Comparison of Students' Learning and Engagement in Teaching an Online Course Before and After COVID-19 Measured Through Emails Using Machine Learning and Deep Learning

Muhammad Abusaqer
Department of Computer Science
North Dakota State University
Fargo, ND, USA
muhammad.abusaqer@ndsu.edu

Kenneth Magel
Department of Computer Science
North Dakota State University
Fargo, ND, USA
kenneth.magel@ndsu.edu

ABSTRACT
Emails provide valuable insights into student learning and engagement in online courses. In this study, we apply machine learning and deep learning techniques to analyze the sentiment of emails sent by students to the course instructor in an introductory computer skills course at a medium-sized U.S. university during Spring 2020. By comparing email sentiments before and after the outbreak of COVID-19, we aim to quantify the pandemic's impact on student perceptions. Sentiment analysis is conducted using the BERT neural network, and topic detection is performed using Hugging Face Transformers. The results reveal a slight decline in average sentiment after COVID-19, suggesting increased uncertainty among students. This study is the first to compare learning engagement in an already online course across the COVID-19 transition.

Keywords: Online Learning, Email Analysis, Deep Learning, Sentiment Analysis, COVID-19

INTRODUCTION
The rapid onset of the COVID-19 pandemic forced universities worldwide to transition rapidly to online learning. Although many studies have examined the shift from traditional to online instruction, little attention has been paid to the impact of the pandemic on courses that were already conducted online [1] [2] [3] [4] [5] [6]. This study investigates student sentiment and engagement in an online introductory computer skills course by analyzing email communications sent to the course instructor. We compare data from two periods: before COVID-19 (January 15–February 15, 2020) and after its spread (March 25–April 25, 2020)

BACKGROUND
We collected the email communications between the instructor and students in an online course. The course was offered at a medium-sized US University in the spring of 2020. Students' enrollment in the class was about 300. The course is an introductory course to computer skills and can be taken almost by all majors at the university.

LITERATURE REVIEW
Mansoor and his colleague reported the sentiment analysis over time for people's tweets related to COVID-19 accompanied by exploration data analysis [7]. The authors also measured the sentiment of the tweets dataset regarding how people perceive the new life requirements of Working from Home (WFH) and Online Learning. To do the sentiment analysis classification, the authors used machine learning models of Long Short-term Memory (LSTM) and Artificial Neural Networks (ANN).

In [8], Desai, Ramasmy, and Kiper demonstrated a prototype for a system to grade students' involvement in online discussion forums in a course. They used machine learning and deep learning algorithms to obtain text classification/category, keyword extraction, and sentiment analysis to analyze posted text on the course online discussion board. The author's ultimate goal is to have an integrated tool into the Learning Management System (LMS) of Canvas perform systematic data collection and analysis to improve automatic grading of students' participation on an online discussion board.

Vargo and his colleagues in [9] examined the various digital tools people use during the epidemic of COVID-19 to avoid direct contact and subsequently to minimize the outbreak of the virus. The authors found that Zoom, Facetime, and WhatsApp, respectively, are the most used tools for video communications.

Social media and email are also used for routine synchronous and asynchronous communications. In addition, people use email, online surveys, and Google Sheets to exchange virtual services at work.

In [10], Mishra, Gupta, and Shree talked about the experience of moving the Mizoram University from a regular face-to-face learning model to an online format because of the spread of the pandemic of COVID-19. In the study, the authors address the requirements to do the transfer. They also described how to utilize the existing resources during the transfer process

Abdel-Salam and his colleague conducted a statistical study to investigate the satisfaction level with online teaching after the COVID-19 outbreak [11]. The authors run the study on a sample of students from a public university in the Gulf region. The results indicated that there is no difference in the stratification level between males and females.

DATA COLLECTION

We collected students' emails in an online course sent to the instructor in Spring 2020, which started on January 13 and ended on May 15. We collected two datasets. The first dataset was a sample of emails sent before the start of the outbreak of COVID-19 in the US. It covers emails sent from January 15, 2020, to February 15, 2020. The second dataset consists of emails sent after the spread of COVID-19 and sent from March 25 to April 25. We marked the beginning of the spread of COVD-19 by the spring break, which started on March 16. Then classes did not meet in person until the end of the semester. We picked these two date ranges to ensure the randomness of the collected dataset and avoid specific emails related to the start and the end of the semester.

