

Assessing COVID-19 Impacts on College Students via Automated Processing of Free-form Text

Ravi Sharma¹, Sri Divya Pagadala², Pratool Bharti², Sriram Chellappan¹,
Trine Schmidt³ and Raj Goyal³

¹*Department of Computer Science and Engineering, University of South Florida, Tampa, FL, U.S.A.*

²*Department of Computer Science, Northern Illinois University, DeKalb, IL, U.S.A.*

³*Ajivar LLC, Tarpon Springs, FL, U.S.A.*

Keywords: COVID-19, Emotion, Mental Health, Natural Language Processing, Semantic Search.

Abstract: In this paper, we report experimental results on assessing the impact of COVID-19 on college students by processing free-form texts generated by them. By free-form texts, we mean textual entries posted by college students (enrolled in a four year US college) via an app specifically designed to assess and improve their mental health. Using a dataset comprising of more than 9000 textual entries from 1451 students collected over four months (split between pre and post COVID-19), and established NLP techniques, a) we assess how topics of most interest to student change between pre and post COVID-19, and b) we assess the sentiments that students exhibit in each topic between pre and post COVID-19. Our analysis reveals that topics like Education became noticeably less important to students post COVID-19, while Health became much more trending. We also found that across all topics, negative sentiment among students post COVID-19 was much higher compared to pre-COVID-19. We expect our study to have an impact on policy-makers in higher education across several spectra, including college administrators, teachers, parents, and mental health counselors.

1 INTRODUCTION

The COVID-19 pandemic has impacted people across the globe irrespective of where they live, how old they are, and what they do. From a physical health perspective, COVID-19 has affected senior citizens the most. However, from a mental health perspective, younger people (especially college students) have been severely impacted due to many factors including transitioning to new modes of instruction, loss of friend circles, cancellation of classes and even entire semesters, financial hardships, shrinking job markets, failing relationships and much more. Also, since mental health services are harder to access now (due to social distancing necessities), students are suffering even more from this pandemic.

On a related note, various stakeholders in higher education are also impacted because there is a need to manage new forms of online instruction, understand and react to the financial hardships of students, ensure classrooms are socially isolated, generate facilities for COVID-19 testing, and more. In overcoming these challenges, there needs to be a mechanism to understand student needs and expectations in real-time, as

students are the most important, while also being vulnerable entities in the higher education ecosystem.

One approach to get student perceptions in such times is to send periodic surveys to them. But these can be sent only a limited number of times to prevent students from simply choosing not to respond. Survey responses also suffer from well-known recency biases (Chiesi and Primi, 2009). Furthermore, surveys cannot proactively capture student perceptions and are sensitive more to the mindset of those creating the surveys, rather than students themselves.

The next viable option to glean student perceptions is mining social media. This is an area rich in exploration today, and there are many studies on mining social media content to understand student mental health needs (Bagroy et al., 2017) (Tsakalidis et al., 2018), relationship tendencies (Sin et al., 2019), academic performance (Paul et al., 2012), political participation (Yang and Lee, 2020), (Kushin and Yamamoto, 2010) and much more. But again, the core issue with mining social media data in the context of our problem in this paper stems from a) social media data being noisy, meaning that there is so much irrelevant information out there, which makes it much

harder to filter out and extract only ones relevant to any particular scenario; b) the fact that there are so many social media sites that are popular today, and so it becomes increasingly harder to make conclusions by processing data from only one or few of them. Additionally, these sites are very hard for developers to crawl massive data due to privacy concerns.

In this paper, we report the results of an experiment we conducted at a four-year college in the US to assess topics and sentiments of students before the COVID-19 pandemic began and after, via processing free-form texts collected via a smartphone app specifically designed for students to express their moods, feelings, opinions via the app. With this dataset, we wanted to answer, broadly speaking, two questions presented below:

1. Before COVID-19, what are the topics that students most posted about, and how did the topics change post-COVID-19?
2. Across these topics and timelines, how did the sentiment of students change between positive, neutral, and negative?

To answer the above questions, we process our dataset using several standard and well-established techniques in Natural Language Processing. Our analysis that compares the post-COVID-19 timeframe with the pre-COVID-19 timeframe revealed Health became a highly trending topic for our subjects which increased sharply post COVID-19 timeframe while Housing, Finance, and Relationships experiencing a marginal increase. Also, Education decreased significantly during the post-COVID-19 timeframe. From the perspective of emotion, universally across each topic, there was a significant increase in negative sentiment in the post-COVID-19 timeframe.

