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.
Can social media data be used to evaluate the risk of human interactions during the COVID-19 pandemic?

Lingyao Li\textsuperscript{a}, Zihui Ma\textsuperscript{a}, Hyesoo Lee\textsuperscript{b}, Sanggyu Lee\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Department of Civil and Environmental Engineering, A. James Clark School of Engineering, University of Maryland, College Park, MD, USA
\textsuperscript{b} University of Maryland School of Dentistry, Baltimore, MD, USA

\section{1. Introduction}

Coronavirus Disease 2019 (COVID-19) pandemic has resulted in 417\textsuperscript{k} deaths in the U.S. as of January 22, 2021 \cite{1}. Current evidence suggests that COVID-19 spreads mainly through direct contact or in close proximity with the infected patients via respiratory droplets as they cough, talk, or sneeze \cite{2}. To prevent the spread of COVID-19, the U.S. government has taken multiple measures, including implementing orders such as lockdown and social distancing \cite{3}. Social distancing is an intervention to reduce close interactions among people, and it has been demonstrated as an effective way to reduce the transmission of COVID-19 \cite{4}. Evaluating the risk of close human interactions is critical for the U.S. government agencies as it helps to understand the public compliance with these essential measures and evaluate the effectiveness of these orders to reduce the spread of COVID-19.

Researchers have evaluated the risk of human interactions based on questionnaire surveys or mobile data. Questionnaire surveys investigated people’s perceptions, behaviors, and plans during the COVID-19 pandemic by asking the respondents whether they are self-quarantining, keeping social distancing, eating out, or visiting relatives and friends \cite{5}. However, collecting the surveys from the respondents is often time-consuming \cite{6}. Other researchers leveraged mobile location data to establish an interactive monitoring system with advanced data processing techniques \cite{7,8}. These mobile data-based systems show the likelihood of viral transmission and provide tools for tracking human interactions with quantitative parameters, such as social distancing indices. However, acquiring the mobile location data often relies on third-party institutions and data distributors, which adds to the cost of implementation.

On social media platforms, people express their opinions and concerns about the COVID-19 government measures such as stay-at-home orders, and these can be used to reveal as well as to predict the public reaction to the measures \cite{9-11}. For example, people who express their fear and anxiety about the pandemic may have a tendency to follow the lockdown measures, whereas people who post that the measures are merely a political play and have a devastating impact on the economy may not follow them. Such risk perception shown on social media can reflect individuals’ behaviors. It has been indicated that the geographic regions with many social media users who opposed the lockdown policy had a higher frequency of close human interactions, thus increasing the

\begin{abstract}
The U.S. has taken multiple measures to contain the spread of COVID-19, including the implementation of lockdown orders and social distancing practices. Evaluating social distancing is critical since it reflects the risk of close human interactions. While questionnaire surveys or mobility data-based systems have provided valuable insights, social media data can contribute as an additional instrument to help monitor the risk of human interactions during the pandemic. For this reason, this study introduced a social media-based approach that quantifies the pro/anti-lockdown ratio as an indicator of the risk of human interactions. With the aid of natural language processing and machine learning techniques, this study classified the lockdown-related tweets and quantified the pro/anti-lockdown ratio for each state over time. The anti-lockdown ratio showed a moderate and negative correlation with the state-level social distancing index on a weekly basis, suggesting that people are more likely to travel out of the state where the higher anti-lockdown level is observed. The study further showed that the perception expressed on social media could reflect people’s behaviors. The findings of the study are of significance for government agencies to assess the risk of close human interactions and to evaluate their policy effectiveness in the context of social distancing and lockdown.

\end{abstract}
risk of viral transmission [12].

Public and private sectors need the information of public responses to assess and manage the risk of close human interactions amid the COVID-19 pandemic. While survey-based analysis and mobile data-based systems have provided multiple benefits in tracking human interactions and indicating the risk of viral transmission, leveraging the opinions expressed on social media can offer an additional and novel resource of such information. While social media data are inherently imperfect, it provides geographically well-distributed information that reflects people’s inclination to comply with the government measures. In particular, social media data are available in a large quantity almost instantaneously, which allows access to the most up-to-date public responses toward the lockdown measure. Moreover, social media data are cost- and time-efficient as they are easy to access and relatively cheap.

For these reasons, this study aims to develop a method that evaluates the state-level risk of human interactions by extracting the perceptions of Twitter users upon the lockdown measures. This study applies natural language processing techniques and machine learning classifiers to identify users’ perceptions. It then conducts a spatial-temporal analysis and a correlation analysis to investigate the associations of anti-lockdown perceptions with the case increase and social distancing index. The proposed approach can provide an additional resource to assess the state-level risk amid the COVID-19 pandemic that further benefits government agencies and health institutions.

2. Literature review

2.1. Social media in managing infectious disease

The extensive application of social media opens up useful insights for decision-makers to manage infectious diseases. One primary research branch focuses on discussing how social media can aid communication. A study conducted by Kim and Hawkins [13] revealed that social media could enhance public awareness and encourage health prevention behaviors. In particular, social media can promote preventive hygiene intention. Yang and Sun [14] investigated China’s Health Code policy under the COVID-19 pandemic to explore why the public performs voice behavior on social media. Their study showed that the public voice plays a critical role in developing policies and represents cooperation between the government and the businesses to maintain social stability.

Another popular research field leverages social media data to signal early warning or conduct surveillance for infectious diseases [15–18]. For example, Aramaki et al. [18] scraped millions of influenza-related tweets and applied machine learning classifiers to classify tweets into either influenza-related (positive) or -unrelated (negative). Their study showed a high correlation between the influenza epidemics and tweet data and thus concluded that Twitter texts could reflect the real-world infections. Ginsberg et al. [19] discussed how to use Google search queries to track influenza-like illness amid the epidemic. Similar to this study, a recent article validated the credibility of using internet searches and social media data to predict COVID-19 outbreak by displaying a strong correlation between search indices with subsequent infections [20].

