of masks to blunt the spread of the virus and prevent deaths. The U-turn regarding masks and the subsequent political divide over them has come to symbolize the chaos of the U.S. response to the still-raging pandemic [38].

Expressions of pro-mask and anti-mask opinions are plentiful and varied in the physical, offline world. In some counties, where the coronavirus has surged out of control, mask mandates have been imposed and this has further outraged their residents. Those opposed to mask mandates have staged protests, and one local health official had to even quit her job after receiving a death threat for a mask order [24]. In addition to expressing their views through their actions by either wearing or not wearing masks in public spaces and/or organizing protests, people have often turned to social media platforms such as Twitter and Facebook to express their support or opposition to masks. These social media platforms have not only been woven tightly into the fabric of our society, but sharing on these platforms has skyrocketed especially during the pandemic, because a number of people are either in self-imposed or government-mandated isolation and lockdown. Therefore, in addition to the offline expression of the pro- and anti-mask opinions, this debate over masks has been playing out vociferously over these platforms as well.

Compliance with masks has been spotty at best through the U.S., even though the CDC and other public health experts have repeatedly indicated, on multiple occasions, that wearing masks could save a significant number of lives [10], [17]. Furthermore, the use of masks is likely to continue despite the approval and roll out of vaccines. In fact, masks and social distancing will probably be recommended at least for a while, because a lot is still unknown about what protections vaccines can afford in terms of preventing infections, their severity and their spread [33]. It is thus believed that masks are and will continue to be an effective tool against fighting the pandemic. Given the vitality of masks, it is then imperative to understand the public outlook towards their use. Based on such understanding we can launch educational and public awareness initiatives to dispel the myths and misinformation, and encourage their adoption

978-1-6654-2463-9/21/$31.00 ©2021 IEEE
DOI 10.1109/COMPSAC51774.2021.00098

broadly. Moreover, understanding the drivers and spread of misinformation can be valuable during future pandemics.

The novelty of this paper lies in developing a classification framework that can detect anti-mask misinformation and lies from Twitter dialogue. Detecting and tagging such misinformation is especially important as social media users are more likely to believe falsehoods about Covid-19 and ignore public health advice [32]. Based on the data collected and labeled during July and August 2020 using anti- and pro-mask hashtags, we extract linguistic, auxiliary, psycho-linguistic, sentiment and social features. We employ a combination of these features to train popular machine learning models. Most ML classifiers, including Support Vector Machines, Random Forest, Gradient Boosting, achieve an accuracy of over 90% in separating the anti-mask tweets from the pro-mask ones. Importance analysis shows that a bulk of the contribution towards classification comes from the text of the tweets, and from the social parameters that indicate the reach and popularity of the tweets and the tweeters.

The rest of the paper is organized as follows: Section II explains the process of collecting and preparing the data. Section III presents the sequence of steps involved in building the classification framework. Section IV discusses the results. Section V compares and contrasts related research. Section VI offers concluding remarks and directions for future research.

II. Data Preparation

This section discusses three steps in the preparation of data: data collection, data labeling, and data pre-processing.

A. Data Collection

Data was collected twice, one month apart, using the crawling seeds *#wearadamnmask*, *#nomaskforme*, *#maskupamerica*, *#masksareforsheep*, *#nomasks*, *#nomaskmandate*, *#antimaskers*, *#maskitorcasket* in July 2020 and August 2020. These two time frames were chosen as they represented two significant epochs in the mask debate. In July 2020, as the country was emerging from the lockdown, masks were viewed as a way to restore a sense of normalcy. Furthermore, masks came into sharp spotlight in this one-month period because of the tussle surrounding the reopening of schools, and students returning to college campuses. Masks also became a hot button issue during this period when the Democratic presidential candidate Joe Biden suggested that if elected he will issue a national mask mandate [29]. In the same period, leading public health experts, including the CDC promoted the use of masks as "life-saving", highlighting that if everyone committed to wearing masks, we could save a significant number of American lives [10]. Thus, the two data collection epochs one month apart occurred during an eventful period for the fate of the masks and their acceptance. Both data sets were collected using the rtweet library in R [21].