To the best of our knowledge, this study is unique in processing free-form text generated via a custom made app (that was designed to glean mental health insights of college students) to infer trending topics and sentiments in the COVID-19 timeframe. Towards the end of the paper, we present how our work enables policy-makers in higher education to cater to student mental health and other needs in these unique times. Additionally, the data will also help the app developers to optimize the AI to deliver individualized resources that help the specific student in their time of need based on their sentiment, etc.

The rest of the paper is organized as follows. Section 2 presents related work in the literature. Section 3 explains our dataset. Section 4 details our underlying research methodology. Section 5 presents our results. Finally, Section 6 presents a related discussion, conclusions, and future work.

2 RELATED WORK

Since December 2019, when the COVID-19 pandemic began, there has been an urgent interest across the globe to understand its impact from many perspectives. In particular, the genetic aspects of the disease, the epidemiological aspects, the impact of contact tracing, the impact on economies, and political aspects across the globe are all actively being investigated today. In this section though, we highlight current research in the realm of COVID-19 that is most relevant to our work in this paper, namely the processing of texts over the Internet concerning the pandemic.

In (Jelodar et al., 2020), researchers applied topic modeling on COVID-19-related Reddit comments to extract meaningful categories from public opinions. They then designed an LSTM recurrent neural network (Hochreiter and Schmidhuber, 1997) for polarity and sentiment classification. They used *sentiStrength* (Thelwall et al., 2012) for sentiment detection and average out the sentiment of COVID-19 related comments obtained from 10 subreddits forums. They labeled the data as per *sentiStrength* scores and achieved an 81.15% classification accuracy. In (Oyebode et al., 2020), comment data obtained from social media is processed and then context-aware themes are extracted. For each theme, the sentiment is extracted using VADER lexicon-based algorithm (Gilbert and Hutto, 2014) and only themes having positive or negative scores are kept, neutral themes are discarded. Later, these themes are categorized into 34 negative and 20 positive broader themes using human reviewers. These themes span different areas related to health, psychological, and social issues and showed the impact of COVID-19 negatively. The authors also recommend possible solutions for negative themes. In another paper (Kabir and Madria, 2020), the authors also mine Twitter to use collected tweets as a repository to analyze information shared by the people to visualize topics of interest, and the emotions of common citizens in the USA as the pandemic is unfolding.

Our work in this paper broadly falls in the above categories. However, there are compelling differences. First, in existing research on processing social media data, the trend is to ignore the demographics of subjects, since demographic information is very hard to find from social media data, and even if a person does post demographic information, there is little evidence to check if the posted information is correct. So, if the goal is to understand the topic and emotions of a particular demographic only, then mining social media data can be problematic. In this paper,

Table 1: Data Cleaning Process.

Students' Texts	Removal of Stopwords and Generic Terms	Final Text After Lemmatization
Freaking out my grades ever since online learning.	['Freaking', 'grades', 'online', 'learning']	['freak', 'grade', 'online', 'learning']
Im constantly thinking about domestic abuse victims who have to stay in their homes now	['constantly', 'thinking', 'domestic', 'abuse', 'victims', 'stay', 'homes']	['constantly', 'think', 'domestic', 'abuse', 'victim', 'stay', 'home']
My sleep schedule is off track because im inside all day.	['sleep', 'schedule', 'track', 'inside']	['sleep', 'schedule', 'track', 'inside']

Table 2: Top 20 Frequent Words for Pre and Post COVID-19.

Pre-COVID-19				Post-COVID-19			
Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
class	90	roommate	71	quarantine	277	covid	210
happy	70	family	69	hope	147	home	145
learn	62	school	56	family	145	pandemic	128
positive	53	college	52	miss	126	class	125
mind	50	stress	49	school	116	virus	105
parent	49	enjoy	48	happy	100	worry	99
live	41	attract	41	sick	99	parent	98
success	40	money	40	sad	96	stay	92
relationship	39	healthy	39	house	84	learn	81
create	37	boyfriend	37	understand	80	job	78

Table 3: Topics and Vocabulary for categorization and Semantic Similarity.