2.2. Social media in understanding People’s thoughts and perceptions

Social media postings contain a great deal of textual information that can reflect people’s perceptions, thoughts, concerns, and mental health. As people’s behaviors are in close connection with their thoughts, the information collected from social media may provide a useful resource for decision-makers to understand the risk of people’s behaviors amid the health crisis. This hypothesis is built upon the existing body of knowledge called Cognitive Behavioral Theory (CBT). According to the CBT, how people think is closely connected with how they behave: thoughts determine how people feel, and it affects how they behave, and their behaviors ultimately impact the situation [21].

Multiple studies have provided empirical evidence to support this theory, especially the relationship between risk perceptions and behaviors. A particular study [22] suggested that the perceived risk can be an antecedent of a specific behavior. For example, risky drivers perceive driving risks as low, and those who perceive risks as high are less likely to undertake dangerous driving behavior. The relationship between the perceived risk and behaviors have been studied in education [23], business [24–26], and other sociology and psychology fields [27–29]. In the context of the COVID-19 pandemic, recent publications reported that high-risk perception could promote individuals’ protective behaviors such as following social distancing guidelines, washing their hands, and wearing masks to reduce close contact with viruses [30,31]. Researchers have also demonstrated that people’s traveling patterns were influenced by their risk perception [32,33].

Since the outbreak of COVID-19, many researchers have applied social media data to investigate individuals’ perceptions, thoughts, concerns, behaviors, and mental health [34–37]. For example, Johnson et al. [38] presented a cluster analysis of the contention surrounding vaccines against SARS-CoV-2. Their results revealed that anti-vaccination users on Twitter became highly entangled with undecided users in the main online network, whereas pro-vaccination users are more peripheral in the network. Li et al. [39] applied machine learning techniques to detect stress symptoms-related tweets and manifested that people’s stress symptoms expressed on Twitter have a strong correlation with increasing cases. They also discovered that the main stressors switched from concerns on the increase of reported deaths to financial burdens and economic downturns. Zhao et al. [40] explored COVID-19 related topics on Sina Microblog in China from December 2019 to February 2020. They observed that the public attention initially focused on protection and first-aid and switched to medical services. Zhong et al. [36] examined the association between social media usage and individuals’ mental health. Their results revealed that social media usage rewarded people who gained informational and emotional support as the virus struck but resulted in mental health issues when they were excessively used. Another study conducted by Hou et al. [41] inspected message data on Weibo, Baidu search engine, and Ali e-commerce and described how people’s awareness and behavioral responses corresponded to the media news and government’s announcement.

2.3. Methods in indicating the risk of human interactions

For a highly contagious virus like COVID-19, a rapid appraisal of risky behaviors (i.e., close human interactions) is critical to contain the disease. To the extent of our knowledge, there are two common approaches to reflect the risk of human interactions: questionnaire surveys and mobility-data-based systems.

Surveys have been used to collect the information by asking to what extent a respondent complies with social distancing and the related behaviors during the pandemic [42–44]. Typical questions often include how they interact with others (e.g., virtual or face to face), the frequency and the purposes of going out, and their experiences of being exposed to or tested for COVID-19 [42–44]. For example, researchers at Stanford University performed an online survey and asked questions about the respondent’s health status and their actions [42]. The results showed that the majority of the respondents avoided social gatherings to reduce the risk of infections. Czeisler et al. [43] attempted to assess public compliance with stay-at-home orders through the investigation of more than 5,000 respondents, and more than 80% of the respondents followed the stay-at-home orders and stayed quarantined. The WalletHub’s survey included questions, such as how the respondents feel about social distancing, how they cope during social distancing, what they wear out in public, how often they go outside, and how they feel about others not practicing social distancing [44]. Their key findings showed that most respondents could keep social distancing and 60% of them wore face masks.
The other popular method addresses this issue by establishing interactive monitoring systems based on spatial-temporal mobile data. The Maryland Transportation Institute developed a platform to visualize and supervise parameters reflecting the pandemic’s impacts on mobility, health, and economy [7]. One of the most important daily updated parameters is the social distancing index, which implies the close-distance interactions. Similar to this social distancing index system, Unacast Company published a social distancing scoreboard to show the risk of close human interactions [8]. Its criteria include the changes in average distance traveled, visits to non-essential venues, and human encounters [45]. These social distancing scores range from A to F (A implies the lowest risk, while F implies the highest risk) to assess each state’s risk level of following the social distancing policies [46].

### 2.4. Research objectives

Nevertheless, with the reviewed studies, there has been a lack of studies that utilized social media data to evaluate the risk of people’s behaviors in the pandemic. While survey-based analysis and mobile data-based systems are useful to assess the risk of human interactions, accessing these data may require a vast amount of time and cost. While assessments based on social media may not be as precise as the survey results or mobile data-based instruments, they have advantages of rapidity, spatial coverage, and cost-effectiveness.

Therefore, this study aims to introduce a social media-based approach through the investigation of lockdown-related Twitter postings and quantifies the anti-lockdown ratio to reflect the risk of human interactions in each U.S. state. The rest of this study is structured as follows. Section 3 presents the data preparation and the development of the text classification pipelines. Section 3 also defines the pro/anti-lockdown ratio and introduces the social distancing index. Section 4.1 exhibits the pro/anti-lockdown ratio from temporal and spatial dimensions. Section 4.2 checks its correlation with the social distancing index and demonstrates the credibility of using social media data to evaluate the risk of human interactions. Section 5 discusses the findings, significance, implications, and limitations.

### 3. Methods

#### 3.1. Data preparation

This study used Twitter Search API with the key search terms “lockdown” and “reopen” to download related tweets from April 21 to July 21, 2020. This scraping process generated 30,851,895 lockdown-related and 10,082,646 reopen-related records. Each record contains a user’s profile information (e.g., username, description, registration location, friends, and followers), posted time, tweet content, and retweet information. Tweets used for the analysis include both original tweets and retweets. As retweets are of the same content as their original tweet, it was considered that a user who retweeted a tweet held the same perception as the user who originally posted it. This study selected these two opposite keywords, “lockdown” and “reopen,” to collect data given that they indicated a change of lockdown status, and both topics raised an extensive discussion in Twitter. Moreover, using this pair of opposite keywords could generate enough data and reduce data bias. While there were other terms to describe people’s opinions regarding the lockdown, such as “stay at home,” “shut down,” “open up,” these phrases were not used as they might not be specific to the lockdown policy and could bring in much noise to the dataset.