B. Data Labeling

This set of crawling seeds was harvested because it included both the anti-mask and pro-mask perspectives. For example, we expected that hashtags such as *#maskupamerica* and *#maskitorcasket* would be used in tweets that support masks, whereas hashtags such as *#nomasksforme* and *#masksareforsheep* would be used to show opposition. We anticipated that the tweets would neatly separate according to support and opposition, consistent with the corresponding hashtags. Such clear, neat separation would obviate the need for manual labeling and facilitate weak supervised learning with the hashtags serving as labels. Skimming through the tweets, however, invalidated this assumption and many hashtags were creatively embedded in both supporting and opposing tweets.

Manual annotation of the tweets seemed inevitable, and was undertaken to classify each tweet into one of two groups –'A' for anti-mask, and 'P' for pro-mask. The entire data set was labeled twice, independently, with a gap of about one week between the two labelings. Duplicates were eliminated before the labeling. Only those tweets where the labels matched on two independent occasions were included in the final corpus, which consisted of 4042 tweets. About 500 tweets were eliminated because of mismatch of labels. In the corpus, about 57% of the tweets are pro-mask, and 43% are anti-mask. This data also contained a number of public safety announcements (PSAs) from schools, colleges and sports teams. There were tweets that expressed political opinion regarding the conventions, wildfires in California, and the BLM protests without the express mention of masks other than the hashtag. In the manual labeling process, we eliminated these tweets to build a high quality data set that truly reflects the public opinion about masks instead of other peripheral and allied political issues.

C. Data Pre-processing

The labeled data was converted to UTF-8 encoding, and transformed to lower case. Then, numbers, punctuation and stop words were removed. After word stemming and stripping white space, domain specific words that occur in both pro-mask and anti-mask tweets with a similar frequency were removed as they are likely to be uninformative.

III. Tweet Classification

Anti-mask tweets can propagate discordant information, and their false narrative can easily convince people to question and forgo the commonsense public health measure. These tweets must thus be detected and tagged in a timely manner to limit their damage. However, given the excessive volume of content that gets shared on these platforms, manual separation of anti-mask tweets is impossible, highlighting the need for automated detection. This section presents a classification approach to distinguish between pro-mask and anti-mask tweets, labeled as 'P' and 'A' respectively.

A. Feature Extraction

The first step is to extract features that abstract away the important properties of the tweets while ignoring the unnecessary details. We considered linguistic, auxiliary, social, psycho-linguistic, and sentiment features as discussed below in the classification framework.

1) Linguistic Features: Tweets were processed using natural language techniques so that the key features including the semantic relationship between the words and the contextual information of the words and sentences were numerically encoded in high-dimensional vectors. We considered a number of vector representations such as bag-of-words [40], n-grams, Term Frequency-Inverse Document Frequency (TF-IDF) [26], and word embeddings [27] that are commonly used for classification. Of these, we used the n-grams/TF-IDF and word embeddings.

In the n-grams method, a sample of text is represented by the most frequent instances of every unique n continuous words as a dimension. The tweets were represented through unigram (1-gram) vectors, and the weight for each unigram is its TF-IDF score which is given by:

$$TF - IDF = tf * log(\frac{T}{df}) \tag{1}$$

In Equation (1), tf is the number of times a particular term occurs in a tweet, T is the total number of tweets, and df is the number of tweets containing that particular term. The main advantage of a $TF - IDF$ score over the simple frequency counts of the n-gram method is that it assigns a higher weight to the terms that occur more frequently through the entire data set. Thus, the $TF-IDF$ score should assign a higher weight to those phrases that are the most important in determining whether a tweet is anti-mask or pro-mask. After pre-processing, the size of our corpus (number of unique words) is over 9000. Of these, we calculated the TF-IDF vector representations of the top 2000 most relevant unigrams. We used the TF-IDF implementations from the NLTK library to extract these features [22].

