Exploring Public Opinions of Remote Study during COVID-19 Pandemic

Jiecheng Gu & Xiaomeng Niu

University of Rochester

jgu13@ur.rochester.edu, xniu5@ur.rochester.edu

1 Introduction

Ever since the first covid-19 case was reported in China, there have been 268 million cases globally, including 5.28 million cases of deaths.[1] Not only causing pains and these shocking numbers, the virus but also changed people's lives significantly, even though not affected, in a lot of ways. The Chinese government imposed a lockdown in Wuhan, Hubei in order to quarantine the outbreak of covid-19 from January 2020 to 8 April 2020.[2] Almost 57 million people's normal life was interrupted. The government also encouraged celebrating the 2021 Spring Festival on the spot, whereas normally people go back home and celebrate the most important festival with their families.[3] Other countries also started their own ways to prevent the virus from spreading since WHO declared the coronavirus outbreak a global pandemic on 11 March 2020.[1] In the United States, governments implemented a nationwide social distancing protocol, which is an efficient way to lower the speed of virus spreading.[1]

Due to this social distance requirement and the growing number of affected and dead cases, companies and schools have to start work and study from home. WFH and online learning became well welcomed because of their safety and convenience. Even at this time, when the situation becomes better thanks to the increasing number of vaccinated individuals, some companies adapt WFH and make it regular. Whereas universities began offline teaching as soon as 2021 Fall due to the inefficiency of remote learning. There are many challenges and problems related to teachers' and students' satisfaction with online classes, such as quality and timely interaction between students and professors, technical support availability, structured online class modules, and modifications to accommodate the conduction of practical classes [4]. Moreover, the remote study might have a negative impact on people's mental health. Studies [5] have found that older and Black and Hispanic children as well as those from families with lower income who attend school remotely may experience greater impairment to mental health than their younger, White, and higher-income counterparts. For Chinese students in the U.S., the situation could be worse: some of them have not been back home for two years due to the limited number and extremely high price of airlines as well as the 14 to 21 days of quarantine policy in China.

In this paper, we will use data from social media Twitter and Weibo — a Chinese Twitter-like platform — to further examine people's attitudes towards online learning during the covid-19 pandemic. We will first examine the correlation between user features and their opinions, and then try to analyze the reasons behind these opinions and suggest ways to improve the efficiency of online learning so that it can be more popular and more adaptable. Because generalized online learning can bring many benefits to our society as well as individuals once it becomes part of everyone's daily life. For example, adapting online learning can lead to a fair distribution of educational resources in rural areas. Besides, as the volume of information grows exponentially, it is necessary for people to keep learning in order to maintain their ability to think independently. Additionally, generalized online learning also brings the chance to develop a personal education plan for everyone, in which students can truly discover their interests and take advantage of them. If remote study is adapted as a normal form in public education, students can have the chance to practice how to arrange their time, which is an important skill in future life.

1.1 Related work

There are few studies about exploring opinions on remote study using social media data, but some researchers have used social media data such as Twitter to discover people's opinions on other social issues such as elections [6], working from home (WFH) [7], COVID-19 pandemic [8], and personal face mask usage [9]. Some studies used the Valence Aware Dictionary and Sentiment Reasoner (VADER) [10] for sentiment analysis. Most studies used the M3-inference model to portray different demographic groups.[11]

Moreover, there are also researches related to remote study and online learning during the COVID-19 pandemic via survey questionnaires. Some studies are related to students' and parents' opinions on remote study. One survey at Jordan University of Science and Technology found that remote E-exam experience has a negative impact on students' habits related to diet, sleep, physical activity and smoking.[12] Another survey conducted on the basis of the opinions of parents found that distance learning education tools are not conducive to the focus of attention on the part of the children.[13] Some studies also focus on the best practices for implementing remote learning.[14] However, due to the method of survey, their samples are often limited.

2 Methodology

2.1 Data collection

In this project, we collect both Twitter data and Weibo data to perform sentiment analysis. The time range of these data is 6 months, from September to November in 2020 and from February to April in 2021 when most students around the world were studying remotely. Our data download links are attached at the end.[16, 17]

2.1.1 Twitter Data

Twitter data is fetched using Twitter APIs for academic research and it contains both user data and tweet data. The filter keywords and hashtags are "online learning", "remote learning", "distance learning", "#onlinelearning", "#remotelearning", and "#distancelearning". In total, 336,556 unique tweets posted by 162,980 unique Twitter users were sampled. The main process is shown in Table 2.1.

