ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE

Paper ID #37322

How Do Engineering Students Characterize Their Educational Experience on a Popular Social Media Platform Before and During the Covid-19 Pandemic?

Zheng Chen

Zheng Chen(Bella) is a bachelor at the University of Hong Kong, majoring in Decision Analytics. She is interested in natural language processing, machine learning, smart city, and computer visions.

Ruiqiu Yao

Andrew Gillen

Andrew L. Gillen is an Assistant Teaching Professor in First Year Engineering at Northeastern University. Previously, he was a Lecturer in the Department of Civil, Environmental, and Geomatic Engineering at University College London. He holds a Ph.D. in Engineering Education from Virginia Tech and B.S. in Civil Engineering from Northeastern University.

How Do Engineering Students Characterize Their Educational Experience on a Popular Social Media Platform Before and During the Covid-19 Pandemic?

Keywords
Social media, sentiment analysis, topic modeling, covid-19, student evaluations of teaching

Abstract

Public social media platforms can supplement our understanding of student perceptions of engineering teaching. Looking to social media can help us build a picture of what steps we can take to improve the learning experience. It has the potential to provide meaningful information without requiring more data collection from students. This is particularly salient in times of crisis when contact with students may be inconsistent and when data such as survey results may be more challenging to obtain.

In this study, we analyzed social media data from Reddit towards developing an understanding of engineering students' attitudes and focus areas around their educational experience before and during the Covid-19 pandemic. Students' attitudes were mainly evaluated by sentiment analysis and students' focuses were explored through topic modeling techniques. Both sentiment analysis and topic modeling are a form of natural language processing. Sentiment analysis is a tool to study the feelings expressed in text while topic modeling allows us to look for groups of related phrasings and to obtain a sense of the topics being discussed. Both are readily available through open-source Python packages. Based on the sentiment analysis, findings were categorized as positive, negative, and neutral. Within these three areas, we used topic modeling to categorize and explore the different emphasis areas brought up by students (e.g., extracurricular activities, school assignments). We present the results of the modeling using a topic visualization and interpretation tool. Although this work illustrates computational methods for analyzing social media data, these tools are seen pragmatically as a means to an end and not the sole purpose of our inquiry.

Social media analyses have limitations and ethical considerations, and this work is not meant to supersede other forms of evaluation. Rather, our study explores the use of social media as a potential complementary source of data for practitioners. Our work has implications for educators and institutions looking to develop low-impact ways to evaluate educational programming in times of crisis and beyond. We hope that by presenting this work to other researchers and practitioners in engineering education, we will engage in mutually beneficial conversations around the pros and cons of using social media data and its potential applications.

1 Introduction & Background

The ongoing COVID-19 pandemic has had a huge impact on students' lives on a variety of levels, including the delivery of the school curriculum (Lyons, 2020). Most educational institutions have been delivering at least some of their teaching online, which has been new and starkly in contrast to the traditional classroom learning, requiring students to quickly adapt to this change. In addition to impact on students' learning, Lyons (2020) illustrates some other effects, such as loss of face-to-face interpersonal communication among peers, as well as financial pressure brought on by reduced part-time employment during the pandemic. These impacts can lower students' well-being and cause further psychological distress, consequently disrupting normal daily life and studies. Collecting data on the student learning experience is common in universities, and during the pandemic this became even more critical.

It is common to collect students' feedback on their educational experience (Richardson, 2005). There are four purposes for obtaining students' evaluations of their learning identified by Marsh and Dunkin (1992) including feedback on teaching effectiveness for improving teaching, administrative procedures, improving future course content, and as practice-based research. Many methods can be used for gathering feedback, but feedback questionnaires are often used (Huxham et al., 2008). However, such a commonly used method might also have some drawbacks. To begin with, the students who need to fill in the questionnaires are passively exposed to those questions set by others. This means instead of subjectively mentioning the contents they care about most, when students are filling in the questionnaires, they have to comment on the prompts provided. What's more, feedback is often not collected in a timely manner to be applied in the moment to improve the learning environment. Lamiaa (2017) argued that collecting a range of data on ongoing students' feedback in a timely fashion is better than only collecting once when the students have finished the course. Due to the lack of timeliness, students must recall their past experiences when finishing a questionnaire. Hence the accuracy and the completeness of students' feedback might be decreased. In addition, questionnaires are usually set to be mandatory evaluations, which might even have some negative impacts on teaching (Johnson, 2000).

Beyond all this, the compensation or even the position of a university teaching faculty member may be determined by their classroom teaching abilities, and the students' feedback could be one of the predominant sources to grade these skills (Marsh & Dunkin, 1992). Hence for teachers, this kind of evaluation might generate "fear, damaged relationships, and self-doubt" (Johnson, 2000, p. 433). Marsh (1984, cited in Johnson, 2000, p. 436) indicated that "mandatory evaluation of teaching using an institutionally authorized, standard procedure is perceived as a form of professional development underpinned by ideas of threat and penalty." Therefore, the questionnaire-based mandatory evaluation may have some negative impacts in terms of an instructor's professional development.

Nowadays, social media platforms, such as Twitter and Reddit, are excellent sources of social data (Tyagi, 2020). The information on social media platforms is of significant volume. Billions of individuals use social media worldwide, including 70% of the US population (Dean, 2021). Therefore, analyzing a large amount of data from social media is meaningful. Students using social media platforms could express their opinions and thoughts about the learning experience. Instead of expressing passively when filling in the questionnaires, students subjectively mention what they are concerned about through social media. Moreover, students could post their thoughts and comments whenever they choose. Therefore, the data could be gathered in a timelier manner than traditional questionnaires. Social media information has been leveraged in professional engineering environments already. It has been identified that social media platforms could benefit the communication and marketing related to engineering organizations. By leveraging social media, students and their instructors can experience new ways of “communication, interaction, and experimentation” (Palmer, 2016, p. 37).

In our study, we explored engineering students' experiences (attitudes and focus areas) during the Covid-19 pandemic by analyzing data from the social media platform, Reddit. We addressed the following research question: *How do engineering students characterize their educational experience on a popular social media platform before and during the Covid-19 pandemic?*

2 Research Design

Focusing on data from the subreddit (r/EngineeringStudents, 2021), this research aims at analyzing students' attitudes during the pandemic by assessing sentiment and focus areas of student posts. Figure 1 depicts an overview of our approach.

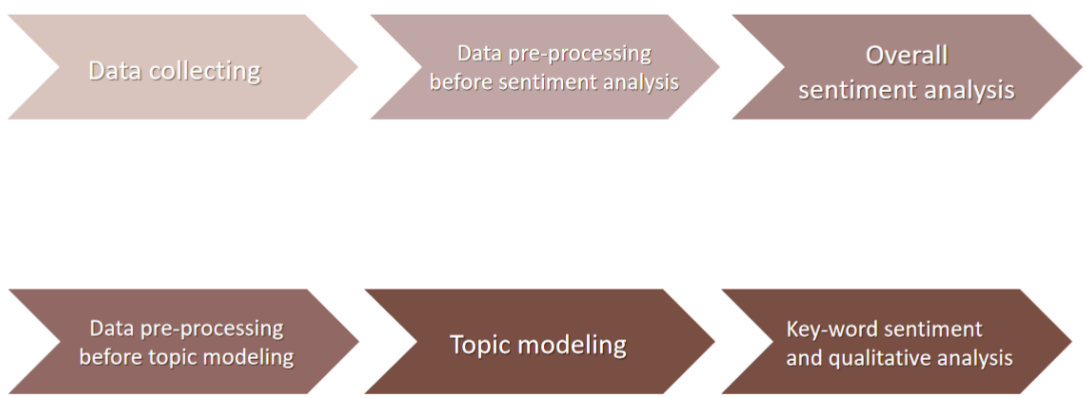


Figure 1. Overview of our research design

In the following paragraphs, we first introduce the data collection process in part 2.1, which includes the choice of social media platform, and the data scoping process. After that, in part 2.2, we briefly explain the functionality of natural language processing (NLP) tools we use for sentiment analysis and topic modeling. Based on understanding the NLP tools mentioned above, we finally explain the necessity and procedures of data pre-processing in part 2.3.