Due to some system issues, the downloaded “reopen” dataset missed three days’ data files from June 3 to June 5, 2021. This study referred to an external dataset [47] (it used more than 90 keywords and hashtags to download COVID-19 pandemic related tweets) to deal with the missing data. All the records were stored in the JavaScript Object Notation (.json) format and converted to Excel (.xlsx) files. The records without the registration location information or not of the U.S. were removed first, and the tweets not containing the keywords “lockdown” or “reopen” were removed. These two steps reduced the size of lockdown dataset to 3,337,435 records and the size of reopen dataset to 3,480,777 records. As noted, some records could contain both “lockdown” and “reopen” words in a tweet, and thus the duplicated records were removed. As a result, the final dataset contained 6,774,678 records.

To build the training dataset for text classification, this study selected the top 5,000 most frequently occurring unique tweets from the lockdown dataset and the reopen dataset respectively and classified each of them into 1) class 1: the attitude opposing lockdown or supporting reopen, 2) class 0: the attitude is not clear from the tweet, and 3) class −1: the attitude supporting lockdown or opposing reopen. Labeling these tweets followed a polling process. Two authors first labeled each of the selected 10,000 tweets separately. If a tweet received the same label from both authors, the label was used for the tweet. Otherwise, the third author labeled the tweet, and the final label followed the majority of the three opinions. Table 1 presents some typical examples of class 1 and class −1 tweets.

In the 5,000 selected lockdown-related samples, there were 2,266,2,375 and 359 tweets labeled as class 1, class 0, and class −1, respectively. In the 5,000 selected reopen-related samples, there were 1,128,2,162, and 1,710 tweets labeled as class 1, class 0, and class −1, respectively. To check the generalization ability of the trained model, this study randomly selected another 1,000 different and unique tweets from each dataset (no overlaps with the training dataset) and labeled them based on the same polling process. Among these 2,000 testing samples, there were 471, 1,184, and 345 samples labeled as class 1, class 0, and class −1, respectively. The research framework to build and implement the pipeline for text classification is presented in Fig. 1.

Before processing the textual data, this study applied several steps to clean the text, as presented in the “Text Cleaning” box (Fig. 1). Short URLs,
@username, RT @username, digits, emojis, and punctuations in a tweet were removed, and then the stop-words which were not informative, such as “the,” “is,” “and,” were stripped from the text. Next, each tweet was tokenized into a list of separate words and characters. As words in a tweet were written in different tenses or forms, the tokenized words were lemmatized to their stemming forms. This cleaning process was completed with the aid of the Natural Language Toolkit (NLTK) python library [48].

3.2. Text augmentation

It was noted that the percentages of class 1, class 0, and class -1 were not balanced in the training dataset. For example, the distribution in the lockdown training dataset was class 1 (2,262), class 0 (2,375), and class -1 (359). Such imbalance in the training dataset could result in a reduced prediction performance for the minority class. In other words, the trained model could achieve 92.7% accuracy ((2,262 + 2,375)/5,000) by predicting only class 1 and class 0 correctly, and thus it lacked the competence to identify class -1 tweets. This study utilized a text augmentation technique called Easy Data Augmentation (EDA) [49] to balance the training samples. The EDA technique does not require the NLP model to be pre-trained on any external dataset and has been demonstrated to enhance the performance for a small dataset effectively. The EDA model consists of four operations: 1) synonym replacement, 2) random insertion, 3) random swap, and 4) random deletion to increase the dataset volume [49], as presented in Table 2.

This augmentation process enlarged the minor class –1 in lockdown training samples from 359 to 2,154, and the minor class 1 in reopen training samples from 1,128 to 2,256. As a result, there were 2,262 class 1, 2,375 class 0, and 2,154 class –1 “lockdown” training samples, and 2,256 class 1, 2,162 class 0, and 1,710 class –1 “reopen” training samples. In the final training set, there were 4,581 class 1 samples, 4,537 class 0 samples, and 3,864 class –1 samples. Following the recommendations by the original article, this study set the parameters for each of the four operations as $\alpha = 0.1$, where $\alpha$ is a parameter that indicates the percent of the words in a sentence are changed [49]. Since tweets can vary in length, a longer tweet can absorb more noise while...
maintaining the original content than a shorter tweet. To compensate for this issue, this study set the number of words in a tweet to be changed as \( n = d \), where \( l \) is the length of a tweet [49]. For a tweet with fewer words, this setting ensured at least one word to be changed.

### 3.3. Text vectorization

Term Frequency-Inverse Document Frequency (TF-IDF) was used to vectorize tweets into features. TF-IDF is a term weighting method implemented in text similarity, text classification, and information retrieval. Although TF-IDF cannot capture semantic structure in a text, it is a useful algorithm to deal with a large set of texts due to its simplicity and fast computation retrieval [50]. In TF-IDF, TF measures the number of words and their frequencies on each of the documents, while IDF is incorporated to reduce the weights of common words in the corpus. The goal of using TF-IDF is to scale down the impact of words that occur more frequently but empirically less informative in a given training dataset retrieval [50]. The mathematical representation of the TF-IDF method is given below:

\[
w(t, d) = tf(t, d) \cdot idf(t, D) = f_{t,d} \log \left( \frac{D}{d_{t}} \right)
\]

where \( w(t, d) \) represents the word \( t \)'s weight in tweet \( d \), \( f_{t,d} \) denotes the frequency of word \( t \) in tweet \( d \), \( D \) is the total number of tweets, and \( d_{t} \) is the number of tweets that word \( t \) appears. Words with high TF-IDF weight generally indicate strong relationships between the tweets in which they appear. Compared to the Bag of Words method that simply counts the frequency of words, TF-IDF contains the information on the more important words and the less important ones as well. Although TF-IDF does not capture text positions or similarities in a document, it is an efficient and simple algorithm for matching words in a query to documents [51]. Due to its simplicity and fast computation, TF-IDF is beneficial when dealing with a large set of Twitter data. In this study, the scikit-learn python library [52] was applied to compute TF-IDF.