Although the TF-IDF score provides a differentiated representation of the words based on their frequency of occurrence, it does not preserve any relationship between the words. Word embeddings are a powerful technique that represent semantically related words as closely related vectors. Words with similar meanings are mapped to low-dimensional, non-sparse vectors that exist near each other in a pre-defined vector space. A good word embedding can preserve the contextual information behind words in a tweet that a n-gram/TF-IDF scheme cannot. We use Word2Vec, which is a popular technique to create distributed numerical representations of word features using a two-layer neural network with back propagation [27]. Word2vec trains words against other words that neighbor them in the input corpus. Word2Vec allows us to encode the context of a given word by including information about preceding and succeeding words in the vector that represents a given instance of a word. Therefore, the results obtained from using Word2vec may result in a much better classification.

We implemented Word2Vec using the gensim library [34]. From the pre-processed tweets, we generated a list of tokens, and built a model to represent each word by a 10-dimensional vector, where the parameter *min_count* is 1. The number of workers, which is the number of partitions during testing is 8. The model considers all the words in the corpus. We created the vector representations for all the tokens, and the total number of epochs used is 25. We used the continuous bag of words (CBOW) model to generate the representations. The other option was to use the skip gram model. Skip gram works well with a small amount of data and is found to represent rare words well. On the other hand, CBOW is faster and has better representations for more frequent words [20], [30]. We chose CBOW based on our earlier success with this model to classify the anti-vaxx dialogue [31].

We also included POS (part-of-speech) tagging using the NLTK library [22]. The NLTK library provides the ability to classify each word as one of 35 parts of speech. POS tagging occurred before removing stop words to capture any differences in the raw text. The occurrences of each part of speech is counted for each tweet and fed as input to our models.

2) Psycho-Linguistic Features: Some studies show that refusal to wear masks may be linked to sociopathic, narcissistic and psychopathic tendencies [39]. These leanings are reflected in an excessive use of first person pronouns "I" and "me" in written and spoken language. Therefore, we considered the use of these first person pronouns in the anti-mask and pro-mask tweets. The use of these pronouns, however, did not appear significantly different in these two groups. In total, the pro-mask tweets used "I" 8 times compared to the use of "I" 14 times in the anti-mask tweets. Counting the instances of both "I" and "me", the pro-mask tweets had 103 occurrences, while the anti-mask tweets had 102. Because the differences appeared insignificant, these first person pronouns were not considered further in the classification.

3) Auxiliary Features: Written texts including social media feeds do not carry with them clues that can be gathered from facial expressions and body language that accompany face-to-face or spoken communication. Therefore, in social media texts, users may use a variety of punctuation marks and other means such as hashtags and emoticons to emphasize their point. These auxiliary features are believed to somewhat substitute the clues that can be learned from communicating in the physical space, and are known to improve classification accuracy [11]. Therefore, we included numbers of hashtags, mentions, punctuations, links, words in upper case letters, question marks, exclamation marks, periods, and quotations as features.

4) Social Features: We used the social features that quantify the reach of the tweets and the tweeters in the classification framework. These include tweet length, favorite count, quoted favorite count, retweet count, followers count, quoted statuses count, quoted friends count, quoted retweet count, list count, statuses count, followers and friends count. Because the values of these features differed widely, we transformed each feature using the MinMaxScaler in sklearn library [1]. This function scales and translates each feature individually such that it lies in the range of 0 and 1. This transformation is often used as an alternative to zero mean, unit variance scaling [1]. We also included binary indicators of whether a tweet is from a verified account, whether it quotes other tweet, whether it is a reply to existing tweet, and if it mentions other users as features.

5) Sentiment Features: Textblob [23] and Vader [19] sentiment scores, computed for each original tweet (before pre-processing) were used in the classification. TextBlob calculates the sentiment polarity for each tweet, which ranges from -1 to $+1$, where -1, 0 and $+1$ indicate negative, positive and neutral sentiment respectively. Vader computes a compound score as a normalized and weighted composite score obtained by analyzing each word in a tweet for the direction of its sentiment - a negative (positive) valency for negative (positive) sentiment. It therefore ranges from -1 to +1 depending on the net sentiment of the tweet.

B. ML Models

We considered the following popular supervised machine learning models for classification. Scikit implementations of these models were used [6], and the parameters chosen are listed below.

