Social media mining to investigate the impacts of the COVID-19 pandemic

by

Mohammad S. Parsa

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2022

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 2 is taken from a published research article: M. S. Parsa, L. Golab, Academic Integrity in Online Education during the COVID-19 Pandemic: a Social Media Mining Study, Educational Data Mining Conference 2021, 777-781.

Chapter 3 is taken from a published research article: M. S. Parsa, L. Golab, S. Keshav, Climate Action During COVID-19 Recovery and Beyond: A Twitter Text Mining Study, International Conference on Social Computing, Behavioral-Cultural Modeling Prediction and Behavior Representation in Modeling and Simulation 2021.

Abstract

The COVID-19 pandemic created a global crisis with devastating social and economic impacts. Firstly, public health measures for COVID-19, such as social distancing affected how we work and study. Secondly, this crisis caused mobility restrictions and shutdowns that impacted our economy. In this thesis, we aim to obtain a better understanding of how these socio-economic impacts have affected people. We, therefore, choose one problem from each of these two areas of impact for further study.

The social distancing mandates shifted working environments and education online. Due to cheating being more prevalent in online education [40], serious issues may arise during the pandemic when classes and examinations are online. In order to understand these issues and their impacts on college students, we ask: how do college students feel about online cheating?

Fuel consumption and carbon emissions declined due to mobility restrictions and economic shutdowns [90]. As a result, air quality improved. Economic shutdowns, however, impacted countries' ability to fight climate change. We are interested in understanding how people's perspectives have changed due to both the positive and negative effects of the pandemic on climate change. To do so, we ask: What is the public's attitude towards climate action during the COVID-19 recovery and beyond?

We answer these questions by analyzing discussions on Twitter and Reddit social media platforms. These online social media platforms are considered essential tools for reflecting and forecasting public opinion on a wide range of topics. Therefore, we answer our questions by mining text messages that were posted during the COVID-19 crisis. We begin by collecting the necessary posts and comments. We then prepare the documents for text mining by using standard pre-processing techniques. As a result, we are able to construct an understanding of the discussions by using these methods.

While investigating the discussions about academic dishonesty on Reddit, we found more discussions related to cheating in 2020 than in 2019. The topics have expanded from plagiarism in programming assignments to online assessments in general. Topic modelling of the Fall 2020 discussions revealed three concerns raised by students: that cheating inflates grades and forces instructors to increase the difficulty of assessments; that witnessing cheating go unpunished is demotivating; and that academic integrity policies are not always communicated clearly.

Investigating the discussions about the climate change and the pandemic on Twitter revealed that most tweets support climate action and point out lessons learned during the pandemic response that can shape future climate policy, although skeptics continue to

have a presence. Additionally, concerns arise in the context of climate action during the pandemic, such as mitigating the risk of COVID-19 transmission on public transit.

Acknowledgements

I would like to thank my advisor Professor Lukasz Golab for his continued guidance through this project.

Dedication

This is dedicated to the voiceless minorities and to minorities whose voices can only be heard through social media.

Table of Contents

List of Figures

List of Tables

Chapter 1

1.1 Introduction

In January 2020, the World Health Organization declared the Coronavirus outbreak as a public health emergency. A global crisis with severe social and economic implications resulted from the outbreak. Firstly, public health measures for COVID-19, such as social distancing, impacted how people studied and worked. Secondly, there were mobility restrictions and shutdowns, which impacted the economy and the climate. We, therefore, attempt to gain a better understanding of how these socio-economic impacts have affected people. From each of these two categories of impact, we have chosen one issue to be studied in depth.

1.1.1 Online learning

Throughout the world, public health measures, such as social distancing were enforced to reduce the spread of Coronavirus. Due to these mandates, working environments and education were shifted online. Because classes and examinations were online during the pandemic, cheating and dishonesty prevailed [40]. Online examinations have received a lot of attention, including methods of cheating and ways to mitigate it [29, 100]. But, there has been less emphasis on the impact of online cheating and the associated prevention mechanisms on students' perceptions of and satisfaction with online education, aside from small-scale interviews. To fill this gap, we ask: how do college students feel about online cheating?

Our answer to the above question comes from social media, specifically from Reddit, which is a social curation platform. There are more than 100,000 user-created discussion communities on Reddit referred to as subreddits. A subreddit allows users to create posts that are commented upon by other users. Subreddit names begin with "r/" and correspond to the subreddit topic, e.g., r/politics or r/relationship_advice. As of 2022, there are over 430 million active monthly users on Reddit. Subreddits are named descriptively to make it easy for users to locate discussions about specific topics. Over 80 subreddits corresponding to Canadian and U.S. universities are of interest to our study, which we call *academic subreddits*. In this study, we collected and analyzed academic subreddits created during the Fall 2019 and Fall 2020 semesters that contained at least one keyword associated with cheating, such as 'cheat' or 'misconduct'. Then, we analyze these discussions in two steps. First, collecting data from the same time period in 2019 and 2020 allows us to compare cheating-oriented discussions from before and during the pandemic. For this, we trained a logistic regression classifier based on the words used to distinguish Fall 2019 content from Fall 2020 content. Finally, we analyze Fall 2020 discussions in detail. To do so, we apply the *topic modelling* algorithm, which segments tweets based on the words they use to identify topics of discussion.

When we investigated the discussions about academic dishonesty on Reddit, we found that there were more threads that discussed cheating in 2020 than there were in 2019. These threads shifted from discussions about plagiarism in programming assignments to discussions regarding online assessments in general. Topic modelling of the Fall 2020 discussions revealed three concerns raised by students: that cheating inflates grades and forces instructors to increase the difficulty of assessments; that witnessing cheating go unpunished is demotivating; and that academic integrity policies are not always communicated clearly.

1.1.2 Climate action

Controlling the spread of the virus required immediate large-scale action, including shutdowns and mobility restrictions. Although these actions have negatively impacted the economy, the reduction in carbon emissions has had positive effects on the environment, such as an improvement in air quality [13, 18, 101] and an increase in wildlife activity and breeding success [62]. With the global recovery from the pandemic, these positive effects on the environment may disappear. Therefore, it has been suggested that COVID-19 recovery programs include climate action such as investments in sustainable infrastructure and technologies. An important aspect of making policy involves understanding public opinion, because the Coronavirus pandemic will require many sacrifices on both the social and economic fronts. As a result, we ask the following question: What is the public's attitude towards climate action during COVID-19 recovery and beyond?

We refer to online social media platforms, specifically Twitter, to answer the above question. This platform has been identified as a critical tool for reflecting [78] and predicting [7] public opinion on a variety of topics. Therefore, we analyze almost 40,000 tweets posted on Twitter during the first wave of the COVID-19 crisis (January to August 2020) for keywords that pertain both to the pandemic and climate change. The following steps comprise our data analysis methodology. We first apply the *topic modelling* algorithm, as we did with our first research question, to identify topics of discussion. Afterward, we measure the sentiment on these topics.

Investigating the discussions about climate change and the pandemic on Twitter revealed that most tweets support climate action and point out lessons learned during the pandemic response that can shape future climate policy, although skeptics continue to have a presence. Furthermore, we found that climate skeptics have integrated the pandemic and the associated economic crisis into their reasons for suspending climate policies, such as the carbon tax.

1.2 Summary of Methods

To answer our two research questions, we leverage topic modelling, a text mining technique that segments documents based on words. Documents assigned to the same segment use similar words, and thus they are likely to discuss the same topic. In this section, we explain the methodology of the topic modelling used in Chapters 2 and 3. The source code for the methods used in this thesis is available on our project Github.

To prepare the documents (i.e., posts and tweets) for topic modelling, we performed standard text pre-processing, as in related work on social media mining [65, 77]. We converted all letters to lower case, and we removed punctuation symbols, hyperlinks and stopwords (which are words that serve a grammatical purpose but do not convey any semantic meaning, such as “and”, “the”, etc.). We then lemmatized the remaining words. Lemmatization is the process of grouping together the inflected forms of a word. For example, words such as “plays”, “played” and “playing” are all lemmatized to “play”. We then vectorized each document based on the remaining words. That is, for each document, the ith entry of its vector corresponds to the Term Frequency-Inverse Document Frequency (TF-IDF) of the ith word in the set of (remaining) words occurring in the dataset. The TF-IDF score of a given word for a given document was computed by dividing the number of times the word appears in the document (TF) by the logarithm of the fraction of documents that contain at least one occurrence of this word (DF)[1]. TF-IDF is frequently used when

[1]For example, suppose that a given document contains two occurrences of the word “economy”, and

vectorizing text since it considers both the uniqueness of a word in the dataset and the importance of the word to the specific document.

We apply the Non-negative Matrix Factorization (NMF) method for topic modelling [109] to the vectorized documents. NMF clusters the documents into topics and produces a list of representative terms for each topic. NMF provided better topics than other topic modelling algorithms, such as Latent Dirichlet Allocation (LDA). Similarly, we used LDA to model topics, but the topics correlated poorly with samples and the most frequent n-grams.

We compute and report the following information for each topic.

- Each representative term is given a "representativeness" score by the NMF algorithm, and we selected the top-10 highest-scoring representative terms for each topic.
- We extract the ten most frequent word n-grams (for n up to three, i.e., sequences of up to three consecutive words) within the documents assigned to each topic.
- To confirm the nature of the topics, we manually examined 100 most frequently retweeted documents for each topic.

NMF requires the number of topics as input. To select an appropriate number of topics, we run NMF to produce between 5 to 90 topics and compute the coherence [75] of each output (higher is better). Coherence measures the extent to which the top representative terms representing each topic are semantically related.

1.3 Thesis Outline

In the following Chapters, we discuss our approach in answering the aforementioned research questions. First, we begin by investigating academic cheating during the pandemic in Chapter 2. Next, in Chapter 3, we study the climate action discussions during the pandemic. Finally, in Chapter 4, we conclude the thesis with the implications of our findings and directions for future work.

this word occurs at least once in 1000 documents. The TF-IDF score of "economy" for this document is $3/(\log(1000))$.

Chapter 2

How do college students feel about online cheating?

2.1 Introduction

Challenges related to online education during the pandemic have recently been documented [29, 100]. The focus has been on online examinations, including cheating methods and how to mitigate cheating. However, there has been less emphasis on how online cheating and the associated prevention mechanisms have affected students' perceptions of and satisfaction with online education, aside from small-scale interview-driven studies (details in Section 3.2).

To fill this gap, the question we study in this Chapter is: *How do college students feel about online cheating?* To answer this question, we turn to social media, specifically the Reddit social curation platform (reddit.com). Reddit hosts over 100,000 user-created discussion communities refereed to as *subreddits*. Within a subreddit, users create posts that other users comment on. Subreddit names begin with "r/" and correspond to the subreddit topic, e.g., r/politics or r/relationship_advice. As of 2021, there are over 430 million active monthly users on Reddit.

Descriptive subreddit names make it easy to locate discussions about specific topics or discussions initiated by various kinds of users. Of interest to our study are over 80 subreddits corresponding to Canadian and U.S. universities, which we call *academic subreddits*. We collected all posts and comments on academic subreddits created during the Fall 2019 and Fall 2020 semesters (September through December inclusive) that match at least one keyword related to cheating, such as 'cheat' or 'misconduct'. We found 2,524 such posts and comments in Fall 2019 and 7,809 in Fall 2020.

Our analysis consists of two steps. First, collecting data from the same time period in 2019 and 2020 allows us to compare cheating-oriented discussions from before the pandemic, when classes were held in person, and during the pandemic, with most courses delivered online. To do so, we train a logistic regression classifier to distinguish between Fall 2019 and Fall 2020 content based on the words used. Next, we analyze Fall 2020 discussions in detail. We apply the Non-negative Matrix Factorization algorithm [110], which clusters posts and comments based on the words used and allows us to identify common discussion topics.