DATA CLEANING

The purpose of the data cleaning is to extract the main body of emails to be distinct from other material such as greeting and thank you words. Therefore, we stripped off all greeting words such as hi, hello, good morning, good afternoon. Likewise, we removed ending phrases such as thanks, thank you, greeting, best luck, stay warm, have a nice day, sincerely. Moreover, we deleted empty spaces and lines. We made all text octopus consist of one paragraph. As expected, we deleted names.

In addition, we did not include emails that only contained pictures illustrating some problems and difficulties in doing a project. Furthermore, we excluded emails from students who were not enrolled in the class.

RESULTS

In the first dataset, we collected 192 emails, whereas the number in the second one is 153. We then used the deep learning algorithm of BERT to get the sentiment classification. BERT is an abbreviation for Bidirectional Representation for Transformers [12] [13] [14] [15]. Researchers at Google AI created the BERT algorithm in 2018. Initially, BERT was used to understand the meaning of queries performed on Google Search. Then, it used various natural language tasks such as sentence pair classification task and question-answer tasks. The following diagram illustrates how BERT works on classifying a sentence

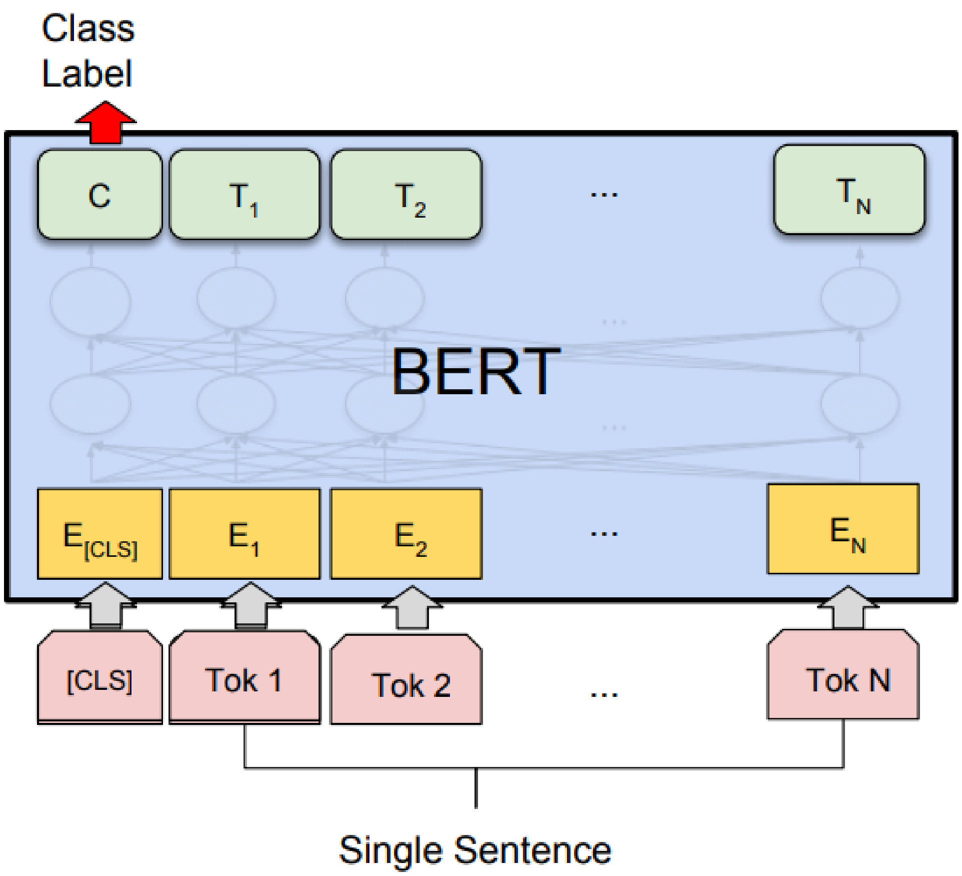

Fig1. BERT single sentence classification (image is depicted from geeksforgeeks.org)

BERT gives the sentiment of a sentence on a scale of 5. A value of 3 of the sentiment means neural, 5 is the happiest/optimistic perspective, and 0 is the most negative/disappointed standing.