- **Random Forests:** The number of trees was set to 100, the number of features in each tree was equal to the square-root of the number of total features, and each decision tree was allowed to grow fully up to its leaves.
- **Support Vector Machines (SVMs):** We used SVMs with RBF kernel, the regularization parameter C is set to 1000, and kernel coefficient gamma is set to 0.01.
- **Multi-Layer Perceptron (MLP):** Multi-Layer Perceptron (MLP) is one of the feed-forward Artificial Neural Networks (ANN) with the numbers of neurons in input, hidden and output layers set to 10, 8, 5, and 2. We used rectifier linear unit (ReLu) activation function.
- **Gradient Boosting (GB):** The number of trees is set to 1600, the fraction of observations to be selected for each tree (subsample) is set to 0.55, the maximum depth of each tree is set to 5, the minimum samples in each leaf is set to 1, the learning rate which determines the impact of each tree on the final outcome is set to 0.05.
- **Long Short-Term Memory (LSTM):** LSTM is an artificial recurrent neural network architecture used in deep learning [18]. We used Keras library to implement the model [9]. We truncated and pad the input sequences to 360. The first layer uses vectors of length 100 to represent each word, the next layer is the LSTM layer with 100 memory units (smart neurons). We use a dense output layer with a single neuron and a sigmoid activation function. Binary cross entropy is used as the loss function. The efficient ADAM optimization algorithm was used, and it uses batch sizes of 64 and 100 epochs.

C. Performance Metrics

Our objective is to identify anti-mask tweets, and hence, to define the performance metrics, we designate the anti-mask and pro-mask classes as positive and negative respectively. Tweets can thus be classified into four groups – true positive (TP) (anti-mask labeled anti-mask), true negative (TN) (pro-mask labeled pro-mask), false positive (FP) (pro-mask labeled anti-mask), and a false negative (FN) (anti-mask labeled pro-mask). These four groups lead to the following metrics to compare classifier performance:

- *Accuracy:* Accuracy is defined as the percentage of tweets that are labeled correctly.
- *Precision:* Precision measures the percentage of the tweets that are actually anti-mask out of all the tweets that are predicted as anti-mask.
- *Recall:* Recall measures how many of the anti-mask tweets are actually labeled as anti-mask.
- *F-score:* F-score seeks a balance between Precision and Recall.

Precision is the percentage of relevant from the set detected and recall is the percentage of relevant from within the global population [25]. Precision is an important measure to determine when the costs of a false positive are high. Applying symmetrical logic, recall would be the metric of significance when the costs of a false negative are high. In the context of detecting anti-mask tweets, false positive labeling implies that a pro-mask tweet is labeled as anti-mask, whereas a false negative labeling implies that an anti-mask tweet is labeled as pro-mask. In false positive labeling, because a pro-mask tweet may be labeled as anti-mask it may be subject to actions such as being censored or tagged for misinformation. However, any additional stringent punitive actions such as removing the tweet altogether may lead to freedom of speech violations. In false negative labeling, an anti-mask tweet will slip through the cracks and will not be tagged for carrying misinformation. While such mislabeling may cause damage by spreading discordant information, it will not lead to any violations of people's individual rights. Therefore, in this problem, precision may be a more important metric than recall. A balance may also be sought between precision and recall to trade off infringing freedom of speech against the spread of discordant information.

IV. Results and Discussion

We split the entire corpus using stratified sampling into two partitions; the training partition consisted of 75% and the testing partition contained 25% of the tweets. All the models listed in Section III-B, except for LSTM, were trained and tested on a combination of linguistic, auxiliary, social and sentiment features. LSTM was fed pre-processed text directly along with auxiliary and social features. We combined all the features for model training, guided by our success in their use in detecting tweets that spread vaccine misinformation [31]. The results of the performance metrics for all the models are noted in Table I.

The table shows that all the classifiers except for SVM can distinguish between anti-mask and pro-mask tweets with

Model	Accuracy	Precision	Recall	F1-Score
RF	95.64	0.9359	0.9913	0.9628
LSTM	93.66	0.9305	0.9640	0.9458
SVM	89.81	0.9056	0.9224	0.9139
GB	95.71	0.9647	0.9441	0.9780
MLP	94.46	0.9382	0.9672	0.9525

TABLE I: Performance of ML Models

accuracy and F1-score over 90%. Moreover, the accuracy of the SVM is only slightly lower than 90%. For some models, the accuracy reaches as high as 96%. These results show that anti-mask tweets that can sow discordant information about masks, and promote non compliance can be accurately separated from social media dialogue. They also show that this accuracy can be achieved even after data from different time periods is combined. Each time period presents a different context or a backdrop against which this dialogue played out; in July it was lifting the lockdowns, and in August it was reopening schools and restarting the sports and other activities. However, without regard to the underlying background information, pro- and anti-mask sentiment can be detected.