Anonymous social media platforms have become a go-to source of public opinion on a variety of topics. In particular, academic subreddits have been analyzed in recent work on students' mental health [8, 80]. To the best of our knowledge, our work is the first to study social media discussions about students' reactions to online cheating.

The remainder of this Chapter is organized as follows. Section 3.2 discusses related work; Section 2.3 describes our methodology; Section 2.4 presents the results; and Section 2.5 discusses actionable insights and directions for future work.

2.2 Related Work

Recent studies have reported that academic misconduct has increased during the COVID-19 pandemic [60, 34, 14, 28, 20, 97]. Studies have been done on how students cheat, with a focus on online examinations; examples include group chats, screen sharing, using Web search engines to answer questions, using websites such as Chegg, as well as impersonation and contract cheating [97, 99, 20, 39, 34]. In response, methods have been proposed to detect and mitigate online cheating [98, 87, 73, 71, 58, 39, 60]. Examples include ensuring that examination questions are open-ended and "un-Google-able", and modifying online examination platforms, such as not allowing to return to previously answered questions [99], restricting the ability to right-click [97], and allowing a short amount of time for each question [14, 70, 20, 28].

There is also recent work on the importance of clear communication between students and instructors in online education [43, 36, 85, 3, 9]. Informing students about software used to detect plagiarism [97] and the consequences of cheating [34], as well as establishing clear rules and giving warnings [39, 1], can be effective in maintaining academic integrity.

The goal of our work is to understand how students feel about the above and other issues related to online cheating. The closest works to ours are those in [28] and [30], which interviewed a small set of undergraduate students and educators. The participants identified some positive aspects of online education, but expressed concerns about cheating and the level of difficulty of online assessments. Our social media analysis explores these and other concerns in detail.

2.3 Data and Methods

2.3.1 Data Collection and Pre-Processing

Previous work on students' mental health [8, 80] identified 83 *academic* subreddits corresponding to major U.S. and Canadian universities. We analyze the same subreddits in this Chapter, listed in the first column of Table 2.1 (U.S.) and Table 2.2 (Canadian). We collected all posts and comments on these subreddits from the Fall 2019 semester, when classes and examinations were held in person, and the Fall 2020 semester, when most campuses moved to online delivery (September-December inclusive). We downloaded the data using a publicly-accessible Reddit interface, PushShift[1].

Next, we retain only those posts and comments that contain at least one of the following keywords: 'cheat', 'plagiari', and 'misconduct'. We perform *substring* matching, meaning that the keyword 'plagari' matches the words 'plagiarize' and 'plagiarism'. Table 2.3 shows the number of posts and comments that include at least one occurrence of each of these three keywords in Fall 2019 and Fall 2020 (note that the same post or comment may match one, two or all three keywords). The substring 'cheat' is significantly more common than the other two.

Tables 2.1 and 2.2 report the number of posts ("P") and comments ("C") on each U.S. and Canadian academic subreddit, respectively, in Fall 2019 and Fall 2020. The "Before" numbers correspond to all posts and comments. The "After" numbers correspond to posts and comments that matched at least one cheating-related keyword, i.e., the posts and comments that will be analyzed in the remainder of this Chapter.

Table 2.1: Number of posts and comments on U.S. academic subreddits (C: Comments, P: Posts)

[1]https://pushshift.io

Subreddits	2020				2019			
	Before		After		Before		After	
	C	P	C	P	C	P	C	P
UIUC	39974	6991	160	21	40556	6431	104	10
berkeley	37355	6343	365	69	28537	4637	114	17
Cornell	36235	8139	165	27	22562	3900	45	8
Purdue	34376	6317	148	15	33322	5273	42	11
UCSD	30589	5798	175	34	28214	5364	106	15
rutgers	29861	6622	269	69	44114	8902	122	16
UMD	21937	4225	206	28	25794	4631	97	6
SBU	20521	4301	163	20	28328	5373	63	13
uofm	19954	3174	79	14	13553	2213	44	5
udub	17867	3487	82	17	18187	3187	59	6
UWMadison	14870	2447	103	18	14236	2039	33	3
UTAustin	13620	3112	53	7	13866	2811	90	6
utdallas	12763	2235	74	7	20731	3109	25	5
PennStateUniversity	12345	1944	64	5	9620	1610	42	2
msu	12052	2104	86	10	15066	2329	23	6
NCSU	11653	1794	72	5	18943	2524	32	1
UVA	11627	2424	79	9	5071	1084	19	5
rit	11603	1577	48	2	10768	1643	6	1
nyu	11034	2952	37	7	5731	1438	10	2
UNCCharlotte	10132	1709	93	12	10700	1508	18	1
USC	9551	1958	82	15	6800	1419	17	4
Baruch	9370	2226	94	16	4851	1144	36	12
UPenn	8886	2083	55	10	4212	997	11	1
UNC	8347	1644	30	8	3800	790	6	2
byu	6951	707	39	2	3165	407	25	3
UGA	6637	1520	20	3	6852	1349	2	0
columbia	6496	1573	55	5	4699	708	22	3
RPI	5652	1220	70	0	7622	1343	5	0
uichicago	4880	894	46	4	6606	1009	84	1
SJSU	4661	1068	27	5	5108	1136	18	3
stanford	3944	1223	13	2	3782	882	10	0
bostoncollege	3493	1006	0	0	753	188	0	0
cmu	3388	657	27	2	2764	517	3	0
washu	3159	572	4	0	1134	259	0	0

Vanderbilt	2581	555	9	1	1447	311	0	0
Harvard	2219	634	1	1	2294	517	1	0
UMBC	2036	457	21	3	2479	464	4	0
duke	2020	469	2	1	1397	317	7	2
mit	1758	532	3	0	1651	373	4	0
BrownU	1363	438	2	1	1315	276	0	0
IndianaUniversity	1225	588	1	1	1797	543	9	1
Caltech	494	130	0	0	220	59	0	0
Total	509479	99849	3122	476	482647	85014	1358	171

Table 2.2: Number of posts and comments on Canadian academic subreddits (C: Comments, P: Posts)

Subreddits	2020				2019			
	Before		After		Before		After	
	C	P	C	P	C	P	C	P
uwaterloo	72244	8372	381	58	88996	9888	130	17
UofT	54343	8460	701	86	67649	9375	171	23
UBC	40058	5281	766	42	39416	5039	109	11
uAlberta	33265	7164	341	58	49494	8270	137	23
McMaster	24556	5188	219	45	14932	2638	27	3
mcgill	21380	3376	167	15	20852	3067	58	6
yorku	15671	4065	228	46	22078	3862	47	6
CarletonU	15455	2531	207	11	16874	2706	43	2
Concordia	10065	2394	192	27	10292	2185	27	7
uwo	9717	1856	122	10	11758	1764	35	2
wlu	8097	1788	97	16	5499	1203	13	4
uvic	7291	1178	85	3	4756	828	11	3
ryerson	6503	2282	87	6	14922	2927	37	8
queensuniversity	5234	1107	18	1	4758	824	6	2
umanitoba	4408	861	66	7	3183	717	3	1
uoguelph	3381	794	51	8	3691	693	5	2
Dalhousie	1807	401	21	4	2019	407	6	2
usask	1177	290	0	0	666	178	0	0
brocku	1007	366	2	0	1442	329	4	2
memorialuniversity	785	183	6	1	637	147	2	0
UdeM	422	90	1	0	174	48	0	0

lakeheadu	119	59	2	1	51	21	0	0
uleth	112	35	0	0	82	33	0	0
University_Of_Regina	96	30	1	0	8	11	0	0
AcadiaU	69	29	1	0	60	15	0	0
UQAM	67	22	0	0	48	17	0	0
uwinnipeg	65	24	2	1	15	10	0	0
unb	62	35	0	1	8	12	0	0
laurentian	33	16	0	0	9	4	0	0
stfx	32	12	0	0	0	1	0	0
SMUHalifax	24	17	0	0	21	9	0	0
nipissingu	13	8	0	0	3	4	0	0
UPEI	12	10	0	0	1	3	0	0
stthomas	6	4	0	0	0	3	0	0
BishopUniversity	5	2	0	0	0	4	0	0
UNBC	3	5	0	0	15	10	0	0
mta	1	0	0	0	6	6	0	0
cbu	0	2	0	0	3	1	0	0
MSVU	0	0	0	0	0	1	0	0
uottawa	0	0	0	0	83	43	0	0
usherbrooke	0	0	0	0	0	2	0	0
Total	337585	58337	3764	447	384501	57305	871	124

Finally, we perform standard text pre-processing. Following previous work on Reddit topic modelling [51, 80], we remove posts and comments with fewer than 40 or more than 4000 characters: short ones are unlikely to be meaningful (and may correspond to URLs), while long ones may mention more than one topic. We also remove stopwords and lemmatize the remaining words (i.e., we group together the *inflected* forms of a word) using the Python NLTK parser.

Table 2.3: Number of posts and comments containing the substrings 'cheat', 'plagiari' and 'misconduct'

Term	2020	2019
cheat	7731	2321
plagiari	681	433
misconduct	450	202

2.3.2 Logistic Regression Analysis

We use supervised learning methods to distinguish between cheating-related discussions before and during the pandemic. Because unsupervised learning cannot be guaranteed to create groups that distinguish between the two datasets, we did not use unsupervised learning methods. Moreover, these groupings may contain topics from both before and during the pandemic. Hence, we use supervised learning methods for this task, namely, logistic regression. As a result of training a logistic regression model, we will be able to examine the features (in this case terms) that are most and least correlated with the prediction. Through the study of these terms, we will be able to distinguish between posts or comments written in Fall 2020 and Fall 2019.

We use term frequency–inverse document frequency (TF-IDF) word scores as features in the model. To interpret the model, we identify words with the most positive coefficients (predicting Fall 2020) and the most negative coefficients (predicting Fall 2019). Our model obtained a 10-fold cross-validation accuracy score of 73%, a precision of 76%, a recall of 96% and an F1-score of 86%.

(We also tested logistic regression models with additional features, including word bigrams, the sentiment of the post or comment (computed using the Valence Aware Dictionary and Sentiment Reasoner (VADER) [46]) and linguistic features computed using Linguistic Inquiry and Word Count (LIWC) [83]. After adding these features, accuracy improved by two percent to 75%. However, none of these additional features were assigned large coefficients and therefore are not considered further in the remainder of the Chapter.)

2.3.3 Topic Modelling

Finally, we apply the Non-negative Matrix Factorization (NMF) topic modelling algorithm [110] (explained in Chapter 1) on the Fall 2020 posts and comments that match at least one cheating-related keyword.

Since NMF requires the number of topics as input, we calculated the *coherence* of the topic descriptors for each topic, as shown in Figure 3.1. However, a preliminary analysis of the NMF output at five topics revealed that most of these topics consisted of several discussion themes. This observation suggested that a larger number of topics may be more appropriate, and thus we selected 20 topics, which has the second-highest coherence score.