● Apply for a Twitter Developer account (for academic research) ● Request authentication and access API ● Return and parse JSON data ● Store returned data in CSV format

Table 2.1: Twitter data collecting methods

2.1.2 Weibo Data

The second research data consists of Weibo posts, which were collected from Weibo API and a web crawler. The API interface is provided from Weibo open platform. However, due to the access restrictions, the number of returned data from API is limited to 2,000. Therefore, we also utilized a web crawler written in Python code. The main process is shown in Table 2.2 and a total of 162,363 original Weibo content was captured for keyword "online learning (网课)".

Weibo API ● Create a Weibo APP and receive a unique key ● Request authentication and access API ● Return and parse JSON data
Web Crawler ● Login using a Weibo account and get cookies ● Simulate login using cookie ● Start searching for key-words-related content ● Store returned data in CSV format

Table 2.2: Weibo data collecting methods

2.2 Data Preprocessing

2.2.1 Twitter Data

The first step in data mining is preprocessing after the data was collected. We first drop duplicates and rows containing None value and then clean the texts of tweets. For the raw texts of tweets, we removed redundant characters such as '@'(mentions), '#' (hashtag), 'RT' (retweets), hyperlinks, emojis, and punctuations. We only keep the punctuations for future analysis.

2.2.1 Weibo Data

For the convenience of the following study, we first unify all the sentences by translating traditional Chinese into Simplified Chinese. Different from tweets, Weibo posts are mostly Chinese sentences that are continuous and do not have spaces to separate words. Therefore, the next step is word segmentation. We use the Jieba word segmentation toolkit, which is highly efficient and has high accuracy. The third step is data denoise. We remove useless characters as we did to Twitter data. We also remove advertisements by recognizing some keywords, such as “price (售价)”, “after discount (券后)”, and “fans (粉丝)”. The last step is to remove stop words, which are functional words without actual meaning. Here we used a stop words dictionary from the Harbin Institute of Technology.

2.3 Feature Inference of Tweets

2.3.1 Sentiment Score

As for Twitter data, we use VADER to measure the sentiment of the tweets. The score ranges from –1 (most negative) to +1 (most positive). The mean sentiment score is 0.289 (std: 0.461; 25th percentile: 0.000, 50th percentile: 0.262, 75th percentile: 0.647). We divide sentiment scores into three groups: positive (sentiment score > 0), neutral (sentiment score= 0) and negative (sentiment score < 0). As shown in Figure 2.1, 55.8% of Twitter users are positive while 21.1% of users are negative about the remote study.

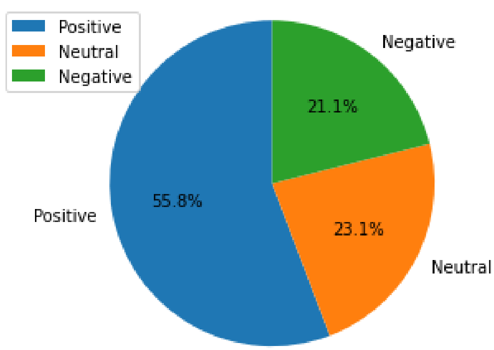

Figure 2.1: Pie chart of sentiment scores in Twitter data

2.3.2 Age

As for Twitter data, we apply the M3-inference model to infer the gender, age, and user status of each Twitter user from their profile name, username (screen name), and profile description. Age is binned into four groups: <18 years old, 19-29 years old, 30-39 years old, and 40 years old. We can see from Figure 2.2 that 57.9% of the users in our dataset are older than 40 years older, 22.3% are between 30 and 39 years old, 14.1% are between 19 and 29 years old, and the rest are younger than 19 years old.

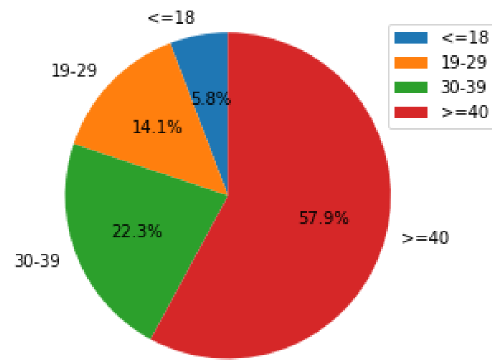

Figure 2.2: Pie chart of age in Twitter data

2.3.3 Gender

According to the previous studies, the gender distribution of Twitter users is biased toward men (71.8%) [15]. A similar pattern is also observed in our data set, where 59.1% of users are men and 40.9% are women.