Topics	Vocabulary
Education	grade, lecture, professor, exam, attendance, homework, quiz, assignment, syllabus, gpa, marks, college, school, study,
Finance	tuition, job, money, rent, debt, grocery, fees, wage, dollars, shopping, car, budget, scholarship, cash, salary, bills, economy, poor, rich, income
Health	health, disease, illness, nutrition, vaccine, medicine, diagnosis, sickness, eat, food, disability, hospital, exercise, gym, covid, pandemic, quarantine, lockdown, virus, corona
Family	mom, dad, grandma, brother, sister, grandparents, aunt, uncle, sibling, son, daughter, kid, wife, relative, family, cousin, stepfather, stepmother, husband, home, homesick, fiance
Relationships	boyfriend, girlfriend, bf, gf, crush, friend, buddy, dude, ex, stranger, people, partner, associate, pal, mate
Housing	hostel, apartment, roommate, house, dorm, cloths, room, bed, window, bathroom, fan, sound

tic textual similarity over recent years. For instance, robust, semantic embeddings for two sentences like “We are very happy now.” and “Our group is enjoying presently.” will be very similar, even though they contain a completely different set of words. Needless to say, such, robust embedding is critical for our problem, since two free-form texts from students may have words completely different from each other, but still could refer to the same topic, for example, sentences like “Have a homework to submit tonight.” and “Working on an assignment due later today”, despite having no common words must have similar semantic embeddings since both refer to the topic of Education.

How to do this effectively is our challenge.

For this challenge, we use SBERT that improves upon BERT and RoBERTa, by optimizing semantic textual similarity using sentence-level embedding from siamese and triplet network structure (Schroff et al., 2015). This allows the model to derive fixed-size vectors for input sentences, and are much more efficient to scale. In our study, posts corresponding to pre and post COVID-19, along with the vocabulary for each Topic identified in Table 3 was fed into the SBERT network to derive sentence-embeddings. Once we derive semantic vector embeddings, we compare pre and post text embeddings with embed-

dings generated from words for each Topic in Table 3 using the notion of Cosine similarity (Li and Han, 2013). The output of this step essentially gives a notion of similarity between the two embeddings compared. For every text, the topic which it belongs to was identified as the one for whose embedding, the similarity with the embedding of the corresponding text was identified as the highest.

4.2 Sentiment Analysis of Topics Pre and Post COVID-19

We now are ready to detail our method to glean the sentiment of a post both pre and post COVID-19. For a comprehensive analysis, multiple states of art NLP models were used for validity analysis, namely, VADER and TextBlob:

4.2.1 VADER

Valence Aware Dictionary and sEntiment Reasoner (VADER) (Gilbert and Hutto, 2014) is a lexicon and rule-based sentiment analysis tool. It does not require any training data and is constructed from a generalized, valence-based human-curated gold standard sentiment. It is fast and gives positive, negative, neutral and an overall sentiment score. It is especially effective for more casual posts (as in free-form text) and is especially useful for mining sentiments of social media posts, YouTube comments, etc.

4.2.2 TextBlob

TextBlob (Loria, 2018) is a Python-based open-source library and returns polarity and subjectivity of a text input. The polarity is in the float range $[-1, +1]$ with -1 being highly negative and +1 being highly positive sentiment. Subjectivity lies in the range $[0, 1]$ where 0 being objective and 1 being subjective. The way it works is that TextBlob goes along finding words and phrases it can assign polarity and subjectivity to based on predefined rules, and it averages them all together for longer text. It is especially useful for more formal texts.

5 RESULTS

In this section, we discuss the results of topic distribution identification and sentiment analysis in the pre vs post-COVID-19 dataset.

5.1 Topic Categorization and Distribution

Fig. 4 presents our results of topics of importance to students before and after COVID-19. We see from the figure that there is a significant shift in the importance of these topics before and after COVID-19. We see that “Education” had a steep drop as a topic in students’ posts after COVID-19 compared to before COVID-19. After COVID-19, it had the second-lowest importance (after Finance) among students. The “Health” topic has the biggest jump among student concerns in post-COVID-19 timeframe compared to the pre-COVID-19 timeframe. There was not much change in the “Family” topic, but we do see an increase in “Housing” and a decrease in “Relationships” topics in the post-COVID-19 timeframe compared to the pre-COVID-19 timeframe.

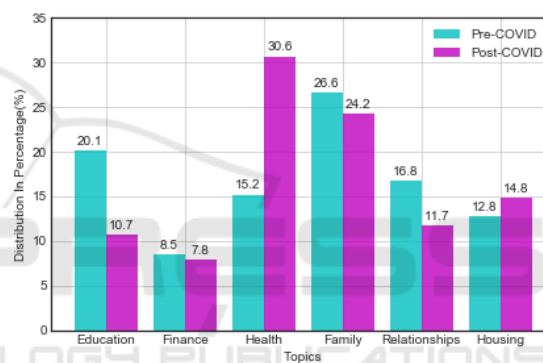


Figure 4: Topic Distribution Comparison Between Pre and Post COVID-19.