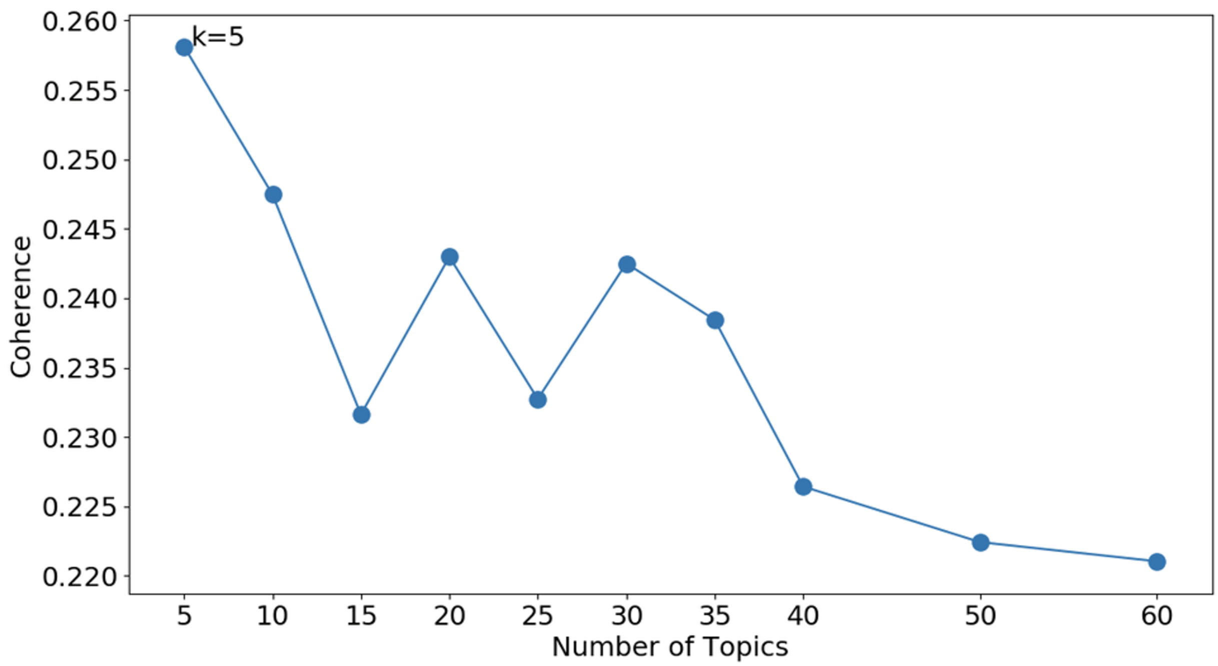

Figure 2.1: Coherence scores for 5-60 topics

2.4 Results

2.4.1 Comparison of 2019 and 2020 Discussions

We begin with Figure 2.2, which plots the weekly number of posts plus comments (related to cheating) in Fall 2019 (blue line) and Fall 2020 (orange line). There are roughly three times as many such posts and comments in 2020 than in 2019 (7,809 vs. 2,524) even though the total number of posts and comments on academic subreddits has not changed much from 2019 to 2020 (see the total "Before" numbers in the last row of Tables 2.1 and 2.2). Note that both lines have spikes in December, corresponding to final examinations.

Next, Table 2.4 lists the most positive and the most negative coefficients from the logistic regression model, with positive coefficients predicting Fall 2020 posts and comments and negative coefficients predicting Fall 2019 posts and comments. The most positive coefficients include 'chegg' (an online platform for answering college and high school questions), as well as words related to online proctoring such as 'proctor', 'proctorio', 'zoom', 'camera', 'webcam' and 'privacy'. The most negative coefficients suggest in-person examinations ('cheat sheet', 'bring', 'sit') and programming assignments and projects ('code',

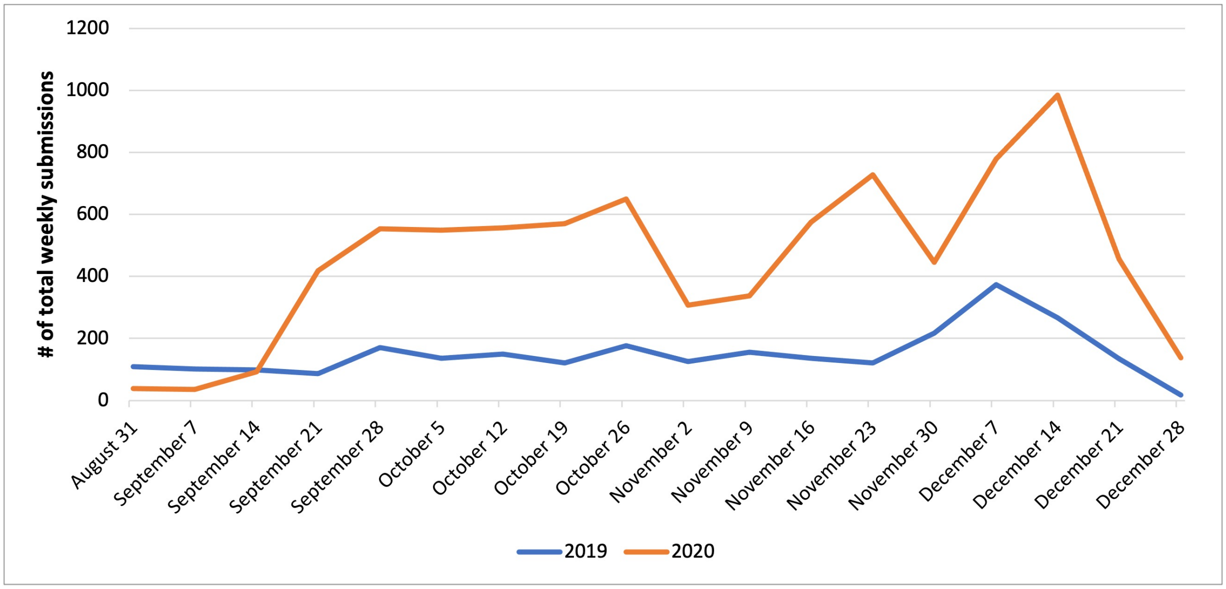

Figure 2.2: Weekly count of posts and comments related to cheating in Fall 2019 and Fall 2020

'program', 'project').

2.4.2 Topic Modelling

Table 2.5 shows the NMF topic descriptors, the frequent n-grams, and the percentage of posts and comments assigned to each topic. We group the topics into the following three categories based on the information in Table 2.5 and manual inspection of a sample of posts and comments.

First, about 40% of the posts and comments include concerns about cheating leading to grade inflation, which in turn leads to assessments becoming more difficult. Students have observed grade inflation (Topic 13) and expressed concerns that Fall 2020 examinations will be more difficult to reduce the class average (Topics 1 and 20). Moreover, students commented on various methods used by instructors to combat cheating and reduce grades, such as grading on a curve (Topics 14 and 15) and using anti-cheating and online proctoring software (Topics 9 and 11). We show three examples below.

"I think online learning is great for the reasons you mentioned, but there are a bunch of problems that need to be fixed. Namely that to counteract cheating some professors are

assigning way too much work and not giving nearly enough time for online exams (according to one of my profs the goal is to make exams so hard that the majority of students don't finish so as to make cheating virtually impossible). I'm pretty good with online learning (i usually only attend less than half of my lectures in a regular year and do very well nonetheless) but I just cant bring myself to care about anything im learning in this 100% online format. ..." (Topic 1)

"I really understand that he is trying to stop cheating from happening but this seems a little way too stressful. Like we already are being zoom proctored and it's open note already there's really nothing we could really do especially with time constraints???" (Topic 9)

"The more students cheat on exams the higher the class average is and the more the class gets curved down." (Topic 14)

Next, students reported feeling demotivated when they know that cheating happens in examinations (Topics 4 and 5) and often goes unpunished (Topics 3, 10 and 17). Students discussed examples of cheating that instructors failed to identify, such as seeking answers on Google and question-answering websites such as Chegg (Topics 7, 8 and 18), and discussing solutions in online chat groups (Topic 19). We give three example comments below.

"Funny, your entire point is that the university doesn't owe students their mental health, yet I don't see why YOU are demotivated about class averages? No one is asking you to compare yourself to others. If you know that others are cheating, then why do you care about the average so much? By not preventing cheating, the university is directly penalizing students with integrity. Right... so you're saying that by not preventing cheating, universities are hurting student's feelings about their integrity. But I thought students are the "only person responsible" for their feelings? Hmmm...." (Topic 4)

"Let us admit, a lot of people do cheat in online exams. I've seen ppl in other universities formed a discord group and live-stream their exams to cheat together, and a lot of dudes using symbolab in the math exam. A lot of prof counter this using very limited exam time, but ppl can always find a way or two to cheat, especially in online exams." (Topic 8)

"I don't recommend you being in a group chat. I've seen several posts here about how people in a course group chat started cheating on tests and got academic dishonesty. It's not a good idea to join a group chat" (Topic 19)

Finally, students reported concerns about new methods used to prevent cheating in online examinations. They worried that some legitimate actions may be misconstrued as cheating: looking away from the computer screen, accidentally pressing a button, or disconnecting from a video meeting due to internet connectivity issues (Topics 6 and 12). Furthermore, some students reported being accused of cheating during online examinations, but did not realize they did anything wrong (Topics 2 and 16). We show three examples below.

"During a test two weeks ago, I accidentally left Quercus for a few seconds because I clicked on Windows update, and also had League of Legends open in the background. Today, our prof mentioned people were cheating during the test, and now I'm worried I'm going to get slapped with a punishment. I'm debating if I should email my prof now to clarify things, but I'm also worried that it'll seem like I actually cheated and was making excuses if I do that. Should I wait and see if they send me an email for academic dishonesty?" (Topic 6)

"just to clarify, does being flagged automatically mean they caught me cheating? Or does it mean they are suspicious? cause the email i got said we are investing you for potential academic misconduct" (Topic 16)

"My professor emailed me saying I been flagged in an active Academic Dishonesty investigation, however I did not cheat nor would I ever. I am worried for I do not want to be accused of cheating if I have not done anything. Anyone have any advice of what to email back?" (Topic 2)

Table 2.5: Fall 2020 topic modelling results

#	Topic descriptors	Frequent N-grams	%
1	work, really, time, way, learn, try, hard, help, school, good	*'feel like', 'work hard', 'first year', 'high school', 'office hour', 'mental health', 'learn material', 'get catch', 'make sure', 'in person'*	10.4
2	say, academic, email, integrity, case, code, worry, report, flag, mean	*'academic integrity', 'academic dishonesty', 'integrity violation', 'academic integrity violation', 'get flag', 'student conduct', 'academic offense', 'would say', 'get catch', 'even though'*	10.3
3	think, probably, pretty, fine, worry, fair, sure, reason, away, good	*'think would', 'think people', 'think get', 'like think', 'get away', 'make sure', 'really think', 'feel like', 'think go', 'think make'*	6.4
4	student, university, honest, case, punish, international, chinese, issue, school, conduct	*'international student', 'student get', 'many student', 'chinese student', 'honest student', 'academic integrity', 'student would', 'mental health', 'academic dishonesty', 'first year'*	5.7

5	know, want, let, wrong, happen, person, tell, need, mean, consequence	*'let know', 'want know', 'know people', 'get catch', 'know would', 'lot people', 'feel like', 'know know', 'know go', 'student know'*	5.5
6	prof, email, mark, ta, ask, tell, send, chance, midterm, try	*'prof make', 'first year', 'email prof', 'open book', 'feel like', 'prof say', 'prof ta', 'prof would', 'make sure', 'ask prof'*	5.4
7	question, answer, time, ask, quiz, look, minute, similar, wrong, google	*'answer question', 'go back', 'multiple choice', 'short answer', 'exam question', 'one question', 'look answer', 'question answer', 'question exam', 'choice question'*	5.1
8	test, open, book, note, close, online, tab, internet, easy, search	*'open book', 'open note', 'make test', 'take test', 'test open', 'close book', 'book exam', 'open book exam', 'exam open', 'book test'*	4.9
9	people, lot, stop, say, agree, mean, proctor, probably, maybe, care	*'people get', 'lot people', 'many people', 'people would', 'get catch', 'people like', 'mental health', 'people go', 'know people', 'feel like'*	4.8
10	like, feel, sound, look, yeah, lol, bad, thing, lot, shit	*'feel like', 'seem like', 'look like', 'sound like', 'something like', 'even though', 'would like', 'make feel', 'online school', 'like people'*	4.8
11	exam, proctor, final, online, open, book, sheet, time, hour, note	*'take exam', 'final exam', 'open book', 'online exam', 'make exam', 'proctor exam', 'write exam', 'take home', 'home exam', 'person exam'*	4.7
12	use, software, proctor, proctorio, computer, browser, note, flag, lockdown, webcam	*'lockdown browser', 'secondary device', 'make sure', 'proctor software', 'take exam', 'get flag', 'student use', 'use respondus', 'virtual machine', 'use note'*	4.5
13	course, year, average, math, midterm, final, assignment, fail, term, quiz	*'first year', 'take course', 'last year', 'math course', 'feel like', 'midterm final', 'year course', 'course average', 'final exam', 'class average'*	4.5
14	class, curve, online, semester, average, fail, homework, lot, easy, problem	*'take class', 'class average', 'online class', 'class get', 'one class', 'feel like', 'math class', 'class take', 'in person', 'make sure'*	4.4