Unlike the TF-IDF method, word embedding techniques have the capability to capture the semantic meanings of words. These alternatives are generally applied to help compute text similarity and perform text classification by converting each word into a pre-trained vector of features. However, this study chose not to use word embedding methods given the following reasons. A word embedding method might treat those augmented data as the same semantic samples. For example, the original tweet was “please reopen America,” and the EDA derived a second tweet using synonym replacement, “please open America.” In this simple case, these two samples might be identical vectors of features according to the word embedding method, and thus the EDA technique might not essentially increase the training samples. As a result, the trained model could still generate a reduced prediction performance for the minor class. Moreover, a word embedding is a much more complicated representation of words and carries more hidden information, which might unnecessarily create more “noisy” patterns in this classification task and increase the difficulty to discriminate the meanings between tweet samples.

### 3.4. Multi-class classifiers

After each tweet was converted into a vector of features, six classifiers were selected to perform the multi-class classification: Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), and Neural Network (NN). Models using these classifiers combined with the TF-IDF technique were constructed based on the scikit-learn python library [52].

A DT is a decision support technique that uses a tree-like structure to represent data and make decisions. In a classification task, a DT classifier can classify a sample by using its leaves to denote class labels and branches to represent conjunctions of features that result in those class labels. A RF is an ensemble learning method that is built on a multitude of DTs [53]. Compared to a single DT, the RFs can reduce variance and control overfitting by training on various subsets of the datasets and outputting the class that is the average prediction of individual trees. The NB classifiers are a family of probabilistic classifiers built based on the Bayes theorem in statistics [54]. In this study, the Multinomial NB (MNB) was applied to perform the multi-class classification task. MNB implements the classification algorithms for data that are multinomially distributed and is suitable for classification with discrete features, such as words frequencies [55]. SVMs are supervised learning models pervasively applied in classification tasks. An SVM intends to apply the kernel functions to map the data in the lower dimensional space into a higher dimensional feature space where it can establish a hyperplane or a set of hyperplanes to search for the maximum margin to split the dataset [56]. LR is a machine learning technique that was originally designed for binary classification. In LR, the core uses a logistic function to model the conditional probability of the predicted label based on independent input variables [57]. In this study’s multi-class case, the training process used a one-vs-rest scheme and applied the multinomial cross-entropy loss. Last, an NN is an interconnected network of neurons that applies mathematical activation functions for information processing. This study applied the Adam solver to update network weights. Adam is an algorithm for the first-order gradient-based optimization of stochastic objective functions [58].

### 3.5. Performance measurement

As the class distributions in the testing set were not balanced, the accuracy itself might not reflect the model’s performance on each class. In addition to testing accuracy, this study used precision, recall, and F1-score to estimate the classification performance. Precision measures the
fraction of true positive cases over the retrieved cases that a model predicts, while 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, representing a combination of recall and precision \[59\]. The mathematical formulas of precision, recall, F1-score, and accuracy are presented from formula 2 to formula 5, in which \( TP = \text{True Positive}, \ TN = \text{True Negative}, \ FP = \text{False Positive}, \) and \( FN = \text{False Negative} \). Each model’s performance on the testing set is exhibited in Table 3.

\[
\begin{align*}
\text{Precision} &= \frac{TP}{TP + FP} \\
\text{Recall} &= \frac{TP}{TP + FN} \\
\text{F1 score} &= \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \\
\text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN}
\end{align*}
\]

According to Table 3, the model trained based on TF-IDF + RF outperformed the other five models, demonstrated by a higher F1-score on every class and a higher testing accuracy. Overall, the F1-score on class -1 is lower than the other two classes, possibly because there are fewer class -1 training samples (some of the training samples were augmented using the EDA technique). It is possible that the variances of textual contents from anti-lockdown data could be smaller. That being said, there could be fewer and less diverse words used to describe anti-lockdown attitudes, such as “Reopen the American,” and “Reopen the States,” but more words used to describe the pro-lockdown possibly because those pro-lockdown tweets often listed specific reasons). Hence, the model presents a better capability to classify an anti-lockdown tweet. Models trained based on DT, LR, and NN show a severe over-fitting problem as the results present a high training accuracy but a low testing accuracy. Therefore, this study selected and applied the TF-IDF + RF model to classify each tweet in the dataset.

Even within the 10,000 labeled samples, the variances inside a class are very high (e.g., tweets in class 0 appear to be very different), which increases the difficulty to train the model. Nevertheless, it is noted that the 10,000 manually labeled tweets constitute 53.3% (3,607,702 out of 6,774,678) of the whole dataset as retweets are of the same textual content of these labeled tweets. Labeling these 53.3% tweet data can obtain the same accuracy as the training accuracy. As a result, the expected accuracy on whole dataset could be 86.5% (97.9% (training) * 53.3% + 73.5% (testing) * 46.7%).

### 3.6. Pro- and anti-lockdown ratio and social distancing index

Social media are often criticized for polarized opinions, which means that there is a possibility that the collected data do not reflect the public opinions \[60\]. Thus, the distribution of opinions on the lockdown orders in various poll results were compared. From the poll data in Table 4, it was discovered that most respondents of the polls were either on the side of supporting or opposing lockdown, and very few people were neutral \[61–63\]. Most of the collected tweets about the lockdown and reopen also showed an opinion either supporting or opposing lockdown, and there were very few tweets showing a neutral or unsure opinion. This indicates that the degree of polarization was not significantly different in the various platforms since people tended to choose a side on the matter of lockdown policy.