We use the Random Forest model to determine the importance scores of the various types of features. The relative scores are summarized in Table II. The table indicates that bulk of the contribution, around 83%, comes from the text of the tweets (includes TF-IDF plus word embeddings plus POS tags). Social features which determine the reach of the tweet and the popularity and level of activity of the tweeters contribute about 14%. Sentiment scores have very little contribution, around 3%.

Feature Type	Importance Score
TF-IDF	0.4666
Embeddings	0.2598
Social Features	0.1398
POS Tags	0.1036
Sentiment	0.0310
Auxiliary	0.000

TABLE II: Importance Scores for Feature Types

V. Related Research

Social media conversations are spontaneous and unfiltered, and hence, can offer genuine insights into people's opinions on a variety of offline events, topics, and policies. Because the donning of masks is relatively recent, very few efforts have analyzed social media conversations around masks. Ahmed *et. al.* [4] build a network of users from mask-related conversations on Twitter, and analyze this network using centrality measures to find the most influential users. Even when face masks were recommended, there remained widespread confusion about who should be wearing a mask – whether healthy people should be wearing it, and for whose protection [36]. A geographical analysis of anti-mask activity based on Twitter content has also been conducted [37].

Overall, social media feeds have been mined to understand the public outlook on hot button medical and other health-related issues, the most notable topic that is related to masks is vaccines. The issue of masks and vaccines are inextricably linked together in the Covid world, especially, because it is believed that there is a significant overlap between anti-vaxxers and anti-maskers. Therefore, we review the work on identifying anti-vaxx dialogue on social media as closely related to this work.

Research at the intersection of vaccines and social media use both unsupervised and supervised learning for harnessing informal opinions, and also classify these perceptions into supporting or opposing. Some works consider specific vaccines such as Dengavaxia [2], MMR [3], Flu [7], and Zika [13], while some mine general attitudes about vaccines (anti-vaxx opinions, adversity and safety signals, fake news and rumors, and interference from trolls) without reference to any particular vaccine [16], [28], [5], [26], [12], [41], and recently the Covid-19 vaccine [35], [31]. To the best of our knowledge, no research has yet reported on understanding the dialogue on masks on social media platforms.

VI. Conclusions and Future Research

This paper analyzes the debate around masks on Twitter using the tweets collected during the months of July and August 2020, just as many states were beginning to lift their stay home, stay safe orders, and plans were being conceived to reopen schools. A classification framework is built which can differentiate between the two groups of tweets with an accuracy over 90%. The classification framework, by the virtue of separating anti-mask tweets from pro-mask ones can label tweets that sow such incorrect information about masks. Such labeling can warn other users that the views promoted by these tweets are not mainstream, and detrimental to public health.

Longitudinal analysis of the mask dialogue, with data collected at several other points during the pandemic, especially after President Trump was hospitalized due to Covid-19 is a topic of the future. A detailed topic modeling [15] framework to discover both the pro- and anti-mask themes, similar to pro-vaxx and anti-vaxx themes is also underway. Finally, collecting data from other social media platforms such as Facebook, and incorporating it in the analysis is also ongoing.

References

[1] "sklearn.preprocessing.MinMaxScaler". https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html, 2007.

[2] A. B. C. Abrigo and M. R. J. E. Estuar. "A Comparative Analysis of N-Gram Deep Neural Network Approach to Classifying Human Perception on Dengavaxia". In *Proc. of 2nd Intl. Conf. on Information and Computer Technologies*, pages 46–51, 2019.

[3] A. Addawood. "Usage of Scientific References in MMR Vaccination Debates on Twitter". In *Proc. of IEEE/ACM Intl. Conf. on Advances in Social Network Analysis and Mining (ASONAM)*, pages 971–979, 2018.