15	grade, curve, average, semester, high, final, letter, higher, better, good	*'good grade', 'letter grade', 'final grade', 'get good', 'get good grade', 'grade get', 'get grade', 'grade inflation', 'grade curve', 'better grade'*	4.2
16	professor, happen, try, evidence, accuse, report, tell, prove, probably, email	*'professor make', 'take exam', 'make exam', 'professor would', 'professor might', 'make sure', 'student professor', 'professor try', 'in person', 'tell professor'*	4
17	catch, happen, wonder, lol, hear, dumb, expel, time, lmao, guy	*'get catch', 'people get', 'people get catch', 'first time', 'catch people', 'catch get', 'use chegg', 'get away', 'without get', 'without get catch'*	3.7
18	chegg, post, account, use, ip, information, address, answer, view, solution	*'use chegg', 'ip address', 'chegg account', 'get catch', 'post chegg', 'question chegg', 'post question', 'chegg exam', 'chegg answer', 'answer chegg*	2.8
19	group, chat, leave, join, share, report, quiz, snitch, post, want	*'group chat', 'share answer', 'get trouble', 'group member', 'join group', 'class group', 'leave group', 'academic integrity', 'group project', 'study group'*	2.5
20	make, harder, sure, sense, hard, easier, difficult, mistake, thing, pretty	*'make sure', 'make harder', 'make exam', 'make sense', 'harder make', 'want make', 'make mistake', 'make difficult', 'make feel', 'want make sure*	1.4

2.5 Discussions

Our logistic regression analysis revealed that cheating-related discussions have expanded from plagiarism in computer programming (representative of Fall 2019) to online assessments in general. The word 'chegg' was associated with Fall 2020 content, suggesting an increase in the use of Chegg and related websites, which is consistent with prior work [34, 20]. Furthermore, words indicating online proctoring were predictive of Fall 2020 content, e.g., 'camera', 'webcam' and 'record'. Inspection of the posts and comments containing these terms revealed students' concerns about their privacy during online examinations. Similar concerns were raised in recent work [28, 48].

Our topic modelling analysis identified three discussion themes in Fall 2020. First, students believe that cheating causes grade inflation, which motivates instructors to make assessments harder and introduce strict anti-cheating protocols such as not being able to scroll back to a previous question on an online examination. Some of these concerns have been highlighted in previous work [97, 99, 14, 28, 70, 20], and our analysis reflects students' opinions on this topic. Second, unpunished cheating lowers students' morale and motivation. Students report feeling demotivated when classmates cheat and obtain high grades. Third, students report not knowing exactly what constitutes cheating and what is allowed. These concerns were often reported in the context of online examinations, with students unsure of how their actions are being monitored.

Table 2.4: Words with the most positive and most negative logistic regression coefficients

Term	coefficient	Term	coefficient
chegg	2.19	sheet	-3
online	1.79	cheat sheet	-2.95
proctor	1.79	code	-1.87
open	1.62	project	-1.68
covid	1.55	plagiarism	-1.51
zoom	1.45	phone	-1.47
prof	1.37	plagiarize	-1.32
pandemic	1.25	relationship	-1.31
proctorio	1.11	sit	-1.1
flag	1.09	talk	-1.02
cheat	1.08	sexual	-0.98
chat	1.06	notice	-0.94
camera	1.03	bring	-0.93
internet	1	textbook	-0.93
privacy	1	international	-0.92
book	1	misconduct	-0.78
cheater	0.98	appeal	-0.78
webcam	0.95	program	-0.79
100	0.93	go	-0.79
format	0.92	front	-0.81
screen	0.9	report	-0.81
open book	0.89	try cheat	-0.81
sem	0.88	ask	-0.81
record	0.88	homework	-0.82
math	0.88	dean	-0.82
term	0.87	practice	-0.83
average	0.86	allow	-0.88
respondus	0.85	partner	-0.88
email	0.83	final	-0.89
semester	0.83	english	-0.9

Chapter 3

What is the public attitude towards climate action during COVID-19 recovery and beyond?

3.1 Introduction

The COVID-19 pandemic has created a global crisis. Controlling the spread of the virus required immediate large-scale action, including shutdowns and mobility restrictions. While these actions have had negative effects on the economy, the corresponding reduction in carbon emissions resulted in some positive impacts on the environment, such as improvements in air quality [13, 18, 101] and an increase in wildlife activity and breeding success [62].

As the world recovers from the pandemic, these positive environmental impacts are at risk of vanishing [41]. It has therefore been suggested that COVID-19 recovery programs should include climate action such as investing in sustainable infrastructure and technologies [22, 88]. Others additionally suggest building on social changes such as working from home [11, 44].

An important aspect of policymaking is an understanding of public opinion, especially now, given the social and economic sacrifices required to combat the Coronavirus pandemic. We therefore ask the following question in this Chapter: *What is the public attitude towards climate action during COVID-19 recovery and beyond?*

Online social media platforms such as Twitter have been identified as critical tools for reflecting [78] and predicting [7] public opinion on a variety of topics. We therefore answer our question by text-mining nearly 40,000 messages posted on Twitter during the first wave of the COVID-19 crisis (January to August 2020) and include keywords related to both the pandemic and climate change. We did not use Reddit for this analysis, since there were only 3,000 posts and comments that contained the keywords.

Our data analysis methodology consists of the following steps. First, we apply a *topic modelling* algorithm that segments the tweets based on the words used in order to identify topics of discussion. Next, we measure the sentiment of opinions expressed on each topic as well as the percentage of tweets classified as inflammatory. Finally, we inspect a sample of tweets belonging to each topic in order to confirm the nature of the topic and the sentiment of the opinions on this topic.

There has been previous social media mining work (mainly using Twitter) on understanding attitudes towards climate science and climate action, as well as understanding public opinion on policies related to mitigating the COVID-19 pandemic. At the intersection of these two topics, there has been qualitative work on the impact of economic shutdowns and social distancing on the environment during the pandemic. To the best of our knowledge, our work is the first data-driven study of public commentary on climate action during the pandemic, as reflected on social media.

The remainder of this Chapter is organized as follows. Section 3.2 discusses related work in more detail; Section 3.3 explains the data and methods used; Section 3.4 discusses the results; and, Section 3.5 concludes with insights and directions for future work.

3.2 Related Work

In this section, we review related work in three areas: social media discussions on climate change, social media discussions on COVID-19, and work done at the intersection of these two topics.

3.2.1 Social Media Discussions on Climate Change

In terms of climate change discussions on Twitter, prior work studied discussion topics [45, 56, 65, 75], spatiotemporal patterns [45, 55, 56, 64, 79], communities of activists and skeptics [53, 105, 54, 108], correlations between Twitter activity, weather and political events [6, 12, 31, 52, 57, 89, 76], political polarization [23, 37, 49, 82, 4], public engagement in climate discussions [32, 82], as well as scientific consensus on climate change [82, 81, 59].

Specifically, there are two studies that applied topic modelling on tweets related to climate change [38, 102]. It was found that majority of the tweets discussed the consequences of climate change for humans. Additionally, Dahal et al. [25] observed that climate change discussions comprise a wide range of topics, such as politics, beliefs, economics, and the environment. While we use similar methods in this Chapter, these topic modelling studies were done before the COVID-19 pandemic.

The sentiment of climate change discussions has also been studied. Williams et al. [105] found that tweets between like-minded users have a positive sentiment whereas tweets between skeptics and activists have a negative sentiment. Another study [4] observed offensive language used in topics associated with politics and climate skeptics. Again, in contrast to our work, these studies were done before the pandemic.

3.2.2 Social Media Discussions on COVID-19

COVID-19 discussions on Twitter and other social media have recently been studied. Areas of focus include frequently asked questions [4, 32], effects of the pandemic on mental well-being [2, 47, 61, 67, 103, 113], impacts of rumors and misinformation related to COVID-19 [5, 27, 19, 33, 84, 93, 96, 47], and health problems caused by economic shutdowns and social distancing [2, 47, 103, 113, 2, 35]. However, climate-related discussions have not been studied in detail, and we fill this gap in this Chapter.

In terms of sentiment analysis, previous work studied the content published on social media during the pandemic. Wang [104] found that the majority of posts published on Weibo that contain discussions about COVID-19, its symptoms, and public health controls have a negative sentiment. Moreover, Sanders et al. [92] found that negative sentiment was mostly associated with political and pro-mask topics, such as discussions related to the former President Trump not wearing a mask. Another study [112], however, showed that most topics published on Twitter in April 2020, such as “stay safe”, have a positive sentiment. In this work, we also perform sentiment analysis, with novel focus on climate action during the pandemic.

Xue et al. [111] investigated COVID-19 discussions and concerns by applying topic modelling and sentiment analysis. The authors reported a feeling of fear associated with topics related to new COVID-19 cases and deaths. The authors did not find any discussions related to the impact of the pandemic on climate action.

Finally, one study [47] investigated social media rumors and conspiracy theories about the Coronavirus in 25 different languages. None of these conspiracy theories and fake news were linked to climate change and the environment.

3.2.3 Climate Change and the Pandemic

Previous work in this area measured air quality in various regions. Improvements in air quality have been reported in China [72, 74], India [94], U.S. [114], Malaysia [68], and Thailand [95]. However, researchers believe that these improvements are temporary and that pollution will increase in the long term [41, 114, 72].

In response to the potential increase in carbon emissions during economic recovery, environmental economists promote climate action as the world recovers from the pandemic [57, 88, 16, 57, 69, 86] and ask for a sustainable economic development [10]. Suggestions include using recovery funds to innovate a low-carbon energy transition, as well as investing in education, training, and green infrastructures [88, 42, 69]. Our work is complementary: we focus on public opinion about these suggestions expressed on social media.

Furthermore, prior work suggests learning from the global response to the COVID-19 pandemic and incorporating these lessons in climate action [63, 86]. Salas [91] points out that prioritising preventive action [63, 15], supporting scientists and experts, and coordinating a rapid global response [63] are the important lessons learned during the pandemic that should be applied to climate action. Again, in contrast, our work focuses on public perception of these suggestions.

3.3 Data and Methods

This study uses data from the Twitter social networking and microblogging platform. Twitter users post messages with up to 280 characters, containing text, images, hyperlinks or hashtags, which are words starting with the symbol '#' and are used to index keywords and topics [24]. Users may follow, i.e., ask to receive tweets from, other users, and may forward, i.e., retweet, messages written by other users. There are currently over 330 million active users on Twitter worldwide, sending approximately 500 million tweets per day [50]).

Recent work has identified over 81 million tweets containing words related to the COVID-19 pandemic, listed in Table 3.1, spanning from January 1, 2020 to July 31, 2020 [18]. This dataset contains only the tweet IDs, and we obtained the full tweets via the public Twitter download interface . Notably, this dataset only contains 'public' tweets and omits those marked as 'private' and therefore only visible to one's followers.