5.2 Sentiment Analysis among Topics

While the topics of interest to students are important for decision-makers in academic environments, it is important to also glean what is the sentiment expressed by students on these topics, and how they change over time. For each of the two sentiment analysis models we utilized, the corresponding sentiments for the topics are presented in Fig. 5. The results from left to right are for VADER and TextBlob. At the top, is the pre-COVID-19 timeframe, and at the bottom is the post-COVID-19 timeframe. In general, overall sentiment values saw a decrease in positive sentiment and an increase in negative sentiment.

6 DISCUSSION & CONCLUSIONS

COVID-19 is still adversely affecting the lives of college students, in the form of numerous uncertainties

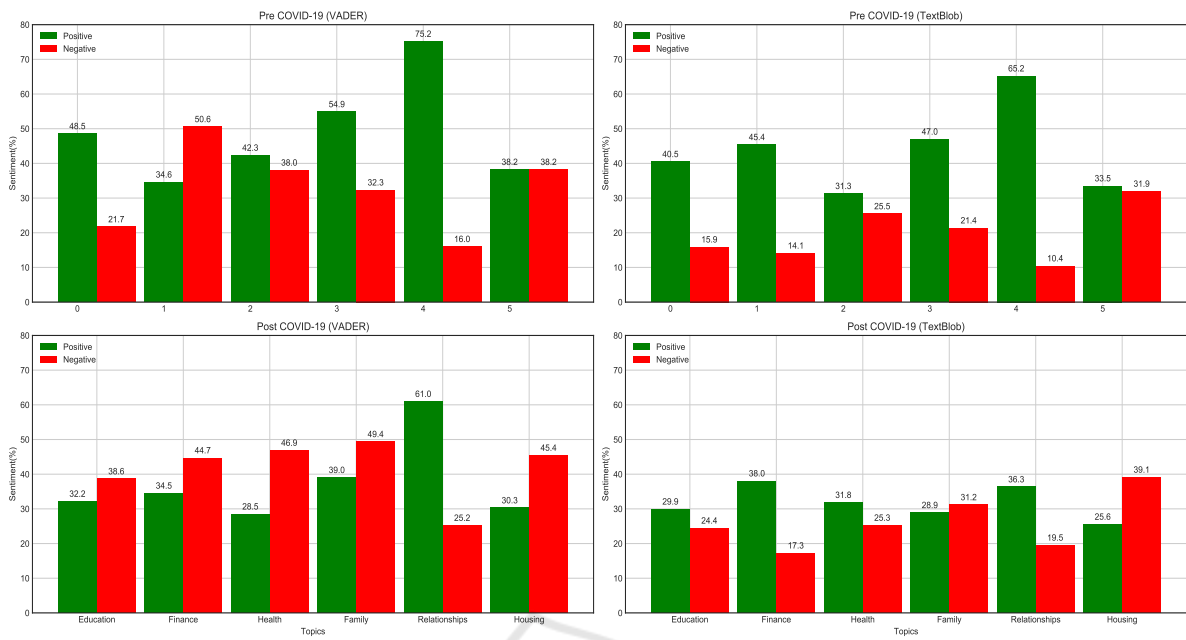


Figure 5: Sentiment across Topics Pre and Post COVID-19.

from the perspective of health, finance, career, relationships, and more, that has taken a toll on their mental health. It is important to understand their issues at the micro and macro levels to provide an effective and scalable solution. Our paper makes important early contributions to this space in terms of gleaning insights on student needs from a custom app, designed and deployed to detect and address student mental health concerns. For instance, our finding that “Education” is not a topic of critical interest to students posts COVID-19 should raise concerns on administrators to ensure that students are motivated to learn and that the quality of instruction must be given priority. Our findings related to how topics of interest changed post-pandemic are also important for administrators to know, to channel their resources in the appropriate directions.

Also, the majority of declination of categories from pre to post-covid-19 such as “Education”, “Relationships”, and “Family” is due to the sharp increase in the “Health” category which further emphasizes the fact that medical concerns and other related topics were affecting the conversations of the students. Although our findings are critical to understanding the college students’ areas of concern, it is important to mention the limitations and underline assumptions of our study.

Firstly, all the participants in the study belong to the same university and hence can have similarities in their context to some degree. This also limits the study to issues related to geographical demographics of one type. Secondly, the free-texts provided by the

participant were not categorized by the age groups and were merged, which may overlook the issues specific to graduate vs undergraduate students. Next, the population considered in our study were overlapped for the two-time frames meaning pre and post-COVID-19 data can be from the same or different students. This restricts us from conducting any statistical studies on the samples as the style and structure of the texts can vary from person to person and for a generalized distribution for categories and sentiments, clear separation between users is needed. We would like to mitigate these issues by collecting more data across universities and demographics.