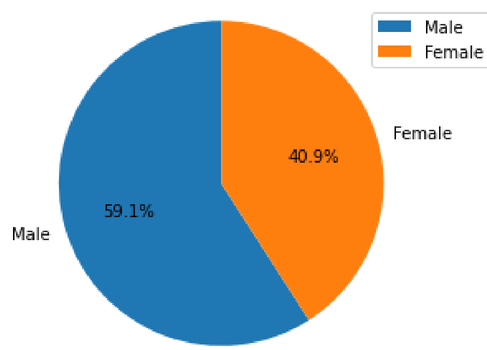

Figure 2.3: Pie chart of gender in Twitter data

2.3.4 User status

According to the M3-inference model, the status of a Twitter user is either organization or individual. Figure 2.4 shows the distribution of Twitter user status. As shown in the pie chart, 53.7% of users are individuals while 46.3% of users are from organizations.

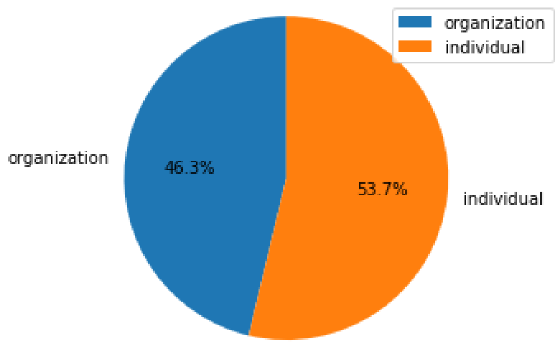

Figure 2.4: Pie chart of user status in Twitter data

2.4 Topic Modeling

To extract major topics from the tweets, we apply Latent Dirichlet Allocation (LDA) using the Gensim package in Python. First, we extend the stopwords from the Gensim package with topic-related words (e.g., “learning”, “onlinelearning”, “remotelearning” and “distancelearning”). Next, we collect nouns, verbs, and adjectives to do the topic modeling. The final model contains 6 topics with a coherence score c_v of 1.417.

3 Analysis and Experiment

3.1 Sentiment analysis of Twitter data

As for Twitter data, we compare the average sentiment score (VADER score) by gender, age, and user status. Figures 3.1, 3.2, 3.3 are the bar charts with these average sentiment scores.

We can see from Figure 3.1 that females tend to be more negative about the remote study. Although the average sentiment scores of male and female groups are both positive and very close, females’ score (0.276) is a little lower than males’ score (0.291).

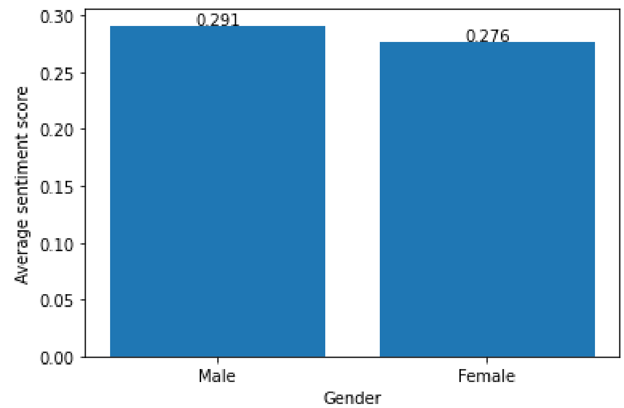

Figure 3.1: Average sentiment score of each gender in Twitter data

From Figure 3.2, it is clear that the average sentiment score of users between 19 and 29 years old is the lowest among all these four age groups, followed by users under 18 years old. Therefore, young people tend to be more negative about remote study. Since most people in these age groups are students, we can infer that students tend to be more negative about remote study, especially college students who are between 19 and 29 years old. Moreover, users over 40 years old get the highest sentiment score (0.312) and it means that they tend to be more positive about remote learning than young people.