Next, following the methodology used in prior work on climate change discussions on Twitter [21, 49, 53], we identified a subset of these tweets that also contain at least one of the following terms: climate change, #climatechange, global warming or #globalwarming.

Table 3.1: Terms used by Chen et al. [18] to identify tweets related to the COVID-19 pandemic

Terms			
Coronavirus	Koronavirus	Corona	CDC
Wuhancoronavirus	Wuhanlockdown	Ncov	Wuhan
N95	Kungflu	Epidemic	outbreak
Sinophobia	covid-19	corona virus	covid
covid19	sars-cov-2	COVID19	COVD
pandemic	coronapocalypse	canceleverything	Coronials
SocialDistancingNow	Social Distancing	SocialDistancing	panicbuy
panic buy	panicbuying	panic buying	14DayQuarantine
DuringMy14DayQuarantine	panic shop	panic shopping	panicshop
InMyQuarantineSurvivalKit	panic-buy	panic-shop	coronakindness
quarantinelife	chinese virus	chinesevirus	stayhomechallenge
stay home challenge	sflockdown	DontBeASpreader	lockdown
lock down	shelteringinplace	sheltering in place	staysafestayhome
stay safe stay home	trumppandemic	trump pandemic	flattenthecurve
flatten the curve	china virus	chinavirus	quarentinelife
PPEshortage	saferathome	stayathome	stay at home
stay home	stayhome	GetMePPE	covidiot
epitwitter	pandemie	wear a mask	wearamask
kung flu	covididiot	COVID__19	

We then removed non-English tweets, leaving 155,716 tweets, and we removed retweeted copies, finally resulting in 39,461 tweets for analysis. While removing retweeted copies, we labelled each distinct tweet with the number of times it was retweeted.

3.3.1 Topic Modelling

To prepare the tweets for topic modelling, we performed standard text pre-processing, as explained in Chapter 1. We then vectorized each tweet based on the remaining words Using TF-IDF. Then, we applied the Non-negative Matrix Factorization (NMF) method for topic modelling [109] to the processed tweets. As shown in Figure 3.1, coherence was highest at 15 topics. Thus, we perform topic modelling for 15 topics.

Next, following recent work on sentiment analysis of COVID-19 discussions on social media [66, 92], we calculated the percentage of tweets having a positive sentiment using the HuggingFace Transformer sentiment analyzer [107]. Then, we calculated the percentage of comments identified as inflammatory using the HateSonar hate speech classifier [26], which was shown to work well on Twitter data.

3.4 Results

Our topic modelling results suggest two main themes in Twitter discussions: climate action during COVID-19 recovery (6 out of 15 topics; summarized in Table 3.3); and lessons learned for future climate action (8 out of 15 topics; summarized in Table 3.4). These two themes cover all but one topic, shown in Table 3.2, containing specific discussions about China's role in the pandemic and in the global carbon footprint.

Tables 3.2, 3.3 and 3.4 contain the following information, from left to right: the topic number, the top-10 representative terms, the frequent n-grams, a topic description based on manual inspection of the most frequently retweeted tweets, the number of tweets assigned to the topic, the percentage of tweets with offensive language (OL), and the percentage of tweets with a positive sentiment (PS). The topics are sorted by size, i.e., the number of tweets assigned to them. In the remainder of this section, we discuss the two main themes in more detail.

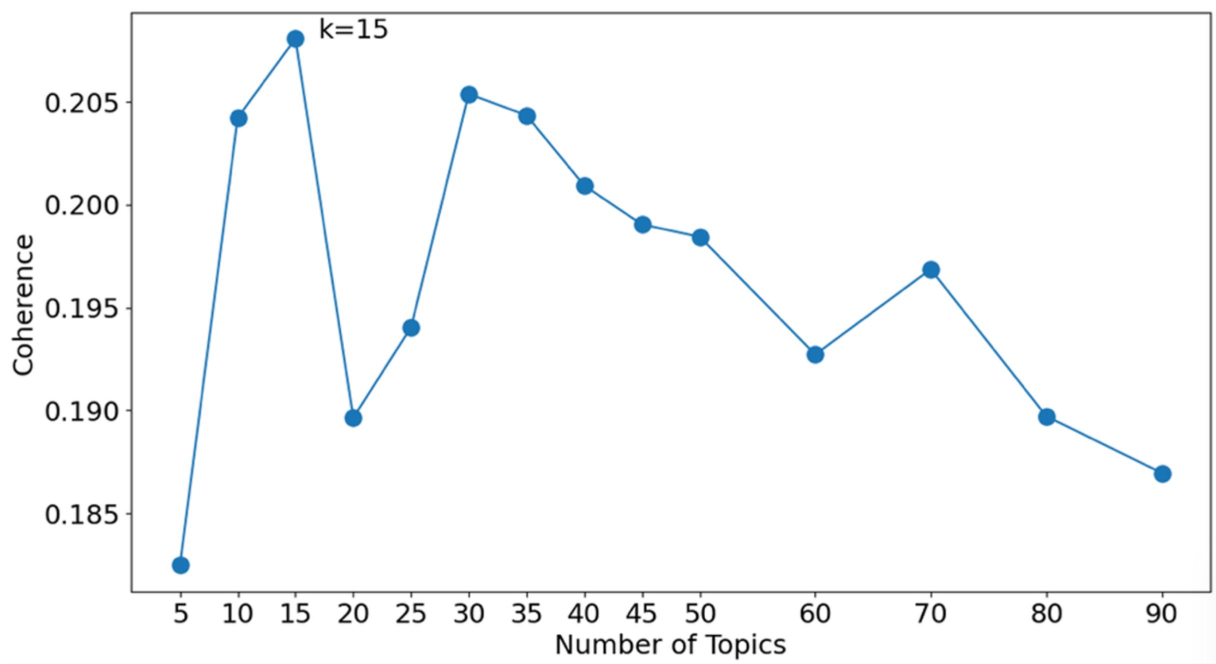

Figure 3.1: Coherence scores for different numbers of topics
An image of a graph

3.4.1 Theme 1: Climate action during COVID-19 recovery

Table 3.3 summarizes the topics related to climate action during the pandemic and beyond. Over 60 percent of the tweets in our dataset are related to this theme. In terms of sentiment, topics 1, and 9 have the highest fraction of positive tweets, encouraging climate action during COVID-19 recovery.

Table 3.3: Topics related to climate action during COVID-19 recovery

	Representative terms	Frequent n-grams	Topic description	#	%OL	%PS

1	need, time, new, threat, tackle, health, action, recovery, make, future	'earthday 2020', 'long term', 'public health', 'flatten curve', 'cop 26', 'economic recovery', 'recovery plan', 'existential threat', 'fossil fuel', 'post covid'	COVID-19 economic recovery plans should include climate action such as reducing fossil fuel use and CO2 emissions.	7901	1.5	31.4
2	human, pollution, emission, nature, earth, china, reduce, way, air, carbon	'air pollution', 'carbon emission', 'co2 emission', 'fossil fuel', 'human die', 'earth day', 'nature wipeout', 'economic demographic', 'mother nature', 'fuel nature'	Economic shutdowns reduced CO2 emissions and air pollution, but more action is needed. However, activities such as using public transit become risky during the pandemic.	4793	1.7	20.7
7	think, thing, know, wait, bad, good, really, cause, hear, happen	'people think', 'would think', 'think go', 'think bad', 'still think', 'bad would', 'think virus', 'also think', 'think serious', 'bad wait'	General discussions about the pandemic and long-term effects of climate change.	2468	4.8	19.4
8	like, look, disease, make, thing, feel, issue, treat, sound, right	'look like', 'feel like', 'disease like', 'issue like', 'thing like', 'seem like', 'sound like', 'infectious disease', 'threat like', 'challenge like'	Climate change may become a crisis similar to or worse than the COVID-19 pandemic.	2340	5.6	21.4

9	fight, help, lesson, money, let, use, warren, elizabeth, deal, way	'help fight', 'must fight', 'fight back', 'elizabeth warren', 'use fight', 'deal chinese', 'chinese must', 'deal chinese must', 'chinese must fight', 'help tackle'	Supportive and discouraging discussions around the use of pandemic recovery funds for climate action.	1932	1.8	31.8
12	stop, work, tax, pay, eat, animal, care, justice, racial, poor	'stop eat', 'eat animal', 'work poor', 'racial justice', 'poor racial', 'work poor racial', 'poor racial justice', 'stop eat animal', 'eat meat', 'care work'	Discussions encouraging a plan-based diet to protect the environment and prevent the spread of animal-borne diseases in the future.	1423	6.1	18.6

Topic 1 and 2, the most frequent topics, discuss recent improvements in air quality and express concerns about their temporary nature. We show two example tweets below.

"#COVID19 locking down the whole world can be considered a LARGE SCALE EXPERIMENT for reduction of emissions. WE all can see the difference in the BLUE skies and through breathing clean air. Q: What will happen when #Covid_19 leaves us alone#Enviroment #ClimateChange"

"COVID-19 shutdowns are clearing the air, but pollution will return as economies reopen — The shutdowns aren't slowing climate change."

As a result of these concerns, many tweets expressed the opinion that economic recovery plans should target climate action. Frequently suggested actions include reductions in the consumption of fossil fuels (Topic 1), encouraging a sustainable lifestyle with a plant-based diet (Topic 12), and investments in green transportation infrastructure such as protected cycle lanes and making public transit safe to use during the pandemic (Topic 9). We give an example tweet below.

"Please support Clean State's campaign to develop a budget that supports sustainable employment and economic recovery package based on addressing a threat even greater than COVID-19: #ClimateChange. Sign up to support our open letter."

Table 3.2: Topic related to China's role in the pandemic and in the global carbon footprint

	Representative terms	Frequent n-grams	Topic description	#	%OL	%PS
6	virus, cure, real, cause, chinese, worry, new, spread, greta, deadly	'chinese virus', 'deadly virus', 'new virus', 'spread virus', 'virus like', 'virus spread', 'kill virus', 'virus go', 'cause virus', 'virus get'	Discussions about China being the origin of the Coronavirus and a significant source of CO2 emissions.	2608	4.6	16.4

In addition to calling on governments and world leaders to support climate action (Topic 4), many tweets suggested individual climate actions that can be done during the pandemic (Topic 8). As an example, the #DarkSelfieChallenge mentioned in one of the most frequently retweeted messages encouraged people to save energy by turning off their lights and taking a selfie in the dark:

"Help the planet from home! With Hyundai, I'm taking part in the #DarkSelfieChallenge. Turn all the lights off and take a selfie with the flash on. Show yourself in the dark to shed light on climate change. #DarkSelfieChallenge #EarthDay #StayHome #HyundaixBTS #NEXO"

Other examples include encouraging cycling as a sustainable mode of transportation that allows social distancing, as expressed in the following tweet.

"All hail the mighty bicycle: France to pay for cycling lessons and bike repairs to fight both coronavirus and climate change - Bicycles promote physical distancing - Paris building 750km (466 miles) more bike lanes - Bogota, Berlin, Brussels, Milan all going big on bicycles too"

In contrast to the supporters of green recovery plans, some tweets argue that recovery plans should only target economic growth (topic 9), at least for now, and should suspend climate action for faster economic recovery. According to these tweets, it would be more dangerous to the environment in the long run if the carbon tax drives businesses into bankruptcy, as this would lead to a weakened economy that is unable to fund climate action plans (Topic 7). We show two example tweets below.

*"The only way that would work is when the economies are stable and we doing it to tackle global warming and climate change. Where we indoors so planet earth can't breathe. Otherwise we ain't doing that **** as a remembrance"*

"Are you seriously that dumb to bring climate change and adding additional carbon tax at a time of this pandemic? The 75% is for small to medium sized businesses. Rgat wont come free. They will pay double at the other end. Same for anyone deferring mortgages, car payments etc."