2.1 Data Collection

2.1.1 Choice of Social Media Platform

Reddit is a popular social media website, with 21 billion average screen views and 330 million active users per month. Users can post, comment, and vote on others' posts in topic-related groups called subreddits (Turcan & McKeown, 2019). In a study from *We Are Flint*, 42% of U.S. Internet users are typical college age (18 to 24) (Foundation, 2021). The platform lends itself well to researching topics related to this age demographic, which is consistent with our study. According to their guidelines, the subreddit EngineeringStudents "is a place for engineering students of any discipline to discuss study methods, get homework help, get job-search advice, and find a compassionate ear when you get a 40% on your midterm after studying all night." (r/EngineeringStudents, 2021, para.1). This subreddit started on May 4, 2011, and has 416,000 members.

We used an application programming interface (API) developed by the moderators of the r/datasets community called pushshift.io. This allowed us to access and aggregate a large volume of data in different ways (Baumgartner, 2018).

2.1.2 Scoping Data Collection

Figure 2 shows the chosen research period. We chose nine months before and during the pandemic, respectively, according to the general semester timeframe.

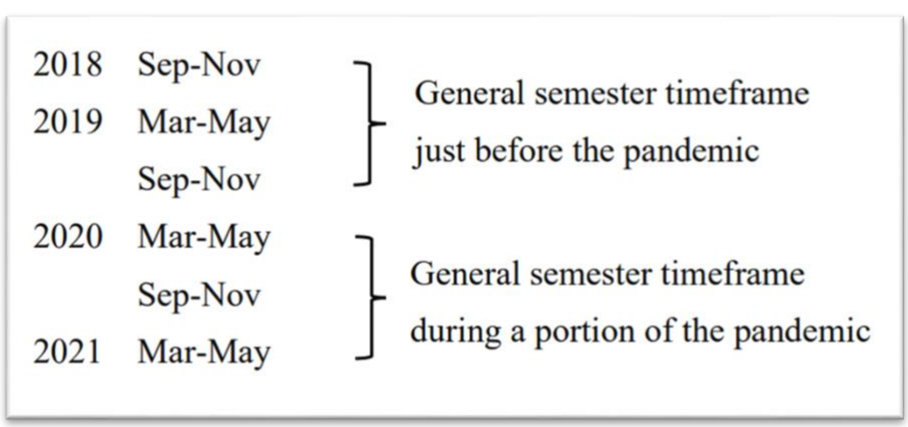


Figure 2. Chosen research period

We collected both submissions and comments in the r/EngineeringStudents subreddit. Posts in this community include submissions around the engineering academics such as study methods, homework help, and experience sharing. In other research on social media analysis, some approaches did not include comments on submissions (e.g., Turcan & McKeown, 2019; Berdanier, 2020; Kaur, 2017), while others include comments (e.g., Okon & Rachakonda, 2019). In our previous research using Twitter data, we chose not to include the comments to avoid duplicate information from retweets (Yao, 2020).

In this research, the comments were still included since they could provide essential information based on the way participants use the platform. As for the content of the comments, since many of the submissions in this subreddit are suggestion-seeking questions, the suggestions or conclusions will be shown in the comments, which are meaningful to analyze. Figure 3 shows the comment numbers for submissions in March 2021. Among the 1119 submissions posted in March 2021, the number of ranges from 0 to 493. Most submissions have comment numbers between 0 to 50, but there is one submission that gets 493 comments. In the research by Kalogeropoulos et al. (2017), submissions with more comments usually contain topics of great concern and vice versa. Therefore, in our case, the submission with 493 comments may contain more often mentioned topics than other submissions. And it is deducible that its comments will also contain keywords of these hot topics. Therefore, we feel including the comments lend validity to the analysis in general.

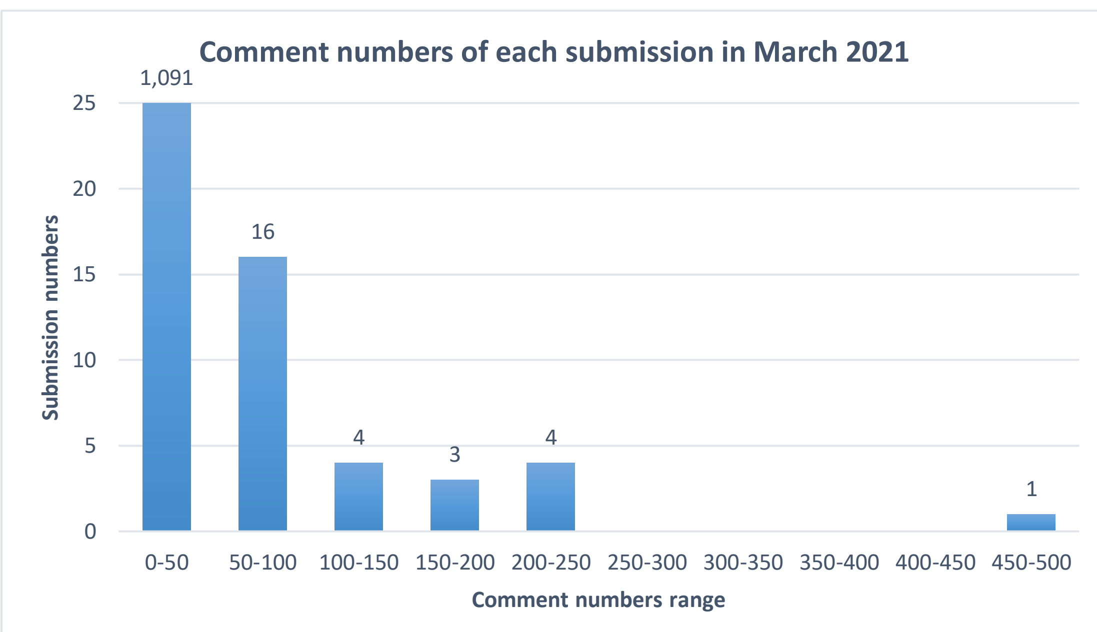


Figure 3. Comment numbers of each submission in March 2021

2.2 Natural language processing tools

2.2.1 Sentiment Analysis

The purpose of sentiment analysis (also known as opinion mining) is to identify the text contents as positive, negative, or neutral (Nasim et al., 2017). Sentiment analysis is commonly used to analyze sentiments, assessments, attitudes, and emotions (Athar, 2021). It is applied to various domains, including social media marketing, business analysis, social sciences, and politics (Nasim, 2017). There are many packages available in Python which use different methods to do sentiment analysis. In this research, we used the packages Textblob and VADER.

Textblob assigns a value for polarity (a scale for negative to positive) and subjectivity. Subjective sentences generally refer to opinion, emotion, or judgment (Athar, 2021). Rufai and Bunce (2020), in their sentiment analysis around Covid-19, found that approximate objective posts account for 64%. Using this as a guide, we considered only the top 36% of the subjectivity scale (100%-64%=36%) in our research.

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a sentiment analysis tool for social media (Stevenson et al., 2018). VADER Sentiment Analysis is suitable for texts from social media since it can adequately handle a lot of slang, emoticons, and special expressions from social media (Hutto & Gilbert, 2014). According to the updated documentation (Hutto, 2021), examples are shown in Figure 4.

- conventional use of **punctuation** to signal increased sentiment intensity (e.g., "Good!!!")
- conventional use of **word-shape** to signal emphasis (e.g., using ALL CAPS for words/phrases)
- using **degree modifiers** to alter sentiment intensity (e.g., intensity *boosters* such as "very" and intensity *dampeners* such as "kind of")
- understanding many **sentiment-laden slang** words (e.g., 'sux')
- understanding many sentiment-laden **slang words as modifiers** such as 'uber' or 'friggin' or 'kinda'
- understanding many sentiment-laden **emoticons** such as :) and :D
- translating **utf-8 encoded emojis** such as 💘 and 💋 and 😁
- understanding sentiment-laden **initialisms and acronyms** (for example: 'lol')

Figure 4. VADER examples

Utilizing the aforementioned two tools, we first used Textblob to analyze the subjectivity of all posts and label them as subjective, neutral, and objective. We then used VADER to analyze the non-objective posts (subjective and neutral posts), and labeled them as positive, neutral, and negative.

2.2.2 Topic Modeling

The output of topic modeling is a set of topics; each topic consists of clusters of words (Jacobi et al., 2015). By calculating the occurrence and distribution of words in a collection of documents, the Python library Gensim identifies potential patterns of words and generates the output, which helps people learn the subjects of documents (Rehurek & Sojka, 2010).