In conclusion, authorities in universities have outreach programs, counselors, and other traditional measures to ease out the mental distress among students. While these are adaptable for the academic persona in general, it can be ineffective with scaling student population, lack of counselors, hesitant target subjects, etc. An effective measure to merge with the traditional methods is to curb these problems and automate the process. Methods to provide students with comfortable sharing and categorize them in defined classes for easy management can not only solve scaling issues of growing student mental health cases but also tackles the lack of counselors’ problem. Though our study is not exhaustive to all students and should not be taken as a systematic assessment, we believe that can be a preliminary step towards providing insight to policymakers, stakeholders, and administrative figures to develop strategies for students’ needs during the pandemic.

REFERENCES

- Ajivar (2019). Life coach powered by artificial intelligence, <https://ajivar.com>.
- AlKandari, N. Y. (2020). Students anxiety experiences in higher education institutions. In *Anxiety Disorders*. IntechOpen.
- Bagroy, S., Kumaraguru, P., and De Choudhury, M. (2017). 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.
- Brooker, A., Brooker, S., and Lawrence, J. (2017). First year students' perceptions of their difficulties. *Student Success*, 8(1):49–63.
- Cer, D., Diab, M., Agirre, E., Lopez-Gazpio, I., and Specia, L. (2017). Semeval-2017 task 1: Semantic textual similarity-multilingual and cross-lingual focused evaluation. *arXiv preprint arXiv:1708.00055*.
- Chiesi, F. and Primi, C. (2009). Recency effects in primary-age children and college students. *International Electronic Journal of Mathematics Education*, 4(3):259–279.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Doygun, O. and Gulec, S. (2012). The problems faced by university students and proposals for solution. *Procedia-Social and Behavioral Sciences*, 47:1115–1123.
- Gilbert, C. and Hutto, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Available at (20/04/16) <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>, volume 81, page 82.
- Head, A. and Eisenberg, M. (2011). How college students use the web to conduct everyday life research. *First Monday*, 16(4).
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Jelodar, H., Wang, Y., Orji, R., and Huang, H. (2020). Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *arXiv preprint arXiv:2004.11695*.
- Kabir, M. and Madria, S. (2020). Coronavis: A real-time covid-19 tweets analyzer. *arXiv preprint arXiv:2004.13932*.
- Kushin, M. J. and Yamamoto, M. (2010). Did social media really matter? college students' use of online media and political decision making in the 2008 election. *Mass Communication and Society*, 13(5):608–630.
- Lai, G., Xie, Q., Liu, H., Yang, Y., and Hovy, E. (2017). Race: Large-scale reading comprehension dataset from examinations. *arXiv preprint arXiv:1704.04683*.
- Li, B. and Han, L. (2013). Distance weighted cosine similarity measure for text classification. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 611–618. Springer.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). Roberta: A robustly optimized bert pre-training approach. *arXiv preprint arXiv:1907.11692*.
- Loper, E. and Bird, S. (2002). Nltk: the natural language toolkit. *arXiv preprint cs/0205028*.
- Loria, S. (2018). textblob documentation. *Release 0.15*, 2.
- Oyebode, O., Ndulue, C., Adib, A., Mulchandani, D., Suruliraj, B., Orji, F. A., Chambers, C., Meier, S., and Orji, R. (2020). Health, psychosocial, and social issues emanating from covid-19 pandemic based on social media comments using natural language processing. *arXiv preprint arXiv:2007.12144*.
- Paul, J. A., Baker, H. M., and Cochran, J. D. (2012). Effect of online social networking on student academic performance. *Computers in Human Behavior*, 28(6):2117–2127.
- Reimers, N. and Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.
- Schroff, F., Kalenichenko, D., and Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823.
- Sin, M., Pyeon, H., Kim, H., and Moon, J. (2019). Interpersonal relationship, body image, academic achievement according to sns use time of college students. *The Journal of the Convergence on Culture Technology*, 5(1):257–264.
- Thelwall, M., Buckley, K., and Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173.
- Tsakalidis, A., Liakata, M., Damoulas, T., and Cristea, A. I. (2018). Can we assess mental health through social media and smart devices? addressing bias in methodology and evaluation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 407–423. Springer.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.
- Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., and Bowman, S. R. (2018). Glue: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint arXiv:1804.07461*.
- Williams, A., Nangia, N., and Bowman, S. R. (2017). A broad-coverage challenge corpus for sentence understanding through inference. *arXiv preprint arXiv:1704.05426*.
- Yang, C.-c. and Lee, Y. (2020). Interactants and activities on facebook, instagram, and twitter: Associations between social media use and social adjustment to college. *Applied Developmental Science*, 24(1):62–78.