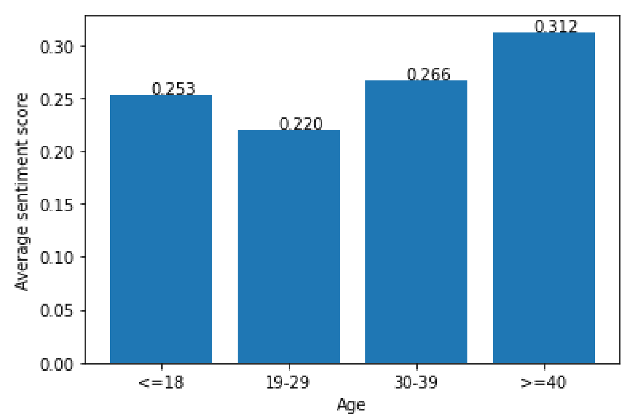

Figure 3.2: Average sentiment score of each age group in Twitter data

As we can see from Figure 3.3, Twitter users from organizations posted more positive tweets than individual users. The average sentiment score of organizations is almost one and half times higher than that of individual users.

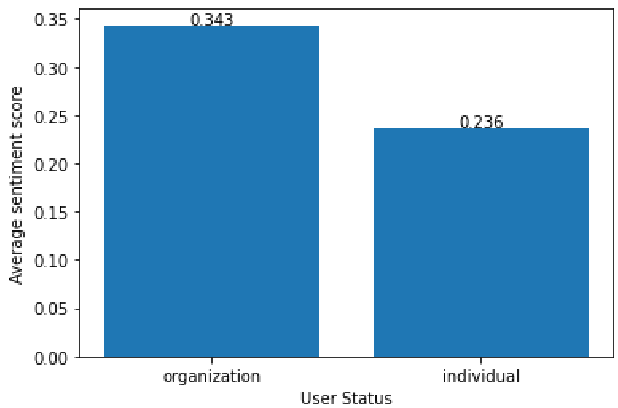

Figure 3.3: Average sentiment score of each user status in Twitter data

In addition, we also perform an ordinary least squares regression analysis on our Twitter dataset to better examine the correlation relationship between VADER sentiment scores and user characteristics. Table 3.1 shows the results of ordinary least squares regression analysis.

Predictor	Coefficient	P-value
Gender (0=female, 1=male)	0.0102	0.207
Age (30-39) (0=no, 1=yes)	-0.1202	0.000
Age (<=18) (0=no, 1=yes)	-0.0749	0.000
Age (>=40) (0=no, 1=yes)	-0.0530	0.000
User Status (organization=0, individual=1)	0.0207	0.005
Constant	0.3109	0.000
F-statistics	23.74 (p value < 0.001)	
R-squared	0.007	
Adjusted R-squared	0.007	

Table 3.1: Ordinary least squares regression using sentiment score against gender, age, and user status

We can see from the table above that the predictor variable of users' age is statistically significant since the p-value is below 0.001 and the predictor variable of users' status is also significant since the p-value is below the common alpha level of 0.05. Moreover, the regression results also show that as age increases (>18), people are significantly more positive about remote study, which is consistent with Figure 3.2.

3.2 Sentiment Analysis of Weibo data

In order to improve the accuracy, we used 3 different dictionaries to calculate the sentiment score, including Boson NLP sentiment words dictionary, a Chinese dictionary that contains words meaning "no" from National Taiwan University, and a degree word dictionary from Dalian University of Technology. The Boson NLP dictionary is extracted and trained from social media data.

Since we concatenate the score from the 3 dictionaries, the score values are pretty scattered. We dropped the first and last 5% of the total value to avoid the influence of extreme values. The maximum score we get is 21.47, the minimum score is -10.39 and the average score is 1.38. We then calculate the first and third quartile based on the average to help define the boundary values. The sentiment score (i) is then divided into 4 ranges: if the score is less than -4.02 (i.e. i < -4.02), it will be in the "Extreme Negative"; if the score is greater than or equal to -4.02 and less than 2.35 (i.e. -4.02 <= i < 2.35), it will fit in range "Negative"; if the score is between 2.35 and 8.72 (i.e. 2.35 <= i < 8.72), it will be "Neutral"; if the score is between 8.72 and 15.10 (i.e. 8.72 <= i < 15.10), the range is "Positive"; if the score is greater than 15.10 (i.e. i >= 15.10), it will be in "Extreme Positive". From figure 3.4, we can tell more than half of the posts are negative, whereas almost ¼ of people hold neutral attitudes and only approximately 10% of the posts are positive about online learning.