The topic numbers of different documents could be adjusted by programmers. To find the most suitable models, we need to refer to the evaluation indexes. According to research from Röder et al. (2015), among the seven most used evaluation methods based on statistical models, the CV coherence score performs best on sentence-level texts. The higher the coherence score is, the better the model suits the documents. CV is "a new combination found by a systematic study of the configuration space of coherence measures" (Röder et al., 2015, p. 400). We referred to the CV coherence score in this research to determine the most suitable topic numbers.

According to Onah and Pang (2021, p. 6), “PyLDAvis is a web-based interactive visualization python package that allows the display of the topics that were identified using the LDA approach.” PyLDAvis presents an overview of the topics that have been generated depending on how they differ from one another. As shown in Figure 5, different circles represent different topics, the greater the distance between circles, the greater the difference in the topics. Our research used this visualization to double-check whether the model decided by the CV coherence score identified topics with suitable distances.

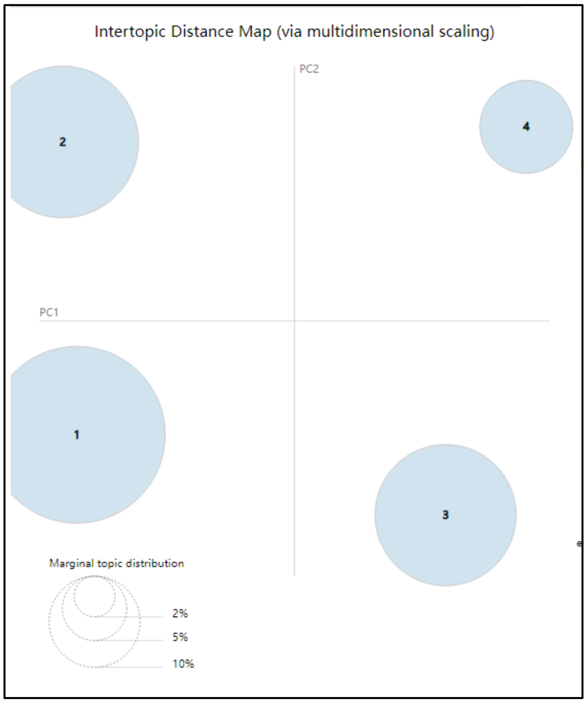


Figure 5. pyLDAvis example

2.3 Data pre-processing

Student feedback data can be considered unstructured text (Nasim et al., 2017). To extract the meaningful information and delete the useless text, several pre-processing steps are applied using Python's Natural Language Toolkit (NLTK) (Bird, 2009). Considering the different functionality of Textblob, VADER, and Gensim, we applied data pre-processing twice, before sentiment analysis and before topic modeling respectively. Before sentiment analysis, 1st pre-processing included dropping all empty texts, removing the URL and noise words, and tokenizing passages into sentences. Before topic modeling, 2nd pre-processing procedures included removing all stop words and converting the texts into lowercases.

2.3.1 Before sentiment analysis (1^{st} pre-processing)

We removed URLs and noise words. URLs include strings containing "http", "www", "com", "net", "org", etc. Noise words include meaningless characters and punctuations that especially appear in our database, for example, "\u", "\n", "/r", "/EngineeringStudents". However, according to the updated VADER documentation, “VADER can now understand many sentiment-laden emoticons such as :) and :D” (Hutto, 2021, p. 3). Therefore, we avoided removing independent punctuations shown in Figure 6, "list of most common emoticons" (Anastasia, 2016), such as ":" and "=", to improve the accuracy of sentiment scores.

Figure 6. List of most common emoticons

We then tokenized texts into sentences. VADER took mostly the sentence-level samples to train the algorithm model. Meanwhile, in our dataset, many posts contained more than 5 sentences. Keeping the whole submission in this length could hide some sentiment information, which can be misleading when we conduct sentiment analysis on specific words.

For example, we reviewed the following submission posted on May 19, 2020. The sentiment score for the whole submission is 0.975; however, we could see the content is generally negative.

> *Those who are moving to remote learning, who else is worried about learning material through PowerPoint and on your own instead of in-class face-to-face?*
>
> *I don't want to [expletive] on other majors because I know every major is hard, but stem/science type majors aren't really things (I think) can be learned online at least very easy. I know business and communication type majors will probably have it easier but I'm pretty worried as I like to ask questions in class and I am just all around more motivated to learn the difficult material when seeing it worked out in the classroom.*

The rest of my semester is online and I'm honestly not sure how well I'm going to perform going forward. If anyone has any tips or tricks for taking online classes or for staying motivated while taking them please share! I'm sure I'm not the only one in this boat. (Posted on r/EngineeringStudents)

This inaccuracy is probably because of the use of positive words, such as “easy”, when describing “business and communication type majors”. It might also be due to VADER’s weak ability of scoring paragraph-level posts. However, if we tokenize this submission into sentence-level (i.e., cut the submission posts into sentences), and then score each sentence, the scores become more accurate. Therefore, we decided to tokenize all posts into sentences and then conduct sentiment analysis for each sentence. Lastly, we dropped all empty excerpts for texts without posts and comments. “Removed” and “deleted” posts and comments were also dropped.

2.3.2 Before topic modeling (2nd pre-processing)

We removed stop words that have no lexical meaning and only serve to connect words in a phrase grammatically (Gustafsson, 2020). Including those words in the topic modeling process will result in meaningless results. For instance, "I" will become a frequently used word, and those results will mislead the topic interpretation. Figure 7 shows the 179 English stop words contained in the NLTK library.

```
print(stopwords.words('english'))
```

```
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', "you're", "you've", "you'll", "you'd", 'your', 'yours', 'yourself',
'yourselves', 'he', 'him', 'his', 'himself', 'she', "she's", 'her', 'hers', 'herself', 'it', "it's", 'its', 'itself', 'they', 'them', 'th
eir', 'theirs', 'themselves', 'what', 'which', 'who', 'whom', 'this', 'that', "that'll", 'these', 'those', 'am', 'is', 'are', 'was', 'wer
e', 'be', 'been', 'being', 'have', 'has', 'had', 'having', 'do', 'does', 'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or', 'bec
ause', 'as', 'until', 'while', 'of', 'at', 'by', 'for', 'with', 'about', 'against', 'between', 'into', 'through', 'during', 'before', 'af
ter', 'above', 'below', 'to', 'from', 'up', 'down', 'in', 'out', 'on', 'off', 'over', 'under', 'again', 'further', 'then', 'once', 'her
e', 'there', 'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few', 'more', 'most', 'other', 'some', 'such', 'no', 'nor', 'n
ot', 'only', 'own', 'same', 'so', 'than', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don', "don't", 'should', "should've", 'now',
'd', 'll', 'm', 'o', 're', 've', 'y', 'ain', 'aren', "aren't", 'couldn', "couldn't", 'didn', "didn't", 'doesn', "doesn't", 'hadn', "had
n't", 'hasn', "hasn't", 'haven', "haven't", 'isn', "isn't", 'ma', 'mightn', "mightn't", 'mustn', "mustn't", 'needn', "needn't", 'shan',
"shan't", 'shouldn', "shouldn't", 'wasn', "wasn't", 'weren', "weren't", 'won', "won't", 'wouldn', "wouldn't"]
```

Figure 7. NLTK library stop words

We then converted the texts into lowercases. In order to properly count the word frequency and further classify topics, we need to recognize “COVID”, “Covid,” and “covid” as the same word. Therefore, we used the NLTK to lowercase all the text.

2.3.3 Why not merge these two procedures?

Stop words disturb topic modeling but seldom disturb sentiment analysis; some even help VADER score better. They include, words expressing sentiment degree, such as "few", "more", "some", "such", "only", "further"; and words expressing negative attitudes, such as "don't", "shouldn't", "needn't", etc. Therefore, we decided not to remove them before sentiment analysis.

Same words in different types of characters (e.g., "COVID" and "covid") disturb topic modeling, but this is important information for sentiment analysis. As shown in VADER's updated documentation (Hutto, 2021), VADER could distinguish and judge the "conventional use of word-shape to signal emphasis (e.g., using ALL CAPS for words/phrases)." Therefore, we did not lowercase texts before sentiment analysis.