[4] W. Ahmed, J. Vidal-Alaball, F. L. Segui, and P. A. Moreno-Sanchez. "A Social Network Analysis of Tweets Related to Masks during the Covid-19 Pandemic". *Intl. Journal on Environ Res Public Health*, 17(21), November 2020.

[5] M. Asghari, D. Sierra-Sosa, and A. Elmagharby. "Trends on Health in Social Media: Analysis using Twitter Topic Modeling". In *Proc. of Intl. Symposium on Signal Processing and Information Technology*, pages 558–563, December 2018.

[6] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and Gaël Varoquaux. "API Design for Machine Learning Software: Experiences from the Scikit-learn project". In *Proc. of ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[7] A. Celesti, A. Galletta, F. Celesti, M. Fazio, and M. Villari. "Using Machine Learning to Study Flu Vaccines Opinions of Twitter Users". In *Proc. of Fourth Intl. Workshop on ICT Solutions for Health*, pages 1103–1106, 2019.

[8] K. K. Cheng, T. H. Lam, and C. C. Leung. "Wearing Face Masks in the Community during the Covid-19 Pandemic: Altruism and Solidarity". *Lancet*, April 2020.

[9] Francois Chollet et al. Keras, 2015.

[10] Centers for Disease Control and Prevention. "CDC Calls on Americans to Wear Masks to Prevent Covid-19 Spread". `https://www.cdc.gov/media/releases/2020/p0714-americans-to-wear-masks.html`, July 2020. Accessed: 2021-01-21.

[11] E. Forslid and N. Wiken. "Automatic Irony- and Sarcasm Detection in Social Media". Master's thesis, Uppsala Univeritet, August 2015.

[12] B. Gambini. "Study: Fake Russian Twitter Accounts Politicized Discourse About Vaccines". `http://www.buffalo.edu/news/releases/2020/03/024.html`, March 2020. Accessed: 2020-06-28.

[13] A. Ghenai and Y. Mejova. "Catching Zika Fever: Application of Crowdsourcing and Machine Learning for Tracking Health Misinformation on Twitter". In *Proc. of Intl. Conf. on Healthcare Informatics*, 2017.

[14] M. Godoy. "Yes, Wearing Masks Helps, Here's Why". `https://www.npr.org/sections/health-shots/2020/06/21/880832213/`, June 2020. Accessed: 2021-01-21.

[15] S. Gokhale. "Monitoring the Perception of Covid-19 Vaccine using Topic Models". In *Proc. of the 13th IEEE Intl. Conf. on Social Computing and Networking*, Virtual, Due to Covid-19, December 2020.

[16] S. K. Habibabadi and P. D. Haghighi. "Topic Modelling for Identification of Vaccine Reactions in Twitter". In *Proc. of ACSW*, Sydney, Australia, January 2019.

[17] Healthline.com. "As Many As 130,000 Lives Could Be Saved the Next 3 Months If Everyone Wore a Mask". `https://www.healthline.com/health-news/as-many-as-130000-lives-could-be-saved-the-next-3-months-if-everyone-wore-a-mask`, November 2020. Accessed: 2021-01-21.

[18] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory". *Neural Computation*, 9(8), November 1997.

[19] C. J. Hutto and C. Gilbert. "Vader: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text". In *Proc. of Intl. AAAI Conference on Weblogs and Social Media*, 2014.

[20] D. Karani. "Introduction to Word Embedding and Word2Vec". `https://towardsdatascience.com/introduction-to-word-embedding-and-word2vec`, September 2018.

[21] M. Kearney. "Collecting Twitter Data". `https://cran.r-project.org/web/packages/rtweet/rtweet.pdf`, January 2020.

[22] E. Loper and S. Bird. "NLTK: The Natural Language Toolkit". *CoRR*, cs.CL/0205028, 2002.

[23] S. Loria, P. Keen, M. Honnibal, R. Yankovsky, D. Karesh, E. Dempsey, et al. "Textblob: Simplified Text Processing". *Secondary TextBlob: Simplified Text Processing*, 3, 2014.

[24] V. Luckerson. "How Mask Mandates were Beaten Down in Rural Oklahoma". `https://www.newyorker.com/news/dispatch/how-mask-mandates-were-beaten-down-in-rural-oklahoma`, May 2020. Accessed: 2021-01-21.