In the next step, this study used the classified tweets to quantify the anti-lockdown ratio. The anti-lockdown ratio is calculated as a fraction of the number of “class 1” tweets divided by the sum of “class 1” and “class –1” tweets. The ratio of higher than 0.5 means that there are more anti-lockdown tweets than pro-lockdown tweets. The anti-lockdown ratio can also simply be converted to the pro-lockdown ratio since the sum of the anti- and pro-lockdown ratios is 1. The mathematical formula is presented below.

\[
\begin{align*}
\text{anti lockdown ratio} &= \frac{N(\text{class 1})}{N(\text{class 1}) + N(\text{class } -1)} \\
\text{pro lockdown ratio} &= \frac{N(\text{class } -1)}{N(\text{class 1}) + N(\text{class } -1)}
\end{align*}
\]

The anti-lockdown ratio indicates people’s negative perceptions about the lockdown policy. In order to examine the relationship between the anti-lockdown ratio and the public’s social distancing behaviors, this study obtained the social distancing index data from the COVID-19 Impact Analysis Platform (https://data.covid.umd.edu/) published by Maryland Transportation Institute \[7\]. According to its documentation, the social distancing index implies the chance of close-distance interactions. The index is computed based on Equation (8) using six mobility metrics: percentage of residents staying at home, the percentage reduction of all trips compared to the pre-COVID-19 benchmark, the percentage reduction of work trips, percentage reduction of non-work trips, percentage reduction of travel distance, and percentage reduction of out-of-country trips. The weight for each aspect was chosen based on the share of residents and visitor trips. The index ranges from zero to a hundred, and a higher number indicates that more residents stay at home while fewer visitors are entering the county or state. A smaller social distancing index implies a higher chance that people stay closer \[7\].
4. Results

4.1. Temporal and spatial analysis

For the temporal analysis, this study binned the data based on the number of classifications and then calculated the daily pro- and anti-lockdown ratios. Fig. 2a presents the trends of the national anti-lockdown ratio and average social distancing index. Fig. 2b exhibits the number of states under the stay-at-home order [64]. Fig. 2c illustrates the relationship between the pro-lockdown ratio and the daily new positive cases [65]. Fig. 2d shows the daily tweet volume.

The average anti-lockdown ratio is 0.596, about 10–15% higher than the polling results (Table 4). A couple of reasons may explain such observation. First, it was noted that the trained model generated a lower F1-score for pro-lockdown than anti-lockdown (Table 3), possibly resulted from fewer pro-lockdown training samples and more diverse words used to describe the pro-lockdown attitude. As a consequence, the trained model has limited capability to identify pro-lockdown tweets. Second, while this study compensated the missing reopen data for June 3 to June 5 by referring to an external dataset, this dataset did not use the keyword “reopen” to download related tweets that might result in a partial loss of pro-lockdown data. Third, the ratio was computed based on the number of tweets rather than the number of accounts, and thus it was possible that accounts with the anti-lockdown attitude published more tweets than accounts with the pro-lockdown attitude. While social
Fig. 4. a. The correlation between the anti-lockdown ratio and social distancing index. b. The correlation between the pro-lockdown ratio and social distancing index.

Fig. 5. Correlation between weekly average social distancing and anti-lockdown ratio.
COVID-19 outbreak that hit the U.S. around June 15 led to the resumption of stay-at-home orders from mid-May (Fig. 2a), the discussion volume on Twitter during April and May (Fig. 2b) is that the social distancing index showed a periodic weekly pattern with positive correlations between April and May. The major opinions in support of pro-lockdown orders argued that 1) the lockdown order was a blessing that helped to contain the spreads of infections, 2) the country could not afford to risk a premature reopening since the curve was not flattening, 3) there lacked a robust testing and tracking program to guarantee a safe reopening, and 4) reopening campaigns used bot accounts to disseminate the information support reopening the country. Conversely, the mainstream views in supportive of anti-lockdown orders defended that 1) the lockdown policy was not necessary to contain the pandemic, 2) the lockdown was an overreaction and a violation of constitutional rights, 3) an extended lockdown could cause irreversible damage to the economy and the society (e.g., bankruptcies, depressions, and suicides), and 4) it was a political play.

Another interesting observation for Fig. 2a is that the anti-lockdown ratio continued to increase in early June. After manual examination of the popular anti-lockdown tweets, it was found that this increase was associated with Black Lives Matter (BLM) movements that occurred in early June. People on Twitter argued that it was unfair to allow BLM protests but forbid other forms of protests and outside gatherings. For example, a tweet that was retweeted frequently stated that “we were forced to stay home, close businesses, stop our kids education, let loved ones die alone and Antifa/BLM can riot and ‘protest’ by the 1000’s and it’s okay.” These people criticized the government for its unequal treatment of BLM protests and anti-lockdown protests. Such voice cast doubt on the necessity of lockdown orders on social media and shook the public trust on the government’s measures.

In addition to the temporal analysis, this study conducted a spatial analysis to compare the anti-lockdown ratio across the states, as exhibited in Fig. 3. The data were binned based on the user’s registration location, and then the anti-lockdown ratio for each state was computed. For example, the following registration locations, “California, USA,” “Los Angeles,” “California,” and “Santa Monica CA,” all indicate California State. Further exploration regarding the state-level anti-lockdown ratio and other demographic factors is presented in the appendix.

The mean anti-lockdown ratio (April 21 – July 21) ranged from the lowest of 50.9% to the highest of 70.3% for the states across the United States. The ten states with the highest anti-lockdown ratio are Alabama (AL, 70.3%), Mississippi (MS, 70.2%), Idaho (ID, 69.8%), Tennessee (TN, 69.6%), North Dakota (ND, 69.1%), West Virginia (WV, 0.68%), Nebraska (NE, 67.4%), Oklahoma (OK, 67.1%), Montana (MT, 66.9%), and South Dakota (SD, 66.3%). The ten states with the lowest anti-lockdown ratio or the highest pro-lockdown ratio include Vermont (VT, 50.9%), Maryland (MD, 53.8%), Oregon (OR, 54.6%), Washington (WA, 54.9%), Massachusetts (MA, 55.8%), District of Columbia (DC, 55.8%), New York (NY, 57.0%), Connecticut (CT, 57.3%), California (CA, 57.4%), and Hawaii (HI, 58.0%). Overall, the states in the North West, Middle West and South East present a higher anti-lockdown ratio, while the states in the West Coast and North East Coast report a lower anti-lockdown ratio.