3 Results and Discussion

Results of data collection, data pre-processing, and overall sentiment analysis are shown in part 3.1. The results of topic modeling for a different group of words are shown in part 3.2. Finally, the keyword sentiment analysis and qualitative analysis results are shown in 3.3.

3.1 Data Collection, Pre-processing, and Sentiment Analysis Results

We analyzed data from the subreddit r/EngineeringStudents of the chosen research period. Figure 8 shows the amount of data collected. In total, 73,691 submissions and 130,286 comments were collected. After data cleaning, there were 203,931 posts, within which 103,716 posts were generated before the pandemic and 100,215 posts were generated during the pandemic. We tokenized them into 251,697 sentences, with 120,477 sentences before the pandemic and 131,210 sentences during the pandemic. Within all those sentences, 90,611 sentences are non-objective (based on Textblob subjectivity analysis).

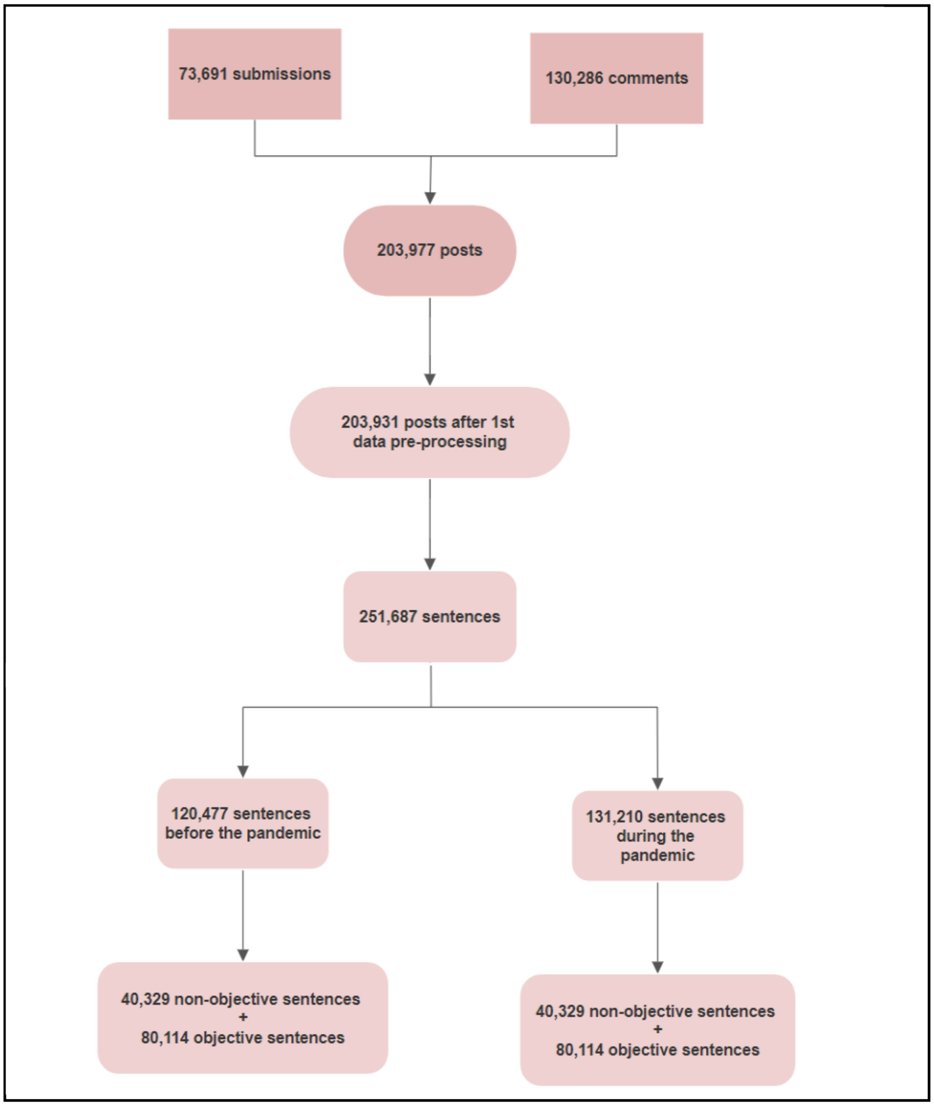


Figure 8. Amount of data

As for the sentiment analysis results, using VADER to evaluate the non-objective posts, the polarity type percentages of 18 months on average were: 57.05% for positive, 23.69% for negative, and 19.22% for neutral. Figure 9 shows the month's sentiment behavior and amount of all non-objective sentences. Each bar represents the sentiment behavior of one month in this bar chart, corresponding to the left ordinate. The red, grey, and blue parts represent the proportion of positive, neutral, and negative attitudes, respectively. For example, in March 2020, positive attitude accounts for 57.10%; and in October 2020, neutral attitude accounts for 26.00%; and in March 2021, negative accounts for 25.90%. The yellow line in this chart shows the number of subjective sentences each month, corresponding to the right ordinate. For instance, the number of September 2020 total subjective sentences is 5,447.

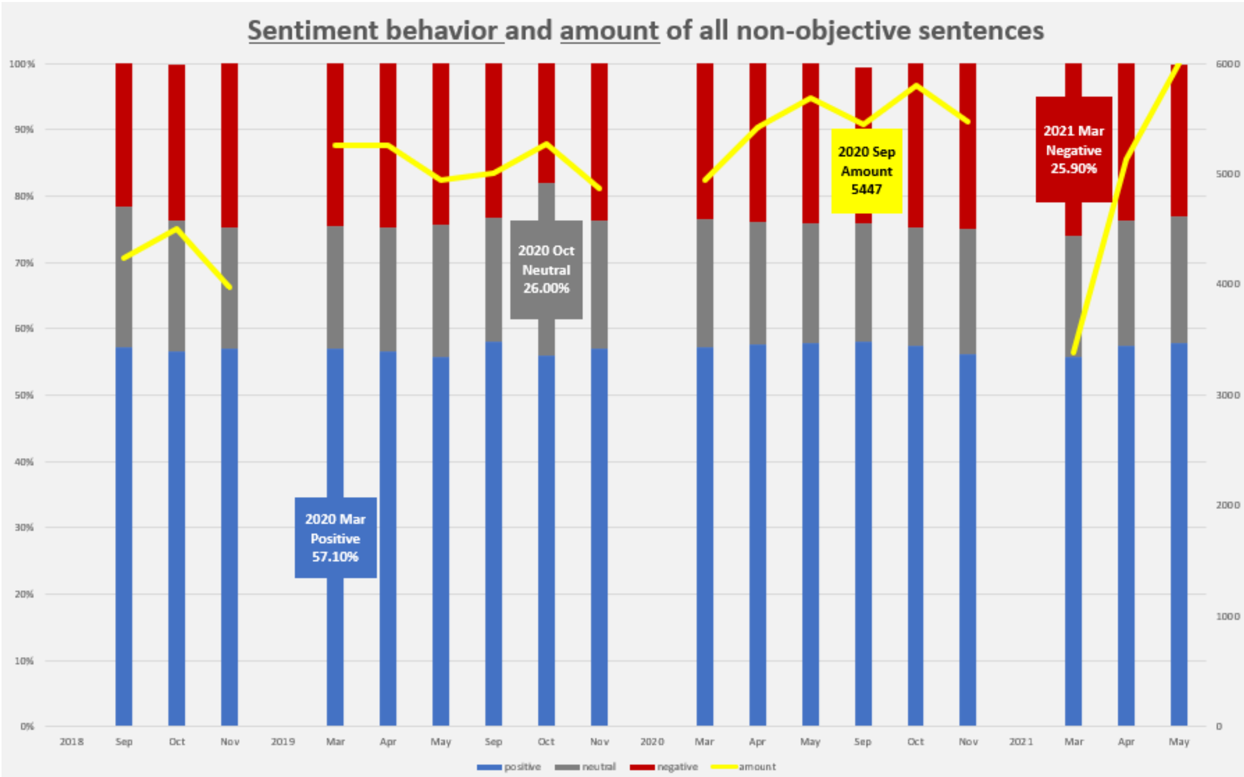


Figure 9. Sentiment analysis results by month.

Despite the pandemic, the sentiment behaviors stayed almost the same among those 18 months.