[25] S. MacAvaney, H. R. Yao, E. Yang, K. Russell, N. Goharian, and O. Frieder. "Hate Speech Detection: Challenges and Solutions". *PLOS ONE*, August 2019.

[26] R. Mahajan, W. Romine, M. Miller, and T. Banerjee. "Analyzing Public Outlook towards Vaccination using Twitter". In *Proc. of Intl. Conf. on Big Data*, pages 2763–2772, December 2019.

[27] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. "Distributed Representations of Words and Phrases and Their Compositionality". *Neural Information Processing Systems*, 2013.

[28] T. Mitra, S. Counts, and J. W. Pennebaker. "Understanding Anti-Vaccination Attitudes in Social Media". In *Proc. of the Tenth Intl. AAAI Conf. on Web and Social Media*, pages 269–278, 2016.

[29] S. Morrison. "Biden Wants A National Mask Mandate. Can He Do That?". `https://www.vox.com/2020/8/21/21395570/biden-mask-mandate-for-all-national-states`, August 2020. Accessed: 2021-01-21.

[30] C. Nicholson. "A Beginner's Guide to Word2Vec and Neural Word Embeddings". `https://pathmind.com/wiki/word2vec`, 2019.

[31] N. Paul and S. Gokhale. "Analysis and Classification of Vaccine Dialogue in the Coronavirus Era". In *Proc. of IEEE Big Data Workshop on Smart and Connected Communities*, Virtual, Due to Covid-19, December 2020.

[32] S. Perry. "Social Media Users are More Likely to Believe False Information About Covid-19 and to Ignore Public Health Advice, Study Suggests". `https://www.minnpost.com/second-opinion/2020/07/social-media-users-are-more-likely-to-believe-false-information-about-covid-19-and-to-ignore-public-health-advice-study-suggests/`, July 2020. Accessed: 2021-01-21.

[33] The Associated Press. "Can I Stop Wearing a Mask After Getting a Covid-19 Vaccine". `https://apnews.com/article/mask-wearing-after-coronavirus-vaccine-f69b720444bd08565bf1f3a42e4a24ef`, December 2020. Accessed: 2021-01-21.

[34] R. Řehůřek and P. Sojka. "Software Framework for Topic Modelling with Large Corpora". In *Proc. of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, May 2010.

[35] R. F. Sear, N. Velasquez, R. Leahy, N. Johnson Restrepo, S. E. Oud, N. Gabriel, Y. Lupu, and N. F. Johnson. "Quantifying Covid-19 Content in the Online Health Opinion War using Machine Learning". *IEEE Access*, 8:91886–91893, May 2020.

[36] Y. Shen. "Covid-19 Outbreak: Tweet Analysis on Face Masks". `https://towardsdatascience.com/covid-19-outbreak-tweet-analysis-on-face-masks-27ef5db199dd`, March 2020. Accessed: 2021-01-21.

[37] Knau Staff. "Twitter Analysis Shows Arizona is #1 in Anti-Face Mask Activity". `https://www.knau.org/post/twitter-analysis-shows-arizona-1-anti-face-mask-activity`, July 2020. Accessed: 2021-01-21.

[38] A. Taylor. "How the Split Over Masks Sums Up America's Chaotic Coronavirus Response". `https://www.washingtonpost.com/world/2020/06/25/face-masks-america-divided/`, June 2020. Accessed: 2021-01-21.

[39] A. Whalen. "Narcissists and Psychopaths Are More Likely to Refuse to Wear Masks, Says New Research". `https://www.newsweek.com/narcissists-psychopaths-face-mask-requirement-mandate-social-distancing-covid-19-coronavirus-1519732`, July 2020. Accessed: 2021-01-21.

[40] R. Zafarani, M. A. Abbasi, and H. Liu. *Social Media Mining: An Introduction*. Cambridge University Press, New York, NY, USA, 2014.

[41] W. Zhao. "Misinformation Correction Across Social Media Platforms". In *Proc. of Intl. Conf. on Computational Science and Computational Intelligence*, pages 1371–1376, 2019.