### 4.2. Correlation analysis

To demonstrate that the pro/anti-lockdown ratio can reflect people’s traveling behaviors and help evaluate the risk of human interactions, this study performed a correlation analysis between the social distancing index and the pro- and anti-lockdown ratios. As previously explained (Section 3.6), the social distancing index is built on the mobile location data and shows people’s traveling behaviors in the COVID-19 pandemic. The correlation coefficient is computed using the mathematical formula below.

$$R_{xy} = \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}$$

(10)

where n equals to 51 in this study, indicating the 50 U.S. states plus the District of Columbia, X denotes the data series of the pro- and anti-lockdown ratios of each state, and Y denotes the social distancing index of each state within a specific time period. The correlation plot is presented in Fig. 4, accounting for all the related tweet data between April 21 and July 21. Retweets were not removed since people who...
retweeted a tweet could hold the same opinion of the original tweet. A tweet with more retweets indicates that more people recognized and shared the viewpoint.

The correlation coefficient between the anti-lockdown ratio and the social distancing index is 0.55. Since the sum of the pro-lockdown ratio and anti-lockdown ratio equals 1, the correlation coefficient between the pro-lockdown ratio and the social distancing index is 0.55. The pro/anti-lockdown ratio shows a moderate and positive/negative correlation with the social distancing index for the study period. In the following analysis, this study only shows the relation between the anti-lockdown ratio and the social distancing index for the avoidance of repetition.

This study further grouped the anti-lockdown tweets on a weekly basis and computed the correlation coefficient for each week. The results are exhibited in Fig. 5 and Table 5. The weekly state-level anti-lockdown ratio shows a moderate and negative correlation with the social distancing index except for three weeks (June 3 ~ June 9, June 17 ~ June 23, and July 15 ~ July 21, as shown in Table 5).

A p-value of less than 0.05 is statistically significant, implying that the weekly anti-lockdown ratio is likely to be associated with the changes of the social distancing index. For most weeks in the study period, the anti-lockdown ratio presents a moderate and negative correlation with the social distancing index. When the social distancing index becomes smaller, the expressed anti-lockdown ratio on Twitter from the state tends to be higher. In other words, people from those states with a higher anti-lockdown ratio were more likely to travel outside during and after the lockdown period. This conclusion demonstrates the credibility of using the anti-lockdown ratio as an indicator to assess the risk of human interactions during the COVID-19 lockdown. The states with a higher anti-lockdown ratio tend to have a higher risk of close human interactions when compared to those with a lower anti-
lockdown ratio.

According to Table 5, the correlation coefficient appears to be smaller and more fluctuated from June 3 to July 21 (most states lifted the lockdown in early June, Fig. 2b) when compared to the coefficients from April 22 to June 2 (many states were still under stay-at-home orders, Fig. 2b). This observation further reveals that the anti-lockdown ratio is presumably more aligned with the social distancing index under the lockdown orders but less associated when the lockdown orders were lifted.

5. Discussion and conclusion

First responders, government entities, and private sectors need information to assess and manage the risk of viral transmission during the pandemic. Current survey-based analysis and mobile data-based systems have provided benefits in assessing the risk of human interactions. However, collecting survey responses and mobile data can be costly and time-consuming. Social media data have the advantages of rapidity, quantity, and spatial coverage and can provide useful information to support the risk assessment. To demonstrate the credibility, this study defined the pro- and anti-lockdown ratios based on Twitter data and explored their associations with the social distancing index and other related metrics (e.g., reported infections, number of states under lockdown). Below are the highlighted findings of this study.

- The trend of the anti-lockdown ratio gradually increased from late April to early June. An extension of lockdown orders might trigger negative emotions expressed on social media. Moreover, there was a concomitant increase of the pro-lockdown ratio with increasing reported COVID-19 infections (Pearson $R = 0.59$, $p < 0.05$), implying that the severity of the pandemic could raise people’s risk awareness.
- The increase of the anti-lockdown ratio in early June was associated with Black Lives Matter (BLM) movements. Some people on social media criticized the government for its unequal treatment for BLM protests and anti-lockdown protests. Such voices further cast doubt on the necessity of lockdown orders.
- A negative association between the anti-lockdown ratio and the state’s social distancing index (Pearson $R = -0.55$, $p < 0.05$) was found. This also indicates that there is a positive association between the pro-lockdown ratio and the social distancing index (Pearson $R = 0.55$, $p < 0.05$).
- The study revealed the connection between the opinions expressed on social media and the pattern of people’s behaviors in the pandemic. People from the states with a higher anti-lockdown ratio were more likely to travel outside during and after the lockdown period.

In addition, this study provided several insights on using social media data to understand public opinions. This study presented the framework on how to build the text classification pipelines and discussed how to address sample imbalance with textual data. The framework can be generalized to quantify the level of perceptions expressed on social media towards a policy or an event.

For implications, the pro- and anti-lockdown ratios reflect people’s understandings and compliances with the rules and regulations during the pandemic, and thus can be used as indicators to evaluate the risk of human interactions in the U.S. states. This social media-based approach is practical and operable since it has the advantages of instantaneity, cost-efficiency, and spatial coverage. The model developed in this study can supplement current mobile location-based monitoring systems and provide government agencies, health officials, and the residents an additional instrument to assess such risk.

Despite the aforementioned benefits, there are some limitations of the study. First, this approach may not be applicable for long-term use if the selected topics are not discussed in social media. For example, if the discussion on lockdown is discontinued and active, few tweets of the subject may sway the results. Second, the anti-lockdown ratio is only used to indicate the risk across U.S. states. In other words, a higher anti-lockdown ratio observed in a specific day does not necessarily imply a higher risk of human interactions in that day since the attitude expressed on social media can be primarily affected by related news and policies. For example, BLM movements contributed to an increase of anti-lockdown ratio in early June. Third, it is noticed that the anti-lockdown ratio computed in this study is higher than the polling results. It possibly results from the limited capability of the training model and the inherently imperfect quality of the social media data. Ongoing and further work will pay attention to the text classification models by preparing more training samples, especially for pro-lockdown tweets, and deploying models with more advanced learning techniques (e.g., recurrent neural network). The other research topic will focus on the improvement of the data cleaning process, such as excluding the bot accounts, to make the results represent the real-world opinions.