3.2 Topic Modeling Results

We used all sentences generated before and during the pandemic (120,477 and 131,210 sentences, respectively) to do the 2nd data pre-processing. Table 1 shows the results of our sample numbers after the 2nd data pre-processing.

	Positive	Neutral & negative	Objective	Total
Before the pandemic	23,601	15,634	79,879	119,114
During the pandemic	25,212	18,772	78,208	122,192
Total	48,813	34,406	158,087	241,312

Table 1. The sample numbers after 2nd data pre-processing

In order to figure out the different focus of sentences with a different sentiment, we conducted topic modeling on positive sentences, negative sentences, neutral sentences, and objective sentences, respectively. We used the CV coherence score to evaluate the model and further decide the number of topics. However, since the data amounts of negative and neutral documents are all quite less, the CV coherence scores of models with possible topic numbers are all relatively low (lower than 0.2) when compared with other documents, which means no model can fit the documents well. Therefore, we merged the negative and neutral documents to conduct topic modeling. Finally, we have six documents in total to conduct the topic modelling.

In the following parts, 3.2.1 shows the results of topic numbers for each document, based on the CV coherence scores. 3.2.2 to 3.2.4 show the results of topic modeling for positive documents, neutral and negative documents, and objective documents, of the two time periods, respectively.

3.2.1 Topic numbers

According to Ostrowski (2015), the appropriate number of topics should be more than 3; therefore, we calculated the CV coherence scores of models with more than or equal to 3 topics. We found that when the topic numbers were larger than 7, the CV coherence scores for all the documents were relatively low (lower than 0.5). Based on this, we did not train the models with topic numbers of more than 7.

Table 2 shows the CV coherence scores for all documents with 3 to 7 topics.

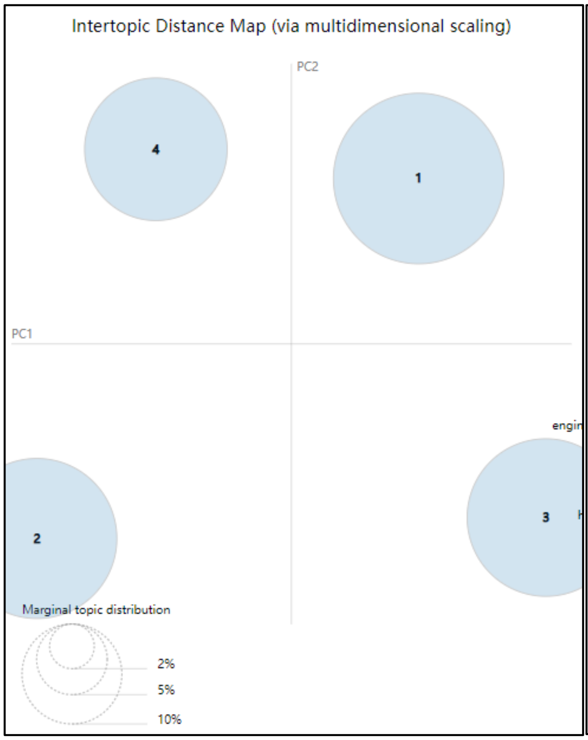
CV coherence score	Before the pandemic			During the pandemic		
	Positive	Neutral & negative	Objective	Positive	Neutral & negative	Objective
3	0.279	**0.322**	0.201	0.491	**0.518**	**0.637**
4	**0.302**	0.314	**0.254**	**0.519**	0.503	0.609
5	0.280	0.304	0.223	0.501	0.482	0.554
6	0.246	0.238	0.242	0.506	0.475	0.529
7	0.239	0.246	0.214	0.474	0.495	0.585

Table 2. CV coherence scores

By comparing all the CV coherence scores, we determined the final topic numbers as follows: positive documents (both four topics), neutral and negative documents (both three topics), and objective documents (four topics for the “before” datasets and three topics for the “during” datasets).

3.2.2 Topic modeling for positive sentiment

Figure 10 depicts a global view of the topic model for positive posts before the pandemic (left) and during the pandemic (right). The closer the topics are, the more relevant the contents will be. In both the two documents, these four topics all have little relationship, which implies that the model fits well for these datasets. The area of the topic circles represents the percentage of those topics in the document, and it sorts the topics in decreasing order of prevalence.



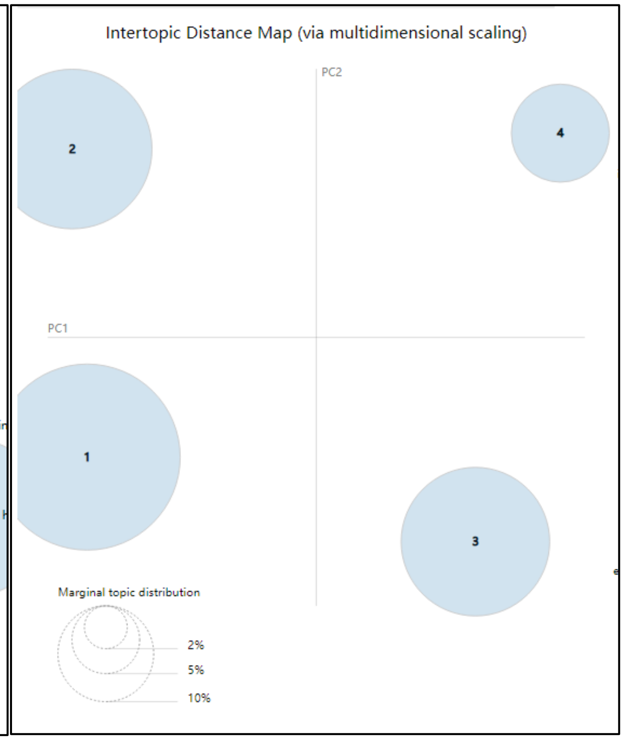


Figure 10. Positive documents' topic modelling visualization of before the pandemic (left) and during the pandemic (right)

Table 3 provides a breakdown of topics and our interpretation for the posts categorized as positive. The table is sorted in descending order according to the topic percentage of the "during the pandemic" document, and topics common to both documents are placed on the same line.

Before the pandemic		During the pandemic		Interpretation
Percentage	Words	Percentage	Words	
29.1%	class, learn, study, exam, semester, year	37.7%	class, school, student, grade, study, learn, friend, exam	Daily engineering study
25.6%	program, career, internship, offer, apply, university, recruiter, graduate, master, hire	27.8%	work, company, internship, experience, apply, interview, degree, program, resume, position, offer	Extracurricular experiences and plans for working or further study
20.4%	math, physics, video, course, easy, static, note	24.1%	book, professor, math, textbook, teach, equation, material	Specific engineering -relevant courses/ subjects
24.8%	course, homework help, question, submission, wiki, chegg, example, coder, ipad	10.5%	survey, mate, brand, strain, laptop, calculator, circuit, solution, method	Specific engineering -relevant questions/ answers

Table 3. Topics for positive documents

Overall speaking, both these two documents contain the same four topics, only the respective percentages vary a little. Before the pandemic, the four topics have smaller percentage gaps; while during the pandemic, the proportion of "daily engineering study" increased a little, and "specific engineering-relevant questions/answers" decreased a lot.

3.2.3 Topic modeling for neutral and negative sentiment

Figure 11 depicts a global view of the topic model for neutral and negative posts before the pandemic (left) and during the pandemic (right). Again, the closer the topics are, the more relevant the contents will be, and the area of the topic circles represents the percentage of those topics. Here in both documents, their three topics have quite a large distance, implying the topic numbers are suitable for these two documents.