3.4.2 Theme 2: Lessons learned during the pandemic to help combat climate change

The second theme focuses on the lessons learned from the pandemic and discusses whether these lessons should be applied to mitigate climate change (Table 3.4). Topic 10 has a significantly high fraction of positive tweets that encourage applying lessons learned from the COVID-19 crisis into fighting climate change. On the other hand, topics 11, 13 and 15, which criticize government policies and express skepticism towards the virus and climate change, have the lowest fraction of positive tweets. The fraction of offensive content is generally low, though it appears that skeptics express their thoughts more negatively and aggressively compared to other users

Table 3.4: Topics related to lessons learned during the pandemic to help combat climate change

	Representative terms	Frequent n-grams	Topic description	#	%OL	%PS
3	say, science, scientist, believe, expert, listen, response, pope, deny, ignore	'scientist say', 'believe science', 'listen expert', 'listen scientist', 'nature response', 'greta thunberg', 'pope francis', 'anti science', 'listen science', 'science denial'	Discussions of statements made by scientists, skeptics, and public figures such as Pope Francis and Greta Thunberg.	3835	3.6	18.5
4	world, end, war, save, leader, react, happen, post, year, control	'around world', 'world war', 'world leader', 'end world', 'world end', 'world health', 'save world', 'world economy', 'happen world', 'world need'	The pandemic and climate change are both issues that must be addressed by the whole world.	3047	3.4	26.8

5	people, die, care, believe, young, make, old, million, worry, imagine	'people die', 'go die', 'young people', 'many people', 'million people', 'people go', 'old people', 'people believe', 'take seriously', 'people would'	The Coronavirus is more dangerous for older adults; climate change is more dangerous for future generations.	2718	5.7	17.7
8	like, look, disease, make, thing, feel, issue, treat, sound, right	'look like', 'feel like', 'disease like', 'issue like', 'thing like', 'seem like', 'sound like', 'infectious disease', 'threat like', 'challenge like'	Climate change may become a crisis similar to or worse than the COVID-19 pandemic.	2340	5.6	21.4
10	crisis, warn, uk, avoid, lesson, opportunity, tackle, learn, climatecrisis, spur	‘health crisis’, ‘economic crisis’, ‘two crisis’, ‘green recovery’, ‘uk warn’, ‘current crisis’, ‘warn avoid’, ‘avoid crisis’, ‘uk warn avoid’, ‘tackle crisis’	Lessons learned from pandemic response should be applied to climate action.	1853	0.4	31.2
11	trump, blame, warren, elizabeth, disease, point, president, donald, long, obama	‘elizabeth warren’, ‘donald trump’, ‘like trump’, ‘disease like’, ‘blame trump’, ‘point blame’, ‘blame disease’, ‘point blame disease’, ‘blame disease like’, ‘disease like trump’	Criticisms of the U.S. pandemic and climate policies.	1808	4.5	12.1

13	hoax, chinese, trump, real, impeachment, democrat, russian, fake, president, russia	‘hoax hoax’, ‘call hoax’, ‘chinese hoax’, ‘say hoax’, ‘hoax like’, ‘think hoax’, ‘virus hoax’, ‘impeachment hoax’, ‘trump call’, ‘hoax impeachment’	Skepticism towards the Coronavirus and climate change, e.g., calling the Coronavirus a ‘hoax’, especially in the context of U.S. politics.	1059	5.6	10.1
14	kill, year, worry, million, end, heat, maybe, thing, right, want	‘go kill’, ‘kill people’, ‘kill million’, ‘kill first’, ‘virus kill’, ‘go die’, ‘kill everyone’, ‘total death’, ‘flu kill’, ‘kill many’	Climate change may be as devastating and deadly as the Coronavirus crisis.	874	7.9	13.3
15	model, wrong, use, computer, accurate, predict, base, year, death, tell	‘computer model’, ‘model model’, ‘model use’, ‘model wrong’, ‘model predict’, ‘virus model’, ‘model accurate’, ‘scientific model’, ‘prove wrong’, ‘accurate wrong’	Skepticism towards pandemic and climate models.	802	1.2	13

A frequently mentioned lesson (Topic 10) is that global consensus and collaboration were critical during the COVID-19 pandemic and will be critical in future climate action. It was observed that some countries implemented pandemic response measures early and were able to control the spread of the virus; similarly, preparation and preventive measures will be important in the context of climate action. Additionally, specific pandemic response actions were mentioned as being helpful in the context of climate action, such as travel restrictions and working from home (Topic 10). We give three example tweets below.

“Great piece on the lessons that the Coronavirus response already has for climate change: early, effective and far-reaching action (don’t wait for consequences to unfold) and international collaboration and solidarity.”

“Inslee: ‘we should not be intimidated by people who say you should not use this COVID crisis to peddle a solution to climate change.’ He’s using this crisis to push his agenda.

He doesn't care about businesses, life savings, careers, lost."

"Global inaction on climate change offers grim lessons in the age of coronavirus: 'Only in hindsight will we really understand what we gambled on' one climate scientist said.' "

Another common opinion reflecting lessons learned was that climate change crisis may be similar to or worse than the Coronavirus, for example, in terms of fatality rates (Topic 14). Additionally, it was observed that while the virus is more dangerous for the older population, climate change will be dangerous for the next generation (Topic 5). The following tweet is one of many that express this opinion.

"Corona Virus is an existential threat to old people. Climate Change is an existential threat to young people. Recipe for a coalition of the many #BernieSurge"

The third common discussion topic in the context of lessons learned was the role of science in decision and policymaking (Topic 3), as expressed by the following tweet.

"Diseases like coronavirus remind us why we need robust institutions and investments in public health, and a government that is ready to respond at any moment. That means using science-based policy and confronting climate change, which will affect how diseases emerge and spread."

On a related note, there were negative reactions to statements made by public figures who underestimated the seriousness of the Coronavirus and global warming (Topic 11). However, Pope Francis' statement that the "Coronavirus pandemic could be nature's response to the climate crisis" received a mixed response. Some tweets agreed with this statement, while others pointed out a lack of scientific proof (Topic 3).

On the other hand, a small minority of skeptics continued to doubt climate change and criticized any use of the pandemic to advance climate action (Topic 13), as shown in the following tweet.

"Ms Swedish environmental goddess her team of alarmed activists may try to link #coronarovirus to #ClimateChange Won't be surprised if they made a speech relating #ClimateEmergency to #coronoavirus outbreak lol #climatechangehoax #ClimateHoax #auspol #auspol2020 #Australia "

Furthermore, some tweets questioned the accuracy of climate and COVID-19 prediction models. They argued that these models are not reliable at estimating deaths caused by the virus and deaths that climate change will cause in the future (Topic 15). We give an example tweet below.

"Why would anyone trust the climate change models when the Covid-19 models have been off by an incredible (perhaps deliberate) amount?"

3.5 Discussions

Our study of Twitter discussions on climate change during the COVID-19 pandemic revealed two main themes: Climate action during COVID-19 recovery and lessons learned for the future.

Climate action during COVID-19 recovery is largely supported on Twitter, with positive opinions on actions such as investing in sustainability and education, as well as social changes such as promoting the use of bicycles and public transit. Many climate actions discussed on Twitter align with those suggested by recent research (recall Section 2.3).

However, some tweets express concerns about the higher risk of contracting the Coronavirus when using public transit. Thus, it appears that there is at least some willingness to continue making environmentally-friendly decisions during the pandemic as long as policies are in place to mitigate risks. This finding should be of interest to local governments wishing to encourage climate-friendly behaviour during the pandemic.

Similarly, many tweets positively reflect on the lessons learned during the pandemic that may shape future climate action, such as the importance of preventive measures, the role of science in public policymaking, and the need for coordinated global action. In particular, we identified a topic (Topic 8; Table 3.4) reflecting the opinion that climate change may become a worse global crisis than COVID-19, underscoring the need for global action. Additionally, we found tweets that criticized politicians who ignore both climate change and COVID-19 safety protocols such as face masks, reinforcing the importance of making scientifically-sound decisions in the future (Topic 11; Table 3.4). Again, many lessons discussed on Twitter align with those reported in recent research.

On the other hand, as observed in pre-pandemic social media studies [17, 106, 82], climate skeptics continue to have a presence on Twitter. We additionally found that skeptics have incorporated the pandemic and the associated economic crisis into their reasoning for suspending climate policies such as the carbon tax. Moreover, some tweets use skepticism for one issue to justify skepticism for another, as evidenced by tweets calling both climate change and the Coronavirus pandemic a hoax, and those that mistrust climate models because COVID-19 models are believed to be inaccurate. An in-depth investigation on how misinformation about climate change creates misinformation about the pandemic and vice versa is an interesting direction for future research.

Chapter 4

Conclusion

As a result of COVID-19, the global economy and social system were devastated. As a result of health measures such as social distance, COVID-19 affected the way we studied. Second, our economy was negatively impacted by mobility restrictions and shutdowns resulting from this crisis. This thesis aimed to understand how these socioeconomic impacts affected people. Thus, we selected one problem from each of these two impact areas, namely online education and climate action, for further study.

4.1 Actionable Insights

The result of this thesis provides the following actionable insights for policy makers and academic institutions.

4.1.1 Online education

1. Online assessments should be designed in a way that reduces cheating without making them excessively difficult.

2. Academic integrity policies should be clearly communicated so that students know what is and is not allowed in various situations, especially during virtually-proctored examinations.

3. Data collection and data use policies should be clearly stated if students are being recorded during online examinations.

4.1.2 Climate action

1. An actionable insight for policy makers is to promote the use of public transportation by mitigating its health risks.

4.2 Limitations and Future Work

One limitation of this thesis is its focus on English language content on Twitter and Reddit platforms. Another limitation is that the study was based on data collected from publicly-accessible social media platforms and therefore only represents the views of users who contributed to those discussions. Lastly, a limitation of this study is that it only examines content that matches specific keywords. In spite of the fact that previous research has shown that these keywords are effective in capturing the majority of the discussions, there is a possibility that some discussions may be overlooked if they do not contain these keywords. Nevertheless, our findings can be used as a starting point for additional focused research, such as the following.

4.2.1 Online education

1. A interesting direction for future research is to mine course discussion forums for more opinions on the pros and cons of online learning tools and platforms.

2. Similar to student who studied from home, many individuals were forced to work from home during the pandemic. Hence, one future direction would be to investigate the issues related to working from home.

4.2.2 Climate action

1. A direction for future work is to compare public attitudes towards climate action during COVID-19 recovery in Asia, Europe and North America, and correlate these opinions with pandemic response policies.

2. It is also important to analyze the role of science in policy making and how public conform to these policies.

3. An in-depth investigation on how misinformation about climate change creates misinformation about the pandemic and vice versa is another interesting direction for future research.

We conducted a data-mining study of the social impacts of the COVID-19 pandemic using a large social media corpus. We found social media platforms useful tools for understanding public opinion towards social issues. By constantly monitoring discussions on social media, policy makers can observe how certain decisions are impacting the public and how people are reacting to these decisions.

This thesis showed the importance of social media mining in understating public opinion towards variety of issues related to the COVID-19 pandemic, from academic dishonesty to climate action. Hence, researchers can use social media to investigate other issues related to pandemic, such as vaccine hesitancy.

References

[1] Giora Alexandron, José A. Ruipérez-Valiente, and David Pritchard. Towards a general purpose anomaly detection method to identify cheaters in massive open online courses. In *EDM*, pages 480–483, 06 2019.

[2] Daniel Allington, Bobby Duffy, Simon Wessely, Nayana Dhavan, and James Rubin. Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, pages 1–7, June 2020.

[3] Sharon Theresa Alston, Carl S. Moore, and Morris Thomas. Strategies for enhancing online teaching in social work education. *Journal of Human Behavior in the Social Environment*, pages 1–10, April 2017.