Author contributions

L. L. contributed to conceptualization, Data curation, Formal analysis, Methodology, Project administration, Visualization, Writing – original draft, and writing – review & editing, revising manuscript. Z. M. contributed to conceptualization, Data curation, Resources, Writing – original draft, writing – review, revising manuscript. H. L. contributed to writing – review & editing, revising manuscript. S. L. contributed to conceptualization, Investigation, Methodology, Project administration, Resources, Writing – original draft, revising manuscript.

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

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors acknowledge the help of Maryland Transportation Institute at University of Maryland to provide the social distancing index data.

Appendix

This study extended the spatial analysis to figure out the associations of state-level anti-lockdown ratio with geodemographic factors. The selected geodemographic factors include educational level (bachelor’s degree %) [68], health (health value 2018) [69], unemployment change (unemployment change from May 2019 to May 2020) [70], household income (average household income 2018) [71], party affiliation (net democratic) [72], economic (GDP per capita 2019) [73], ethnic group (non-white percentage 2018) [74], age (median age 2018) [75], and gender (male to female ratio 2018) [76]. The correlation results are exhibited in Fig. 6.

According to Fig. 6, the state-level anti-lockdown ratio displays a moderate to strong negative correlation with health value ($R = -0.597$, $p < 0.001$), bachelor degree ($R = -0.676$, $p < 0.001$), net democratic ($R = -0.718$, $p < 0.001$), GDP per capita ($R = -0.585$, $p < 0.001$), and average household income ($R = -0.600$, $p < 0.001$). People in the states with higher health value, higher educational level, higher GDP per capita, higher average household income, and more democratic inclined were more likely to hold a perception for supporting lockdown policy. Moreover, people were likely to oppose lockdown policy when the state saw a lower unemployment rate since the state-level ratio shows a moderate correlation with unemployment change ($R = 0.463$, $p < 0.001$). Among these investigated factors, the anti-lockdown ratio doesn’t show a significant correlation with median age, male to female ratio, or the population density. The authors declare that they have no potential competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.
interesting observation for Fig. 6 is that the state-level perception shows U.S. states, moderate correlation with the factors related to the perception of COVID-19 mortality risk perceptions and prosocial behavior, J. Econ. Behav. Organ. (Jan. 2021) doi.org/10.1016/j.jebo.2021.01.007.

25. L. Bode, Feeling the pressure: attitudes about volunteering and their effect on civic and political behaviors, J. Adolesc. 57 (Jun. 2021) 23-30, doi.org/10.1016/j.adolescence.2021.01.004.

26. J. Kim, B.-J. Kim, N. Kim, Perception-based analytical technique of evacuation behavior under radiological emergency: an illustration of the Kori area, Nucl. Technol. 167 (Aug. 2020) doi.org/10.1016/j.nucste.2020.08.012.

27. I. Cojourecenco, G. Cornelissen, N. Karelia, Yes, I can: feeling connected to others increases perceived effectiveness and socially responsible behavior, J. Environ. Psychol. 48 (Dec. 2016) 75-80, doi.org/10.1016/j.jenvp.2016.09.002.

28. M. Abel, T. Byker, J. Carpenter, “Socially optimal mistakes? Debiasing COVID-19 mortality risk perceptions and prosocial behavior”, J. Econ. Behav. Organ. (Jan. 2021) doi.org/10.1016/j.jebo.2021.01.007.

29. Y. Zhao, S. Cheng, X. Yu, H. Xu, Chinese public’s attention to the COVID-19 epidemic on social media: demonstration on flu and cancer, J. Med. Internet Res. 22 (5) (May 2020) e18825, doi.org/10.2196/18825.

30. M. Abel, T. Byker, J. Carpenter, “Socially optimal mistakes? Debiasing COVID-19 mortality risk perceptions and prosocial behavior”, J. Econ. Behav. Organ. (Jan. 2021) doi.org/10.1016/j.jebo.2021.01.007.

31. L. Bode, Feeling the pressure: attitudes about volunteering and their effect on civic and political behaviors, J. Adolesc. 57 (Jun. 2021) 23-30, doi.org/10.1016/j.adolescence.2021.01.004.

32. J. Kim, B.-J. Kim, N. Kim, Perception-based analytical technique of evacuation behavior under radiological emergency: an illustration of the Kori area, Nucl. Technol. 167 (Aug. 2020) doi.org/10.1016/j.nucste.2020.08.012.

33. I. Cojourecenco, G. Cornelissen, N. Karelia, Yes, I can: feeling connected to others increases perceived effectiveness and socially responsible behavior, J. Environ. Psychol. 48 (Dec. 2016) 75-80, doi.org/10.1016/j.jenvp.2016.09.002.

34. M. Abel, T. Byker, J. Carpenter, “Socially optimal mistakes? Debiasing COVID-19 mortality risk perceptions and prosocial behavior”, J. Econ. Behav. Organ. (Jan. 2021) doi.org/10.1016/j.jebo.2021.01.007.

35. L. Bode, Feeling the pressure: attitudes about volunteering and their effect on civic and political behaviors, J. Adolesc. 57 (Jun. 2021) 23-30, doi.org/10.1016/j.adolescence.2021.01.004.

36. J. Kim, B.-J. Kim, N. Kim, Perception-based analytical technique of evacuation behavior under radiological emergency: an illustration of the Kori area, Nucl. Technol. 167 (Aug. 2020) doi.org/10.1016/j.nucste.2020.08.012.

37. I. Cojourecenco, G. Cornelissen, N. Karelia, Yes, I can: feeling connected to others increases perceived effectiveness and socially responsible behavior, J. Environ. Psychol. 48 (Dec. 2016) 75-80, doi.org/10.1016/j.jenvp.2016.09.002.

38. M. Abel, T. Byker, J. Carpenter, “Socially optimal mistakes? Debiasing COVID-19 mortality risk perceptions and prosocial behavior”, J. Econ. Behav. Organ. (Jan. 2021) doi.org/10.1016/j.jebo.2021.01.007.