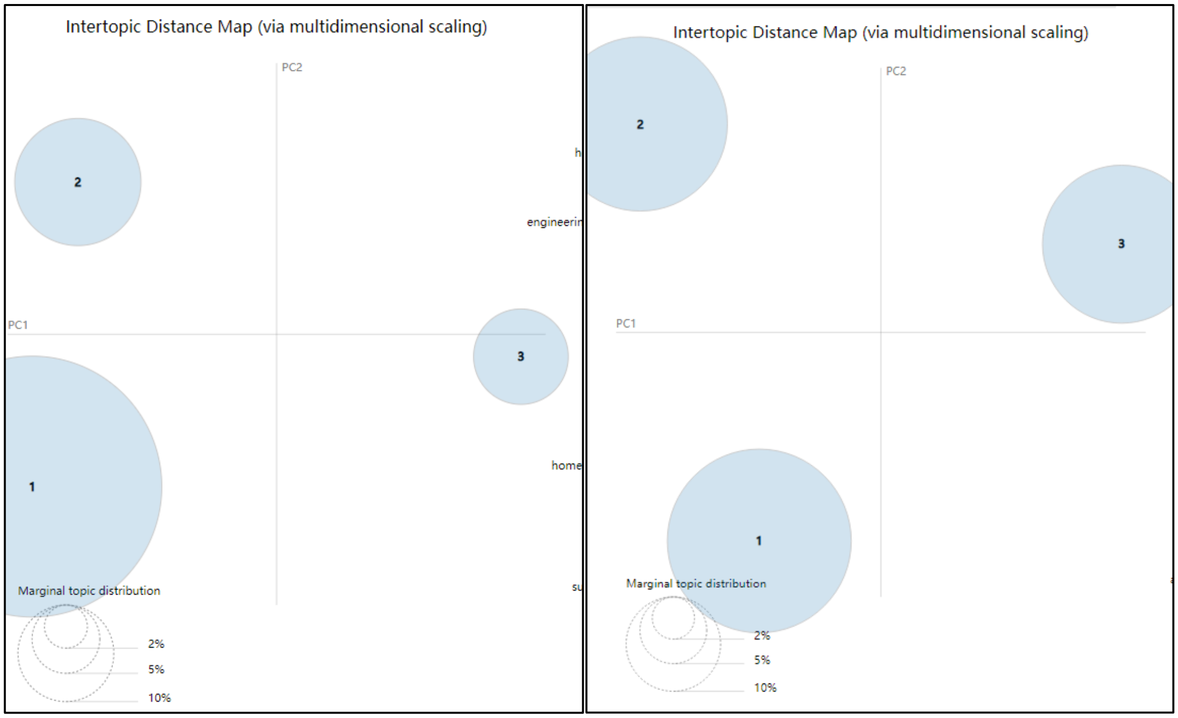


Figure 11. Neutral and negative documents' topic modelling visualization of before the pandemic (left) and during the pandemic (right)

Table 4 provides a breakdown of topics and our interpretation of the topic for the posts categorized as neutral or negative. The table is sorted in descending order according to the topic percentage of the "during the pandemic" document, and topics common to both documents are placed on the same line.

Before the pandemic		During the pandemic		Interpretation
Percentage	words	Percentage	Words	
		37.9%	hard, year, job, bad, [expletive], hate	Direct sentiment expressions.
64.1%	class, exam, problem, learn, study, school, degree	34.1%	online, course, exam, grade, fail, hard, problem, test, year	Daily engineering study, but the during the pandemic document mentions more about online.
21.6%	job, company, guideline, major, apply, program			Extracurricular experiences and plans for working or further study
14.3%	course, submission, math, course, explain, number	28.0%	equation, answer, calculate, Matlab, velocity, pressure,	Specific engineering-relevant questions/ answers

Table 4. Topics for neutral and negative documents

It's worth mentioning that topics that appear most frequently in the second document do not appear in the first document, which is the "direct sentiment expressions". Meanwhile, "extracurricular experiences and plans for working or further study" only frequently appeared in the "before the pandemic" period, but not in the "during" period. Although the "daily engineering study" is frequently mentioned in both periods, the "during" documents mentions more about "online".

3.2.4 Topic modeling for objective sentiment

Figure 12 depicts a global view of the topic model for objective posts. This time, topic 1 and 2 overlap each other a little bit, however, using other topic numbers could only make the model worse, so we kept topic numbers as 3.

Figure 12. Objective documents' topic modelling visualization of before the pandemic (left) and during the pandemic (right)

Table 5 provides a breakdown of topics and our interpretation of the topic for the posts categorized as objective. The table is sorted in descending order according to the topic percentage of the "during the pandemic" document, and topics common to both documents are placed on the same line.

Before the pandemic		During the pandemic		Interpretation
Percentage	Words	Percentage	Words	
30.20%	offer, career, research, program	50.9%	work, year, school, job, internship, experience, company, time	Extracurricular experiences and plans for working or further study
29.53%	exam, problem, study, class	41.3%	exam, semester, course, work, class, learn, study.	Daily engineering study
21.75%	homeworkhelp, submission, project, question, chegg, wiki	7.8%	homework, concern, perform, message, wiki, faq, submission, meme	Specific engineering-relevant questions/ answers
18.52%	math, space, variable, video, note			Specific engineering -relevant courses/ subjects

Table 5. Topics for objective documents

Overall, the topics contained in both documents are similar. The topics in “during the pandemic” document are all covered in the “before the pandemic” document, while the latter include one more about the “specific engineering-relevant courses/subjects”.

3.2.5 Overall topic modeling results

Figure 13 shows the five topics appeared across all six documents. The topic variation is not huge across these six documents, which makes sense since all the posts are related to one unique topic, engineering study, whose subtopics do not vary much.

The topic "extracurricular experiences and plans for working or further study" appears the most in the objective document, then appears majorly in the positive document (for more than a quarter). Before the pandemic, 21.4% of the neutral and negative words related to this topic. However, during the pandemic, this topic is not obviously reflected in the neutral and negative words; instead, more words tend to directly express their negative sentiments. Therefore, it is deducible that even though students' negative sentiments were expressed more directly, these negative emotions were not strongly relevant to "extracurricular experiences and plans for working or further study".

To investigate what content these directly expressed negative sentiments are related to, we mainly focused on the neutral and negative documents during the pandemic. The topic “daily engineering study” appears in all these documents with quite a high percentage (at least for 29%). But only the negative document frequently mentions more about “online.” Meanwhile, "online" is closely related to the new teaching model under the pandemic. Therefore, it encouraged us to conduct sentiment analysis and qualitative analysis on posts containing "online" in the research period (18 months). The results are shown in part 3.3.