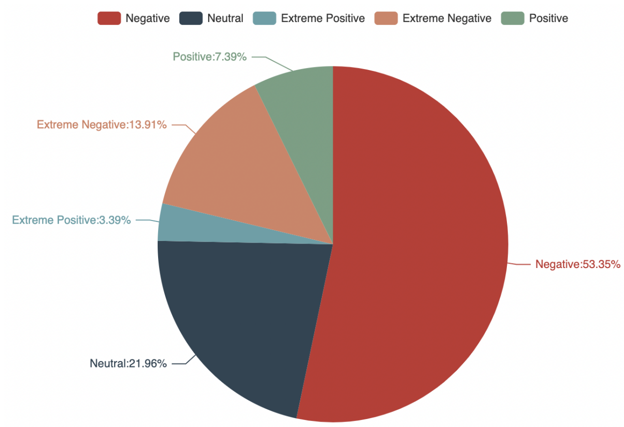

Figure 3.4: Pie charts of sentiment scores distribution

3.3 Topic analysis

Table 3.2 shows the 6 topics extracted by the LDA model from Twitter data. We listed the top keywords of each topic and tried to generalize each topic from these keywords. Figure 3.5 shows the percentage of tweets of each topic. Notice that one tweet may cover more than one topic and in this study, we chose the dominant topic of a tweet as its topic.

Topic #	Topic title	Topic keywords
1	Pandemic and schools	cases, covid, schools, pandemic, positive, coronavirus, students, university
2	Online teaching classes (e.g., zoom)	zoom, today, video, live, morning, google, students, check, class, week
3	Online certificates and webinars	free, join, webinar, certificate, course, register, virtual, training
4	Health and safety	health, safety, mental, education, kids, open, higher, spread
5	Support and help	support, help, access, need, resources, internet
6	Studying at home	kids, homeschooling, education, math, problem, energy efficiency, chance

Table 3.2: Topics and keywords of tweets using LDA

Topic 1 is about pandemic and schools, and contains keywords like "covid", "university" and "case". Topic 1 accounts for 23.1% of the total tweets. An example of such a tweet is as follows: "*Nearly ⅔ of US students are unable to read at grade level. School closures, remote learning, and other necessary COVID protocols have made a bad situation worse*."

Topic 2 (online teaching classes) and topic 3 (online certificates and webinars) both focus on the types of online education. The difference between them is that topic 2 is more about online zoom classes which were in-person courses before the covid-19 pandemic, and topic 3 is more about online certificates and webinars that are less influenced by the pandemic.

Topic 4 is about health and safety, and it accounts for about 10.4% of the total tweets. It contains keywords like "health", "mental", "kids", and "spread". An example of such a tweet is as follows: "*Thanks to unscientific school shutdowns, kids are falling back academically, slipping into loneliness-induced depression, dying by suicide at insanely high rates; are subject to a higher likelihood of domestic abuse. Lockdowns are terrible leadership.*"

Topic 5 (support and help) accounts for 13.6% of the total tweets, where people discuss the support, resources, and even difficulties of remote learning. An example of such a tweet is as follows: "*Students in Fort Worth and Dallas ISD are reportedly having some trouble connecting for the first day of online learning.*"

Topic 6 (studying at home) is about people's feelings about kids studying at home and it accounts for 10.4% of the total tweets. An example of such a tweet is as follows: "*Remote learning is making my kid hate me. Or maybe it's cuz he's a teenager now but it sucks.*"

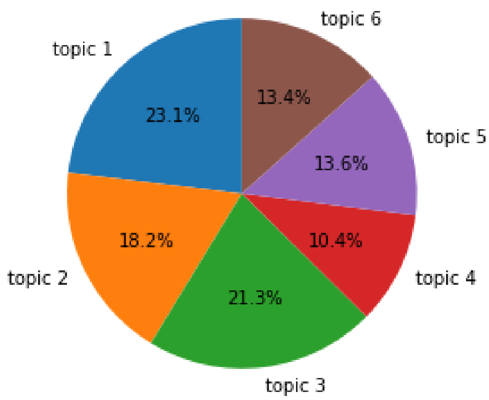

Figure 3.5: Distribution of each topic in tweets

Figure 3.6 illustrates the average sentiment score of tweets of each topic in Twitter data. The average score of all tweets we fetched is 0.239. From Figure 3.6, we see that the average sentiment scores of tweets of topic 1, 4, and 6 are under the average score. Among these topics, the first topic has the lowest score with 0.137, and we can infer that the covid pandemic has a negative impact on people's attitudes towards remote study. Also, there are more negative tweets related to people's health and safety, and studying at home.