[4] Ashley A. Anderson. Effects of social media use on climate change opinion, knowledge, and behavior, 03 2017.

[5] Oberiri Destiny Apuke and Bahiyah Omar. Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, 56:101475, January 2021.

[6] Tina Askanius and Julie Uldam. Online social media for radical politics: climate change activism on YouTube. *International Journal of Electronic Governance*, 4(1/2):69, 2011.

[7] Sitaram Asur and Bernardo A. Huberman. Predicting the future with social media. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. IEEE, August 2010.

[8] Shrey Bagroy, Ponnurangam Kumaraguru, and Munmun De Choudhury. A social media based index of mental well-being in college campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1634–1646. ACM, 2017.

[9] Craig J. Bailey and Karen A. Card. Effective pedagogical practices for online teaching: Perception of experienced instructors. *The Internet and Higher Education*, 12(3-4):152–155, December 2009.

[10] Edward B. Barbier and Joanne C. Burgess. Sustainability and development after COVID-19. *World Development*, 135:105082, November 2020.

[11] Kristine Belesova, David L Heymann, and Andy Haines. Integrating climate action for health into covid-19 recovery plans. *BMJ*, page m3169, September 2020.

[12] Peter Berglez and Walid Al-Saqaf. Extreme weather and climate change: social media results, 2008–2017. *Environmental Hazards*, pages 1–18, October 2020.

[13] Jesse D. Berman and Keita Ebisu. Changes in u.s. air pollution during the COVID-19 pandemic. *Science of The Total Environment*, 739:139864, October 2020.

[14] Eren Bilen and Alexander Matros. Online cheating amid COVID-19. *Journal of Economic Behavior & Organization*, 182:196–211, February 2021.

[15] Wouter Botzen, Sem Duijndam, and Pieter van Beukering. Lessons for climate policy from behavioral biases towards COVID-19 and climate change risks. *World Development*, 137:105214, January 2021.

[16] Thijs Bouman, Linda Steg, and Thomas Dietz. Insights from early COVID-19 responses about promoting sustainable action. *Nature Sustainability*, October 2020.

[17] Constantine Boussalis and Travis G Coan. Text-mining the signals of climate change doubt. *Global Environmental Change*, 36:89–100, 2016.

[18] Emily Chen, Kristina Lerman, and Emilio Ferrara. Tracking social media discourse about the COVID-19 pandemic: Development of a public coronavirus twitter data set. *JMIR Public Health and Surveillance*, 6(2):e19273, May 2020.

[19] Matteo Cinelli, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. The COVID-19 social media infodemic. *Scientific Reports*, 10(1), October 2020.

[20] Ted M. Clark, Christopher S. Callam, Noel M. Paul, Matthew W. Stoltzfus, and Daniel Turner. Testing in the time of COVID-19: A sudden transition to unproctored online exams. *Journal of Chemical Education*, 97(9):3413–3417, July 2020.

[21] Emily M. Cody, Andrew J. Reagan, Lewis Mitchell, Peter Sheridan Dodds, and Christopher M. Danforth. Climate change sentiment on twitter: An unsolicited public opinion poll. *PLOS ONE*, 10(8):e0136092, August 2015.

[22] Maurie J. Cohen. Does the COVID-19 outbreak mark the onset of a sustainable consumption transition? *Sustainability: Science, Practice and Policy*, 16(1):1–3, March 2020.

[23] R. M. Colvin, Luke Kemp, Anita Talberg, Clare De Castella, C. Downie, S. Friel, Will J. Grant, Mark Howden, Frank Jotzo, Francis Markham, and Michael J. Platow. Learning from the climate change debate to avoid polarisation on negative emissions. *Environmental Communication*, 14(1):23–35, July 2019.

[24] Michael D. Conover, Clayton Davis, Emilio Ferrara, Karissa McKelvey, Filippo Menczer, and Alessandro Flammini. The geospatial characteristics of a social movement communication network. *PLoS ONE*, 8(3):e55957, March 2013.

[25] Biraj Dahal, Sathish A. P. Kumar, and Zhenlong Li. Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, 9(1), June 2019.

[26] Thomas Davidson, Dana Warmsley, Michael Macy, and Ingmar Weber. Automated Hate Speech Detection and the Problem of Offensive Language. *arXiv e-prints*, page arXiv:1703.04009, March 2017.

[27] Anneliese Depoux, Sam Martin, Emilie Karafillakis, Raman Preet, Annelies Wilder-Smith, and Heidi Larson. The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine*, 27(3), March 2020.

[28] Jessica R. Deters, Marie C. Paretti, and Jennifer M. Case. How implicit assumptions about engineering impacted teaching and learning during COVID-19. *Advances in Engineering Education*, 8(4):1–5, 2020.

[29] Shivangi Dhawan. Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems*, 49(1):5–22, June 2020.

[30] Biljana Dorić, Marija Blagojević, Milos Papic, and Nebojša Stanković. Students' attitudes regarding online learning during COVID-19 pandemic. In *Information Technology and Education Development*, pages 157–160, 10 2020.

[31] Patrick J. Egan and Megan Mullin. Turning personal experience into political attitudes: The effect of local weather on americans' perceptions about global warming. *The Journal of Politics*, 74(3):796–809, July 2012.

[32] Miriam Fernandez, Lara S. G. Piccolo, Diana Maynard, Meia Wippoo, Christoph Meili, and Harith Alani. Talking climate change via social media. In *Proceedings of the 8th ACM Conference on Web Science*. ACM, May 2016.

[33] Emilio Ferrara. What types of COVID-19 conspiracies are populated by twitter bots? *First Monday*, May 2020.

[34] Kelum A.A. Gamage, Erandika K. de Silva, and Nanda Gunawardhana. Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences*, 10(11):301, October 2020.

[35] Junling Gao, Pinpin Zheng, Yingnan Jia, Hao Chen, Yimeng Mao, Suhong Chen, Yi Wang, Hua Fu, and Junming Dai. Mental health problems and social media exposure during COVID-19 outbreak. *PLOS ONE*, 15(4):e0231924, April 2020.

[36] Alanna Gillis and Laura M. Krull. COVID-19 remote learning transition in spring 2020: Class structures, student perceptions, and inequality in college courses. *Teaching Sociology*, 48(4):283–299, September 2020.

[37] A. Hajibagheri and G. Sukthankar. Political polarization over global warming : Analyzing twitter data on climate change. volume 74, pages 1–5. ASE, May 2014.

[38] Robin Haunschild, Loet Leydesdorff, Lutz Bornmann, Iina Hellsten, and Werner Marx. Does the public discuss other topics on climate change than researchers? a comparison of explorative networks based on author keywords and hashtags. *Journal of Informetrics*, 13(2):695–707, 2019.

[39] Bao-xun HE. How to prevent online learners from cheating? *EELSS*, pages 72–76, 07 2020.

[40] Michael Heberling and M Flint. Maintaining academic integrity in online education. *Online Journal of Distance Learning Administration*, V(1):1–6, 2002.

[41] Dieter Helm. The environmental impacts of the coronavirus. *Environmental and Resource Economics*, 76(1):21–38, May 2020.

[42] Cameron Hepburn, Brian O'Callaghan, Nicholas Stern, Joseph Stiglitz, and Dimitri Zenghelis. Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change? *Oxford Review of Economic Policy*, 36(Supplement_1):S359–S381, 2020.

[43] Craig Hlavac and Jeffrey L. Buller. Leading the academic department during the COVID-19 crisis. *The Department Chair*, 31(1):17–19, June 2020.

[44] Candice Howarth, Peter Bryant, Adam Corner, Sam Fankhauser, Andy Gouldson, Lorraine Whitmarsh, and Rebecca Willis. Building a social mandate for climate action: Lessons from COVID-19. *Environmental and Resource Economics*, 76(4):1107–1115, July 2020.

[45] Ting Hua, Feng Chen, Liang Zhao, Chang-Tien Lu, and Naren Ramakrishnan. Automatic targeted-domain spatiotemporal event detection in twitter. *GeoInformatica*, 20(4):765–795, August 2016.

[46] C.J. Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*, pages 216–225, 01 2015.

[47] Md Saiful Islam, Tonmoy Sarkar, Sazzad Hossain Khan, Abu-Hena Mostofa Kamal, S. M. Murshid Hasan, Alamgir Kabir, Dalia Yeasmin, Mohammad Ariful Islam, Kamal Ibne Amin Chowdhury, Kazi Selim Anwar, Abrar Ahmad Chughtai, and Holly Seale. COVID-19–related infodemic and its impact on public health: A global social media analysis. *The American Journal of Tropical Medicine and Hygiene*, 103(4):1621–1629, October 2020.

[48] Marnie V. Jamieson. Keeping a learning community and academic integrity intact after a mid-term shift to online learning in chemical engineering design during the COVID-19 pandemic. *Journal of Chemical Education*, 97(9):2768–2772, August 2020.

[49] S. Mo Jang and P. Sol Hart. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter big data. *Global Environmental Change*, 32:11–17, May 2015.

[50] S. Kemp. *Digital 2020 October Global Statshot Report (October 2020).* 2020.

[51] Abeer Khan and Lukasz Golab. Reddit mining to understand gendered movements. In *Proc. EDBT Workshop on Data Analytics Solutions for Real-Life Applications*, pages 3:1–3:8, 2020.

[52] Andrei P. Kirilenko, Tatiana Molodtsova, and Svetlana O. Stepchenkova. People as sensors: Mass media and local temperature influence climate change discussion on twitter. *Global Environmental Change*, 30:92–100, January 2015.

[53] Andrei P. Kirilenko and Svetlana O. Stepchenkova. Public microblogging on climate change: One year of twitter worldwide. *Global Environmental Change*, 26:171–182, May 2014.

[54] P. Kitcher. The climate change debates. *Science*, 328(5983):1230–1234, May 2010.

[55] Anna Kovacs-Györi, Alina Ristea, Ronald Kolcsar, Bernd Resch, Alessandro Crivellari, and Thomas Blaschke. Beyond spatial proximity—classifying parks and their visitors in london based on spatiotemporal and sentiment analysis of twitter data. *ISPRS International Journal of Geo-Information*, 7(9):378, September 2018.

[56] Caglar Koylu, Ryan Larson, Bryce J. Dietrich, and Kang-Pyo Lee. CarSenToGram: geovisual text analytics for exploring spatiotemporal variation in public discourse on twitter. *Cartography and Geographic Information Science*, 46(1):57–71, September 2018.

[57] Caroline Kuzemko, Michael Bradshaw, Gavin Bridge, Andreas Goldthau, Jessica Jewell, Indra Overland, Daniel Scholten, Thijs Van de Graaf, and Kirsten Westphal. Covid-19 and the politics of sustainable energy transitions. *Energy Research & Social Science*, 68:101685, October 2020.

[58] Jihyun Lee, Ryan Jinyoung Kim, Shin-Young Park, and Marcus A. Henning. Using technologies to prevent cheating in remote assessments during the COVID-19 pandemic. *Journal of Dental Education*, pages 1–3, August 2020.

[59] Stephan Lewandowsky, John Cook, Nicolas Fay, and Gilles E. Gignac. Science by social media: Attitudes towards climate change are mediated by perceived social consensus. *Memory & Cognition*, 47(8):1445–1456, June 2019.

[60] Daniel M Low, Laurie Rumker, Tanya Talkar, John Torous, Guillermo Cecchi, and Satrajit S Ghosh. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: Observational study. *Journal of Medical Internet Research*, 22(10):e22635, October 2020.

[61] Renato Assis Machado, Natália Lins de Souza, Rayane Maria Oliveira, Hercílio Martelli Júnior, and Paulo Rogério Ferreti Bonan. Social media and

telemedicine for oral diagnosis and counselling in the COVID-19 era. *Oral Oncology*, 105:104685, June 2020.