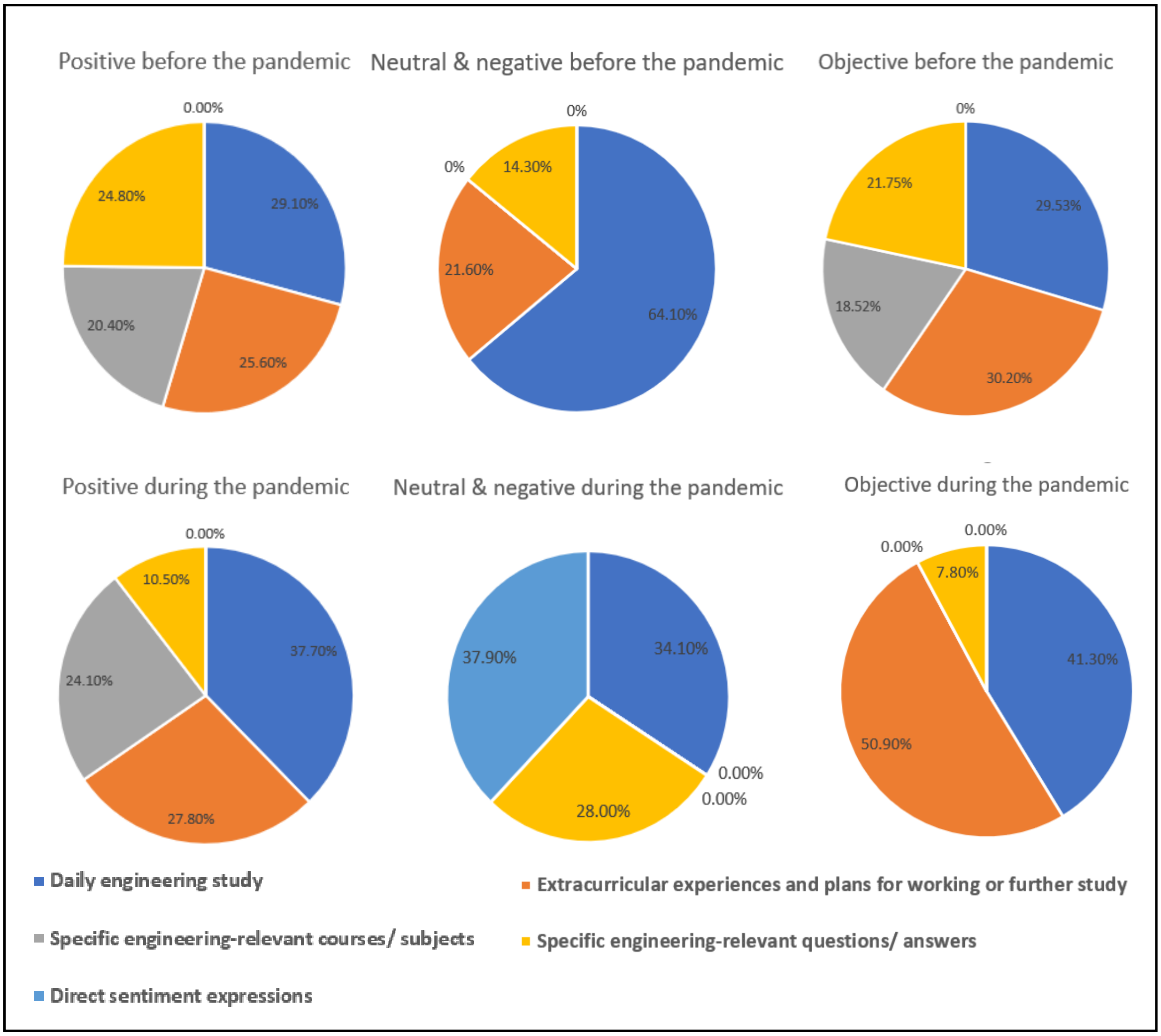


Figure 13. Topic proportions in the six documents

3.3 Key-word Sentiment and Qualitative Analysis Results

3.3.1 Sentiment Analysis Results

Using the 90,611 non-objective sentences (as mentioned in 3.1), we filtered the sentences containing the word "online" and conducted sentiment analysis by month using VADER. The results are shown in Figure 14. Each bar represents the sentiment behavior of one month, corresponding to the left ordinate. The red, grey, and blue parts represent the proportion of positive, neutral, and negative attitudes, respectively.

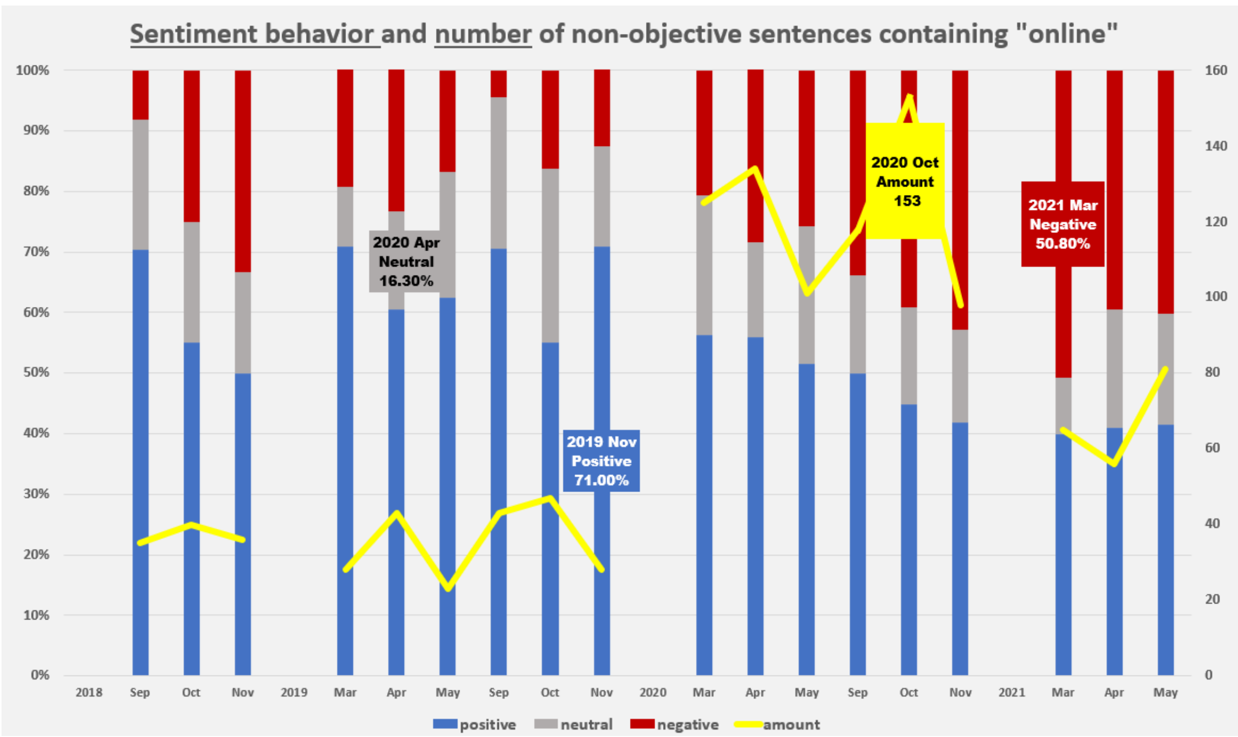


Figure 14. Sentiment behavior and number of non-objective sentences containing "online"

Although compared with the whole dataset the sentence numbers for each month are relatively low (within the range of 25 and 153), we could still observe some obvious tendency from Figure 13.

Firstly, the sentence numbers increase after the outbreak of Covid-19. The numbers before Covid-19 (in the first nine chosen months) fluctuate around 30. However, it keeps to more than 100 for the whole year 2020, even reaching the peak of 153 in October 2020. Although in 2021, the number drops back to around 70, it still doubles the data before the pandemic. We infer that engineering students talk more about "online" after the outbreak of the pandemic. This is the probable result of the new teaching model taken by universities during the Covid-19 pandemic.

In terms of the sentiment behavior, overall, the sentiment behavior worsens during the pandemic. We conclude that the positive proportion decreases, and the negative proportion increases in general before and during the pandemic. Referring to the data, the average positive proportions before the pandemic are 58.43% and 65.10% respectively for 2018 and 2019, however, it decreases to 50.07% in 2020, and further to 40.80% in 2021. More obviously, the negative proportion increases from 22.13% and 15.47% in 2018 and 2019 to 31.75% in 2020 and further to 43.47% in 2021, which is more than double the negative proportion in 2018. To further investigate the reason for this sentiment deterioration, we read most of the original posts, especially those with extreme positive or negative sentiment grades, and summarized the following contents.

3.3.2 Qualitative Analysis Results

Contents summary of the positive posts:

1. Creative online activities promotion: Promotion of online note-taking tools, study resources platform, etc. Promotion of competitions held online, for example, Online International Science Engagement Challenge, e-commerce business start-up competition, etc. Promotion of online socializing, online parties, etc.

2. Being able to save study time and adjust study pace: Students could speed up the study videos, and skip or retake some parts, especially for their own study needs. Students could work as interns and take courses simultaneously since the time arrangements are more flexible. Could save the time of commuting.

3. Feel relaxed since universities provide a pass/fail policy.

Contents summary of the negative posts:

1. Study quality drops: Universities do not update the arrangements timely and effectively, which leads to many inconveniences. Students are more likely to miss online assignments and other study requirements. It's hard for students to focus on home. Professors are concerned about course quality since students tend not to show up in courses.

2. Communication quality drops. Students have many questions to ask, but sometimes it's harder online. Students cannot communicate effectively with peers, and more "free-riders" appear during group work.

3. Examination cheating. Both students and professors complain about students cheating during

online exams.

4. Doubt the choice of study disciplines (especially for first-year students). Students cannot practice the skills they learned online, and it's harder for them to choose future study disciplines.

5. Ruin the internships, research programs, socializing, travels, and other plans. Students complain about losing their work-life balance.

6. Job opportunities decrease. Students complain about the difficulty of finding an internship or a job.

4 Conclusions

Based on the sentiment analysis, most of the students still characterize their engineering experiences positively, despite the pandemic. And the overall sentiment percentage did not change meaningfully during the 18 tested months.

Based on the topic modeling, the topic of documents diverse little; they contain topics related to daily engineering study, extracurricular experiences or plans, and specific engineering courses or questions. However, the neutral and negative documents during the pandemic does not contain much content related to "extracurricular experiences or plans." The words "online" only frequently appear in the neutral and negative documents during the pandemic.