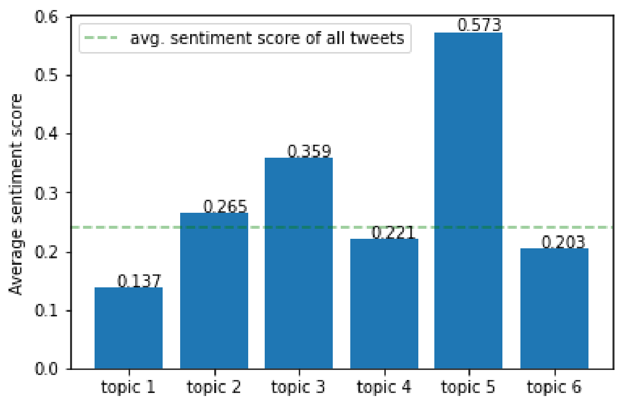

Figure 3.6: Average sentiment score of each topic in Twitter data

We also perform LDA analysis on Weibo data based on text rank keywords. Text rank is the algorithm that calculates the weight among words in a sentence and returns the most important k words to the sentence. We first perform a text rank on each Weibo post and extract 10 keywords and use those keywords to build out a dictionary that can be used later. After all, we use the LDA model to extract topics, which can be generalized into 5 categories.

Topic #	Topic	Keywords
1	Qualification Exams	教资 考研 资格证 专升本 面试 小学 粉笔
2	Schoolwork	作业 考试 复习 开学 学期 打卡 上课 课程 发现
3	Entertainment	视频 微博 直播 推荐
4	School and covid-19	军训 第一课 校园网 宿舍 开学 学校 疫情 教室
5	Teacher and students	老师 学生 孩子 教师 同学

Table 3.3: Topics and keywords extracted from Weibo

The first topic refers to qualification or certificate exams, containing keywords like "teaching certificates (教资)" and "graduate school entrance exam (考研)". And there exists a special word "粉笔" which refers to a popular training center that provides courses for teaching certificates and civil service exams in China. It is not an uncommon thing in China that people take these exams after university in order to avoid intensive competition during job hunting.

Topic 2 can be summarized as schoolwork containing words such as "homework (作业)", "exam (考试)", "review (复习)", and "courses (课程)". Similarly, topic 4 and topic 5 are close to topic 2: topic 4 has keywords such as "school (学校)", "dormitory (宿舍)", "classroom (教室)", "school begins (开学)", and "military training (军训)"; topic 5 contains keywords like "teachers (老师)" and "students (学生)". However, keywords in topic 4 tend to describe buildings or events on campus, and topic 5 is related to people in school, while topic 2 focuses more on study assignments. Additionally, topic 4 also includes "covid-19 (疫情)".

Topic 3 is unique among these topics, which has keywords such as "video (视频)", "Weibo (微博)", "streaming (直播)" and "recommendation (推荐)". We used the word entertainment to represent this topic.

The table below shows an example post as well as the expressed attitude from each topic category.

Topic	Original Posts	Meaning	Attitude
1	Day86 1.上日语课2.每日英语听力3.教资科目二网课。加油吧，考完教资就能全身心实习啦！如果我资评，教资都过了，那我岂不是很厉害，嘿嘿嘿	Describe what the user does while preparing for the teaching qualification exam and her hope for the future.	Positive
2	不愿再笑我的周末支离破碎了已经周六上午考试下午网课周日上午值班下午复习求求了赶紧五一吧	This user had a busy weekend and is urgent for holidays.	Negative
3	笑死了在看姜云升直播室友问我在听什么网课	The user thinks it is funny when she watches idol streaming but her roommate thinks she is taking an online class.	Extreme Positive
4	好烦好烦马来西亚疫情不知道何时才能结束之后又要恢复网课我的学生能学到东西才有鬼了	The user is in Malaysia and thinks students can learn nothing from online classes.	Extreme Negative
5	网课的魅力不在于课程内容，而是老师的颜值和声音	The user thinks the quality of online courses depends on teachers' personalities.	Neutral

Table 3.4: Contents and attitudes from 5 topics

3.4 Word frequency analysis

Additionally, we also extract the most frequent words from the Twitter data and Weibo data. Notice that we dropped the keywords like "remote learning", "distance learning" and "online learning" from the list of the top words since they are already the top words. We use word cloud to visually show the list. In a word cloud, the frequency of each word is represented by its size. This means that the bigger the word is, the more people use it in remote-study-related posts.