39. L. Bode, Feeling the pressure: attitudes about volunteering and their effect on civic and political behaviors, J. Adolesc. 57 (Jun. 2021) 23-30, doi.org/10.1016/j.adolescence.2021.01.004.

40. J. Kim, B.-J. Kim, N. Kim, Perception-based analytical technique of evacuation behavior under radiological emergency: an illustration of the Kori area, Nucl. Technol. 167 (Aug. 2020) doi.org/10.1016/j.nucste.2020.08.012.

41. I. Cojourecenco, G. Cornelissen, N. Karelia, Yes, I can: feeling connected to others increases perceived effectiveness and socially responsible behavior, J. Environ. Psychol. 48 (Dec. 2016) 75-80, doi.org/10.1016/j.jenvp.2016.09.002.
[52] F. Pedregosa, et al., Scikit-learn: machine learning in Python, Mach. Learn. PYTHON (2011) 2825-2830.

[53] Tin Kam Ho, The random subspace method for constructing decision forests, IEEE Trans. Pattern Anal. Mach. Intell. 20 (8) (Aug. 1998) 832-844, https://doi.org/10.1109/34.709601.

[54] I. Rish, An Empirical Study of the Naive Bayes Classifier, 2003, p. 6.

[55] C.D. Manning, P. Raghavan, H. Schütze, Introduction to Information Retrieval, Cambridge University Press, New York, 2008.

[56] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (Sep. 1995) 273-297, https://doi.org/10.1007/BF00994018.

[57] D.G. Kleinbaum, M. Klein, Logistic Regression, Springer New York, New York, NY, 2010.

[58] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, ArXiv14126980 Cs (Jan. 2017) [Online]. Available, http://arxiv.org/abs/1412.6980. (Accessed 24 January 2021).

[59] J. Lever, M. Krzywinski, N. Altman, Classification evaluation, Nat. Methods 13 (8) (Aug. 2016) 603-604, https://doi.org/10.1038/nmeth.3945.

[60] C.A. Bail, et al., “Exposure to opposing views on social media can increase political polarization, Proc. Natl. Acad. Sci. Unit. States Am. 115 (37) (Sep. 2018) 9216-9221, https://doi.org/10.1073/pnas.1804840115.

[61] Monmouth University Polling Institute, President inconsistent on covid. https://www.monmouth.edu/polling-institute/reports/monmouthpoll_us_050520/, May 05, 2020.

[62] T. Beaumont, H. Fingerhut, AP-NORC poll: few Americans support easing virus protections, AP NEWS (Apr. 22, 2020). https://apnews.com/article/9ed271ca1304/898522180/dehsple-mass-wars-americans-support-aggressive-measures-to-stop-covid-19-poll-fla. (Accessed 21 January 2021).

[63] B. Mann, Despite Mask Wars, Americans Support Aggressive Measures to Stop COVID-19, Poll Finds, NPR.org, Apr. 04, 2020. https://www.npr.org/2020/08/04/898522180/dehsple-mass-wars-americans-support-aggressive-measures-to-st op-covid-19-poll-fla. (Accessed 21 January 2021).

[64] Coronavirus closures: map of where U.S. states are tightening restrictions. https://www.usatoday.com/storytelling/coronavirus-reopening-america-map/. (Accessed 21 January 2021).

[65] Data Download, The COVID tracking Project. https://covidtracking.com/data/download, (Accessed 21 January 2021).

[66] Nearly Half of the Twitter Accounts Discussing ‘Reopening America’ May Be Bots, Carnegie mellon school of computer science. https://www.cs.cmu.edu/news/early-half-twitter-accounts-discussing-reopening-america-may-be-bots, May 20, 2020. (Accessed 24 January 2021).

[67] A.B.C. News, Reopening the Country Seen as Greater Risk Among Most Americans: POLL, ABC News, May 08, 2020, https://abcnews.go.com/Politics/reopening-coun try-greater-risk-americans-poll/story?id=70055060. (Accessed 23 January 2021).

[68] List of U.S. States and Territories by Educational Attainment, Wikipedia, Jun. 29, 2020 [Online]. Available, https://en.wikipedia.org/w/index.php?title=List_of_U.S._states_and_territories_by_educational_attainment&oldid=965165403. (Accessed 30 June 2020).

[69] Findings state rankings | 2018 annual report, America’s Health Rankings, https://www.americashealthrankings.org/learn/reports/2018-annual-report/findings-state-rankings. (Accessed 30 June 2020).

[70] State Employment and Unemployment Summary. https://www.bls.gov/news.release/lau.nr0.htm. (Accessed 30 June 2020).

[71] List of U.S. States and Territories by Income, Wikipedia, Mar. 29, 2020 [Online]. Available, https://en.wikipedia.org/w/index.php?title=List_of_U.S._states_and_t erritories_by_income&oldid=947939623. (Accessed 30 June 2020).

[72] G. Inc, Democratic states exceed republican states by four in 2018, Gallup.com (Feb. 22, 2019). https://news.gallup.com/poll/247025/democratic-states-exceed-republican-states-four-2018.aspx. (Accessed 30 June 2020).

[73] List of states and territories of the United States by Gdp - Wikipedia. https://en.wikipedia.org/wiki/List_of_states_and_territories_of_the_United_States_by_GDP. (Accessed 24 January 2021).

[74] Population Distribution by Race/Ethnicity | Kff. https://www.kff.org/other/state-indicator/distribution-by-raceethnicity/?currentTimeframe=0&sortModel=%7B%22collId%22%3A%22Location%22%2C%22sort%22%3A%22asc%22%7D. (Accessed 10 July 2020).

[75] List of U.S. States and Territories by Median Age, Wikipedia, Jun. 29, 2020 [Online]. Available, https://en.wikipedia.org/w/index.php?title=List_of_U.S._states_and_territories_by_median_age&oldid=965073937. (Accessed 30 June 2020).

[76] U.S. population: male to female ratio, by state 2018 | Statista. https://www.statista.com/statistics/201946/us-population-males-per-100-females-by-state/. (Accessed 10 July 2020).