Based on the sentiment and qualitative analysis for sentences containing "online," we discovered after the outbreak of Covid-19, the amount of content related to "online" increased, and their overall sentiment scores decreased. In further qualitative work, we summarized the commonly mentioned aspects in positive and negative documents.

We've found some corresponding conclusions in other papers. According to Pelargos et al. (2020), during the pandemic, most North American neurosurgical residents "were not concerned that the changes would have long-term negative effects on their overall education or future career prospects" (p. 386). This could be the reason why students' sentiment behavior did not change a lot during the pandemic period. What's more, according to the results of a topic modeling on tweets, Mujahid et al. (2021) found that students' main worries are a lack of technical skills and network issues when taking online education during COVID-19. We also reached similar conclusions after conducting the qualitative analysis, and we pointed out some other major concerns, such as concerns about examination cheating.

5 Limitations and Future Works

In this research, the data amount (251,697 sentences) is not that large compared to similar works, therefore some results might not be accurate enough (for example, the sentiment behavior of sentences containing "online"). In future studies using Reddit, there are other engineering study-related subreddits, such as AskEngineers, AskScience, Education, Engineering, and HigherEducation. It is possible to use keyword searching functions in those subreddits to get more meaningful information on this topic.

Additionally, for sentiment analysis, we use VADER, which is an unsupervised method. The unsupervised method does not require a database to be annotated with proper sentiment labels, while supervised machine learning requires a labeled dataset of text documents to train classifiers (Nasim et al., 2017). Therefore, the supervised analyzer might perform more accurately to analyze sentiment polarity and subjectivity.

What's more, in topic modeling testing, we choose the topic numbers by referring to the CV coherence score. However, the CV coherence score is just one of the evaluation indexes, and more evaluation methods could be considered in future studies.

Ultimately, we explored this research method because it is hard to know how to help students during the pandemic with traditional methods and evaluations. This research provides a preliminary view into the potential for utilizing social media data during a crisis.

References

Baumgartner, J. M. (2018). *Pushshift Reddit API v4.0 documentation*. Pushshift Reddit API v4.0 Documentation - Pushshift 4.0 documentation. Retrieved July 3, 2021, from http://reddit-api.readthedocs.io/en/latest/

Berdanier, C. G., Whitehair, C., Kirn, A., & Satterfield, D. (2019). Analysis of social media forums to elicit narratives of graduate engineering student attrition. *Journal of Engineering Education*, *109*(1), 125–147. https://doi.org/10.1002/jee.20299

Dean, B. (2021, June 10). *How many people use Social Media in 2021? (65+ statistics)*. Backlinko. Retrieved July 3, 2021, from https://backlinko.com/social-media-users

Giachanou, A., & Crestani, F. (2016). Like it or not. *ACM Computing Surveys*, *49*(2), 1–41. https://doi.org/10.1145/2938640

Gustafsson, M. (2020). Sentiment Analysis for Tweets in Swedish : Using a sentiment lexicon with syntactic rules (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-91832

Hutto, C. J. (2021). *Welcome to Vadersentiment's documentation!*. Welcome to VaderSentiment's documentation! - VaderSentiment 3.3.1 documentation. Retrieved January 10, 2022, from https://VADERsentiment.readthedocs.io/en/latest/

Hutto, C., & 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. Retrieved July 3, 2021, from https://ojs.aaai.org/index.php/ICWSM/article/view/14550

Huxham, M., Laybourn, P., Cairncross, S., Gray, M., Brown, N., Goldfinch, J., & Earl, S. (2008). Collecting student feedback: A comparison of questionnaire and other methods. *Assessment & Evaluation in Higher Education*, *33*(6), 675–686. https://doi.org/10.1080/02602930701773000

Jacobi, C., van Atteveldt, W., & Welbers, K. (2015). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89–106. https://doi.org/10.1080/21670811.2015.1093271

Johnson, R. (2000). The Authority of the Student Evaluation Questionnaire. *Teaching in Higher Education*, *5*(4), 419–434. https://doi.org/10.1080/713699176

Kalogeropoulos, A., Negredo, S., Picone, I., & Nielsen, R. K. (2017). Who shares and comments on news?: A cross-national comparative analysis of online and social media participation. *Social Media + Society*, *3*(4), 205630511773575. https://doi.org/10.1177/2056305117735754

Lyons, Z., Wilcox, H., Leung, L., & Dearsley, O. (2020). Covid-19 and the mental well-being of Australian medical students: Impact, concerns and coping strategies used. *Australasian Psychiatry*, *28*(6), 649–652. https://doi.org/10.1177/1039856220947945

Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. *Higher education: Handbook of theory and research, 8,* 143-233.

Mujahid, M., Lee, E., Rustam, F., Washington, P. B., Ullah, S., Reshi, A. A., & Ashraf, I. (2021). Sentiment analysis and topic modeling on tweets about online education during COVID-19. *Applied Sciences, 11*(18), 8438. https://doi.org/10.3390/app11188438

Nasim, Z., Rajput, Q., & Haider, S. (2017). Sentiment analysis of student feedback using machine learning and lexicon based approaches. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*. https://doi.org/10.1109/icriis.2017.8002475

Onah, D., & Pang, E. (2021). MOOC design principles: Topic modelling-pyldavis visualization & summarization of learners' engagement. *EDULEARN Proceedings*. https://doi.org/10.21125/edulearn.2021.0282

Ostrowski, D. A. (2015). Using latent Dirichlet allocation for topic modelling in Twitter. *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015)*. https://doi.org/10.1109/icosc.2015.7050858

Palmer, S. (2016, July 18). *The Representation of Engineering Education as a social media topic on Twitter*. PAEE/ALE 2016 : Proceedings of the 8th International Symposium on Project Approaches in Engineering Education and 14th Active Learning in Engineering Education Combined Conference and Workshop. Retrieved July 10, 2021, from http://hdl.handle.net/10536/DRO/DU:30084976

Pelargos, P. E., Chakraborty, A., Zhao, Y. D., Smith, Z. A., Dunn, I. F., & Bauer, A. M. (2020). An evaluation of neurosurgical resident education and sentiment during the coronavirus disease 2019 pandemic: A north american survey. *World Neurosurgery, 140*. https://doi.org/10.1016/j.wneu.2020.05.263

R/engineeringstudents. reddit. (n.d.). Retrieved June 14, 2021, from https://www.reddit.com/r/EngineeringStudents/

Richardson, J. T. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment & Evaluation in Higher Education*, *30*(4), 387–415. https://doi.org/10.1080/02602930500099193

Rufai, S. R., & Bunce, C. (2020). World leaders’ usage of Twitter in response to the COVID-19 pandemic: A content analysis. *Journal of Public Health*, *42*(3), 510–516. https://doi.org/10.1093/pubmed/fdaa049

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. https://doi.org/10.1145/2684822.2685324

Stevenson, K., Kregor, J., Gomez, B., & Kelly, J. S. (2018). Grassroots market research on grass: Predicting cannabis brand performance using social media scraping. *Proceedings of the 2nd International Conference on Advanced Research Methods and Analytics (CARMA 2018)*. https://doi.org/10.4995/carma2018.2018.8348

Team, F. (2022, January 19). *Reddit statistics for 2021: Eye-opening usage & traffic data*. Foundation Marketing. Retrieved July 3, 2021, from https://foundationinc.co/lab/reddit-statistics/

Turcan, E., & McKeown, K. (2019). Dreaddit: A reddit dataset for stress analysis in Social Media. *Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019)*. https://doi.org/10.18653/v1/d19-6213

Tyagi, A., Field, A., Lathwal, P., Tsvetkov, Y., & Carley, K. M. (2020). A computational analysis of polarization on Indian and Pakistani social media. *Lecture Notes in Computer Science*, 364–379. https://doi.org/10.1007/978-3-030-60975-7_27

White, B. (2020, May 27). *Sentiment analysis: Vader or textblob?* Medium. Retrieved June 15, 2021, from https://towardsdatascience.com/sentiment-analysis-vader-or-textblob-ff25514ac540

Yao, R. (2020). *Investigating public opinions of civil infrastructure projects using social media analysis: A case study of High Speed 2 (HS2) Twitter data* [Unpublished master's dissertation]. University College London.

Youssef, L. S. (2017). What can regular and timely student feedback tell us about the teaching and learning processes? *Reflective Practice*, *18*(6), 750–771. https://doi.org/10.1080/14623943.2017.1351350