Figure 3.5 shows the word cloud of tweets. Top noun words contain "teacher", "student", "class", "work", "parents", "school", "child", and "home". Therefore, we believe that most tweets are related to teachers and students, parents and children, school and home. Pandemic-related words like "covid19" and "pandemic" also exist in the list of the top words. Moreover, we also notice that there are some positive words like "thank" and "love".

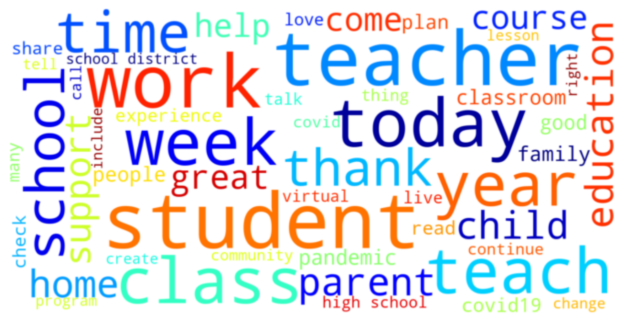

Figure 3.5: Word Cloud of tweets

Figure 3.6 shows the word cloud of Weibo posts. There are obvious education-related words such as "teacher (老师)", "school (学校)", "study (学习)", "homework (作业)", and "take classes (上课)". There are simple verbs as well: "speak (说)", "do (做)", "look (看)", "think (想)", and "feel (感觉)". We can also see emotional words "sound of crying (呜呜呜)" is in mid-size on the top left and "happy (快乐)" in a smaller size on the left side.

Figure 3.6: Word Cloud of Weibo posts

4 Conclusion and Future Work

4.1 Primary Attitudes Twitter VS. Weibo

There are significant differences in the sentiment scores distributions between Weibo and Twitter. The primary attitude among Twitter users is positive, but the main attitude from Weibo users is negative.

We believe that one of the reasons is that the United States has a longer history of online learning than China, and people get used to it. Platforms such as Udemy and Duolingo emerged in 2010, and Coursera started to provide the opportunity for students to finish bachelor's and master's degrees online ever since 2012. Last year, Edge Pathways was established in New York, which cooperates with enterprises to provide internships and study opportunities to high students who are interested in STEM courses. Edge Pathways aims to replace the traditional lecture-style college education. Education technology companies develop quickly with innovations in the United States in recent decades.

Back in China, online learning platforms did not appear until 2012 and only became popular in 2019. Besides, their aiming customers – unlike in the U.S. – are teenagers and children. However, as the policy of double reduction of students' homework and off-campus training was announced, these Ed-Tech companies gradually closed or transformed their main business. Therefore, people in the U.S. are used to the remote study style and have fewer negative opinions.

Secondly, as we mentioned before, the two countries apply different epidemic prevention policies. The Chinese government imposed a lockdown on high-risk areas to reduce the spread, and that's the main reason why students study remotely. Remote study accompanied by quarantine cannot be enjoyable: people can only stay at home and cannot go elsewhere. While in the U.S., even though schools are closed, people can still go to supermarkets and restaurants.

Despite the regional factor, even Chinese students who stayed in the U.S. may also feel depressed during remote study time. At the beginning of 2020, the medical resources were in shortage because of the increasing number of affected cases and inactions from the government. The high medical expenses and the huge number of death cases led to a huge panic among the public. The travel ban from the Chinese government and extremely high prices of airlines forced them to stay in the United States. Later, the Trump government once proposed that F-1 students who take online classes have to leave the country and cannot return in the Fall semester. Even though this proposal was rejected only after a week, the unstable political environment may also cause these negative attitudes. Although these situations were smoothed when the vaccine was invented and the policy revised, external factors can strongly affect people's feelings and opinions.