[62] Raoul Manenti, Emiliano Mori, Viola Di Canio, Silvia Mercurio, Marco Picone, Mario Caffi, Mattia Brambilla, Gentile Francesco Ficetola, and Diego Rubolini. The good, the bad and the ugly of COVID-19 lockdown effects on wildlife conservation: Insights from the first european locked down country. *Biological Conservation*, 249:108728, September 2020.

[63] Rubén D. Manzanedo and Peter Manning. COVID-19: Lessons for the climate change emergency. *Science of The Total Environment*, 742:140563, November 2020.

[64] Yago Martín, Zhenlong Li, and Susan L. Cutter. Leveraging twitter to gauge evacuation compliance: Spatiotemporal analysis of hurricane matthew. *PLOS ONE*, 12(7):e0181701, July 2017.

[65] Paolo Missier, Alexander Romanovsky, Tudor Miu, Atinder Pal, Michael Daniilakis, Alessandro Garcia, Diego Cedrim, and Leonardo da Silva Sousa. Tracking dengue epidemics using twitter content classification and topic modelling. In *Current Trends in Web Engineering*, pages 80–92. Springer International Publishing, 2016.

[66] Vasee Moorthy, Ana Maria Henao Restrepo, Marie-Pierre Preziosi, and Soumya Swaminathan. Data sharing for novel coronavirus (COVID-19). *Bulletin of the World Health Organization*, 98(3):150–150, March 2020.

[67] Teagen Nabity-Grover, Christy M.K. Cheung, and Jason Bennett Thatcher. Inside out and outside in: How the COVID-19 pandemic affects self-disclosure on social media. *International Journal of Information Management*, 55:102188, December 2020.

[68] Mohd Shahrul Mohd Nadzir, Maggie Chel Gee Ooi, Kemal Maulana Alhasa, Mohd Aftar Abu Bakar, Anis Asma Ahmad Mohtar, Mohd Fadzil Firdzaus Mohd Nor, Mohd Talib Latif, Haris Hafizal Abd Hamid, Sawal Hamid Md Ali, Noratiqah Mohd Ariff, Johary Anuar, Fatimah Ahamad, Azliyana Azhari, Norfazrin Mohd Hanif, Mohammed Ahmed Subhi, Murnira Othman, and Mohd Zaim Mohd Nor. The impact of movement control order (MCO) during pandemic COVID-19 on local air quality in an urban area of klang valley, malaysia. *Aerosol and Air Quality Research*, 20(6):1237–1248, 2020.

[69] Robert Newell and Ann Dale. COVID-19 and climate change: an integrated perspective. *Cities & Health*, pages 1–5, July 2020.

[70] Curtise Kin Cheung Ng. Evaluation of academic integrity of online open book assessments implemented in an undergraduate medical radiation science course during COVID-19 pandemic. *Journal of Medical Imaging and Radiation Sciences*, 51(4):610–616, December 2020.

[71] Joseph G. Nguyen, Kristopher J. Keuseman, and Jonathan J. Humston. Minimize online cheating for online assessments during COVID-19 pandemic. *Journal of Chemical Education*, 97(9):3429–3435, August 2020.

[72] Dongyang Nie, Fuzhen Shen, Junfeng Wang, Xiaoyun Ma, Zhirao Li, Pengxiang Ge, Yang Ou, Yuan Jiang, Meijuan Chen, Mindong Chen, Tijian Wang, and Xinlei Ge. Changes of air quality and its associated health and economic burden in 31 provincial capital cities in china during COVID-19 pandemic. *Atmospheric Research*, 249:105328, February 2021.

[73] Navid Ibtehaj Nizam, Shan Gao, Mengzhou Li, Hisham Mohamed, and Ge Wang. Scheme for cheating prevention in online exams during social distancing. *Preprints*, April 2020.

[74] Nima Norouzi, Gerardo Zarazua de Rubens, Saeed Choupanpiesheh, and Peter Enevoldsen. When pandemics impact economies and climate change: Exploring the impacts of COVID-19 on oil and electricity demand in china. *Energy Research & Social Science*, 68:101654, October 2020.

[75] Derek O'Callaghan, Derek Greene, Joe Carthy, and Pádraig Cunningham. An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42(13):5645–5657, August 2015.

[76] Alexandra Olteanu, Carlos Castillo, Nicholas Diakopoulos, and Karl Aberer. Comparing events coverage in online news and social media: The case of climate change. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, 2015.

[77] David Alfred Ostrowski. Using latent dirichlet allocation for topic modelling in twitter. In *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015)*. IEEE, February 2015.

[78] Brendan O'Connor, Ramnath Balasubramanyan, Bryan Routledge, and Noah Smith. From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1), May 2010.

[79] Jose J. Padilla, Hamdi Kavak, Christopher J. Lynch, Ross J. Gore, and Saikou Y. Diallo. Temporal and spatiotemporal investigation of tourist attraction visit sentiment on twitter. *PLOS ONE*, 13(6):e0198857, June 2018.

[80] M. Parsa and Lukasz Golab. Social media mining to understand the impact of co-operative education on mental health. In *EDM*, pages 653–657, 2020.

[81] Warren Pearce, Brian Brown, Brigitte Nerlich, and Nelya Koteyko. Communicating climate change: conduits, content, and consensus. *Wiley Interdisciplinary Reviews: Climate Change*, 6(6):613–626, September 2015.

[82] Warren Pearce, Sabine Niederer, Suay Melisa Özkula, and Natalia Sánchez Querubín. The social media life of climate change: Platforms, publics, and future imaginaries. *Wiley Interdisciplinary Reviews: Climate Change*, 10(2):e569, 2019.

[83] James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. The development and psychometric properties of LIWC 2015. Technical report, 2015.

[84] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7):770–780, June 2020.

[85] Jill M. Price, Joy Whitlatch, Cecilia Jane Maier, Melissa Burdi, and James Peacock. Improving online teaching by using established best classroom teaching practices. *The Journal of Continuing Education in Nursing*, 47(5):222–227, May 2016.

[86] Bruce Prideaux, Michelle Thompson, and Anja Pabel. Lessons from COVID-19 can prepare global tourism for the economic transformation needed to combat climate change. *Tourism Geographies*, 22(3):667–678, May 2020.

[87] Sonali Raje and Shannon Stitzel. Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. *Journal of Chemical Education*, 97(9):3436–3440, August 2020.

[88] Daniel Rosenbloom and Jochen Markard. A COVID-19 recovery for climate. *Science*, 368(6490):447–447, April 2020.

[89] Nicholas Roxburgh, Dabo Guan, Kong Joo Shin, William Rand, Shunsuke Managi, Robin Lovelace, and Jing Meng. Characterising climate change discourse on social media during extreme weather events. *Global Environmental Change*, 54:50–60, January 2019.

[90] Saeida Saadat, Deepak Rawtani, and Chaudhery Mustansar Hussain. Environmental perspective of COVID-19. *Science of The Total Environment*, 728:138870, aug 2020.

[91] Renee N Salas. Lessons from the covid-19 pandemic provide a blueprint for the climate emergency. *BMJ*, page m3067, August 2020.

[92] Abraham C. Sanders, Rachael C. White, Lauren S. Severson, Rufeng Ma, Richard McQueen, Haniel C. Alcânatara Paulo, Yucheng Zhang, John S. Erickson, and Kristin P. Bennett. Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 twitter discourse. September 2020.

[93] Karishma Sharma, Sungyong Seo, Chuizheng Meng, Sirisha Rambhatla, Aastha Dua, and Yan Liu. Coronavirus on social media: Analyzing misinformation in twitter conversations. *CoRR*, abs/2003.12309, 2020.

[94] Ramesh P. Singh and Akshansha Chauhan. Impact of lockdown on air quality in india during COVID-19 pandemic. *Air Quality, Atmosphere & Health*, 13(8):921–928, July 2020.

[95] Dimitris Stratoulias and Narissara Nuthammachot. Air quality development during the COVID-19 pandemic over a medium-sized urban area in thailand. *Science of The Total Environment*, 746:141320, December 2020.

[96] Samia Tasnim, Md Mahbub Hossain, and Hoimonty Mazumder. Impact of rumors and misinformation on COVID-19 in social media. *Journal of Preventive Medicine and Public Health*, 53(3):171–174, May 2020.

[97] David Michael Telles-Langdon. Transitioning university courses online in response to COVID-19. *Journal of Teaching and Learning*, 14(1):108–119, May 2020.

[98] L. Tiong and H. Lee. E-cheating prevention measures: Detection of cheating at online examinations using deep learning approach - a case study. *ArXiv*, abs/2101.09841, 2021.

[99] Nik Ani Afiqah Tuah. Is online assessment in higher education institutions during COVID-19 pandemic reliable? *Siriraj Medical Journal*, 73(1):61–68, December 2020.

[100] Shem Unger and William Meiran. Student attitudes towards online education during the COVID-19 viral outbreak of 2020: Distance learning in a time of social distance. *International Journal of Technology in Education and Science*, 4(4):256–266, September 2020.

[101] M. Urrutia-Pereira, C.A. Mello da Silva, and D. Solé. COVID-19 and air pollution: A dangerous association? *Allergologia et Immunopathologia*, 48(5):496–499, September 2020.

[102] Giuseppe A. Veltri and Dimitrinka Atanasova. Climate change on twitter: Content, media ecology and information sharing behaviour. *Public Understanding of Science*, 26(6):721–737, November 2015.

[103] Abdullah Wahbeh, Tareq Nasralah, Mohammad Al-Ramahi, and Omar El-Gayar. Mining physicians' opinions on social media to obtain insights into COVID-19: Mixed methods analysis. *JMIR Public Health and Surveillance*, 6(2):e19276, June 2020.

[104] Tianyi Wang, Ke Lu, Kam Pui Chow, and Qing Zhu. COVID-19 sensing: Negative sentiment analysis on social media in china via BERT model. *IEEE Access*, 8:138162–138169, 2020.

[105] Hywel T.P. Williams, James R. McMurray, Tim Kurz, and F. Hugo Lambert. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, 32:126–138, May 2015.

[106] Hywel TP Williams, James R McMurray, Tim Kurz, and F Hugo Lambert. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global environmental change*, 32:126–138, 2015.

[107] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, R'emi Louf, Morgan Funtowicz, and Jamie Brew. Huggingface's transformers: State-of-the-art natural language processing. *ArXiv*, abs/1910.03771, 2019.

[108] Gabrielle Wong-Parodi and Irina Feygina. Understanding and countering the motivated roots of climate change denial. *Current Opinion in Environmental Sustainability*, 42:60–64, February 2020.

[109] Wei Xu, Xin Liu, and Yihong Gong. Document clustering based on non-negative matrix factorization. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*, SIGIR '03, page 267–273, New York, NY, USA, 2003. Association for Computing Machinery.

[110] Wei Xu, Xin Liu, and Yihong Gong. Document clustering based on non-negative matrix factorization. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 267–273, 2003.

[111] Jia Xue, Junxiang Chen, Ran Hu, Chen Chen, Chengda Zheng, Yue Su, and Tingshao Zhu. Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. *Journal of Medical Internet Research*, 22(11):e20550, November 2020.

[112] Hui Yin, Shuiqiao Yang, and Jianxin Li. Detecting topic and sentiment dynamics due to covid-19 pandemic using social media, 2020.

[113] Meng Yu, Zhiyong Li, Zhicheng Yu, Jiaxin He, and Jingyan Zhou. Communication related health crisis on social media: a case of COVID-19 outbreak. *Current Issues in Tourism*, pages 1–7, April 2020.

[114] Shelby Zangari, Dustin T. Hill, Amanda T. Charette, and Jaime E. Mirowsky. Air quality changes in new york city during the COVID-19 pandemic. *Science of The Total Environment*, 742:140496, November 2020.