4.2 Correlation between user features and sentiment

In our study of user features, we find that females tend to be more negative about the remote study than males. Also, young people between 19 and 29 years old tend to be more negative about remote study than people in other age groups, which could be owing to the fact that most people in this age group are students and they suffer a lot from remote study during the covid-19 pandemic. They have to face computers all day and receive limited help for schoolwork. Moreover, as age increases (>18), people are significantly more positive about remote study. Additionally, sentiment scores vary in different topic categories. We find that when it comes up with the covid-19 pandemic, people tend to be more negative. Therefore, this pandemic has a negative impact on opinions of remote study.

4.3 Future work

Potential future work could include applying an algorithm to filter some ads-related tweets. In our study, we notice there are some tweets introducing online courses and certificates, which are not useful to our sentiment analysis.

In addition, more features could be included in future analyses. In our study, we only use age, gender, and user status to discover the relationship between these features and sentiment score. Other features like ethnicity, location, and family income may also be the factors to affect people's experience with remote study. We can also perform fine-grained sentiment analysis on data in order to figure out the exact emotions people hold, such as anger, sadness, or satisfaction.

5 Bibliography

[1] Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D. and Roser, M., 2021. Coronavirus Pandemic (COVID-19). [online] Our World in Data. Retrieve December 1 from: https://ourworldindata.org/coronavirus.

[2] Echo Xie. Build-up to coronavirus lockdown: inside China's decision to close Wuhan. (April 2020). Retrieved December 1 from https://sg.news.yahoo.com/build-coronavirus-lockdown-inside-china-145139273.html.

[3] 2021年就地过年. 2021. Wikipedia: The Free Encyclopedia. Retrieved from https://zh.wikipedia.org/wiki/2021年就地过年.

[4] Nambiar, D., 2020. The impact of online learning during COVID-19: students' and teachers' perspectives. *International Journal of Indian Psychology*, 8(2), pp.783-793.

[5] Hawrilenko, M., Kroshus, E., Tandon, P. and Christakis, D., 2021. The Association Between School Closures and Child Mental Health During COVID-19. *JAMA Network Open*, 4(9), p.e2124092.

[6] Wang, Y., Zhang, X. and Luo, J., 2017. When Follow is Just One Click Away: Understanding Twitter Follow Behavior in the 2016 Presidential Election. *arXiv*.

[7] Xiong, Z., Li, P., Lyu, H. and Luo, J., 2021. Social Media Opinions on Working From Home in the United States During the COVID-19 Pandemic: Observational Study. *JMIR Medical Informatics*, 9(7), p.e29195.

[8] Duong, V., Pham, P., Yang, Y., Wang, Y. and Luo, J., 2020. The Ivory Tower Lost: How College Students Respond Differently than the General Public to the COVID-19 Pandemic. *arXiv*.

[9] Yeung, N., Lai, J. and Luo, J., 2020. Face Off: Polarized Public Opinions on Personal Face Mask Usage during the COVID-19 Pandemic. *arXiv*.

[10] Hutto, C. and Gilbert, E. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*. 8, 1 (May 2014), 216-225.

[11] Zijian Wang et al. 2019. Demographic inference and representative population estimates from multilingual social media data. The World Wide Web Conference (2019). DOI: http://dx.doi.org/10.1145/3308558.3313684

[12] Elsalem, L., Al-Azzam, N., Jum'ah, A., Obeidat, N., Sindiani, A. and Kheirallah, K., 2020. Stress and behavioral changes with remote E-exams during the Covid-19 pandemic: A cross-sectional study among undergraduates of medical sciences. *Annals of Medicine and Surgery*, 60, pp.271-279.

[13] Trzcińska-Król, M., 2020. Students with special educational needs in distance learning during the COVID-19 pandemic – parents' opinions. *Interdyscyplinarne Konteksty Pedagogiki Specjalnej*, (29), pp.173-191.

[14] Morgan, H., 2020. Best Practices for Implementing Remote Learning during a Pandemic. *The Clearing House: A Journal of Educational Strategies, Issues, and Ideas*, 93(3), pp.135-141.

[15] Henderson, C., Kim, G., Zarrella, G., and Burger, D.. 2011. Discriminating Gender on Twitter. *Association for Computational Linguistics*, pp.1301–1309. URL: https://aclanthology.org/D11-1120

[16] Download our Twitter data from: https://drive.google.com/drive/folders/10ExUxl-lGcQ0d9ZUpf-SdLWeUVMI9EuB?usp=sharing

[17] Download our Weibo data from: https://drive.google.com/file/d/1FEtpXQDfILp34R20Eii6uaoORGpJsORf/view?usp=sharing