Fig 2 shows the code implementation with some run outputs.

```
from transformers import AutoTokenizer, AutoModelForSequenceClassification
import torch
import re
```

```
def sentiment_score(Emails):
    tokens = tokenizer.encode(Emails, return_tensors='pt')
    result = model(tokens)
    return int(torch.argmax(result.logits))+1
```

```
sentiment_score(df['Emails'].iloc[1])
```

```
5
```

```
df['sentiment'] = df['Emails'].apply(lambda x: sentiment_score(x[:512]))
```

```
df
```

	Emails	sentiment
0	Perfect!	5
1	Thank you so much for your help.	5
2	I am wondering if we can still submit the extr...	3
3	Sorry to reach out to you so late--I realized ...	2
4	I was wondering how I got 15 points off for my...	2
...	...	...
148	I tried restarting it and it still doesn't wor...	2
149	I was working on number 15C and when I finishe...	1
150	Thank you for the assistance.	5
151	I found out what was going on.	3
152	I'll try that again. I did once and it wouldn'...	1

153 rows × 2 columns

Fig 2. Code implementation of BERT

We did very simple compassion taking the average of each set sentiment. The average sentiment before the COVID is 2.52 and is after 2.38. We can notice there is a little drop in the sentiment after the outbreak. Our interpretation of the almost negative average sentiment of both sets is due to three reasons. 1) Students were uncertain about the semester and the grades they could have. The uncertainty was even deepened in the second half of the semester with incoming news of the wide-spreading of a global pandemic. 2) The weather in that area is typically freezing during the spring, which could be a contributing factor to a negative impact on the students. 3) Even though the weather was getting much better in the area in the second half of the semester; nonetheless, young people were so concerned about how it could sever on them. Our conclusion is those three factors need to be separated to identify the real reason for the negativity of the sentiment. However, the kind of study we conducted cannot separate them from one of the other. Identifying the sort of impact of each of these three factors is a plan for future work to investigate.

We took the average of the *daily* sentiment and drew a line graph to illustrate the changes over time in Fig 3 and Fig 4

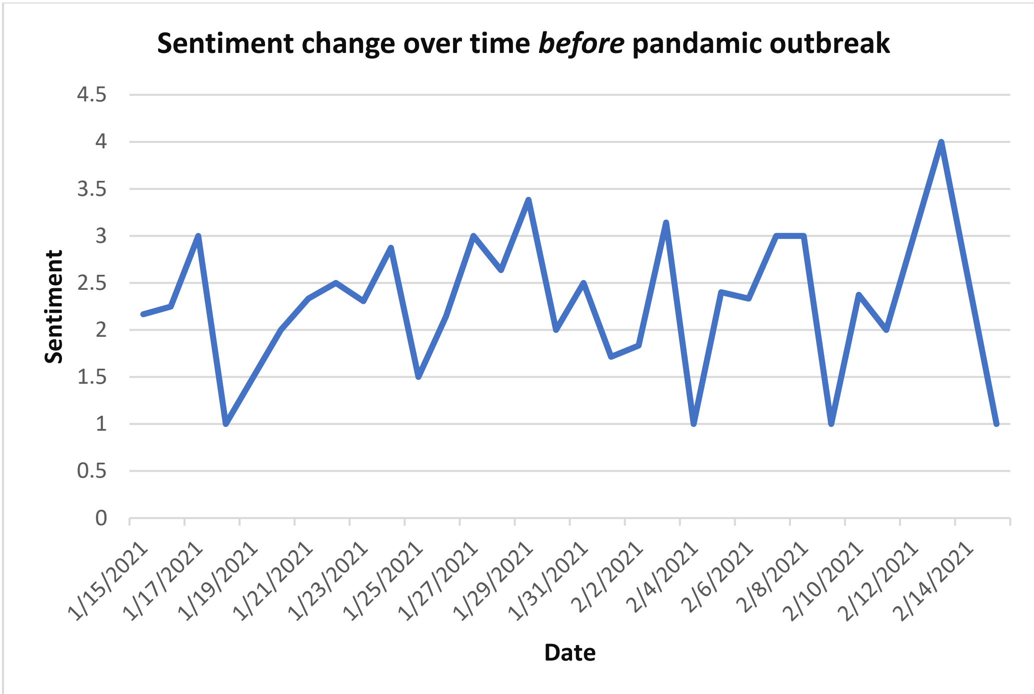

Fig 3. The graph shows the sentiment changes over time *before* the spread of the pandemic

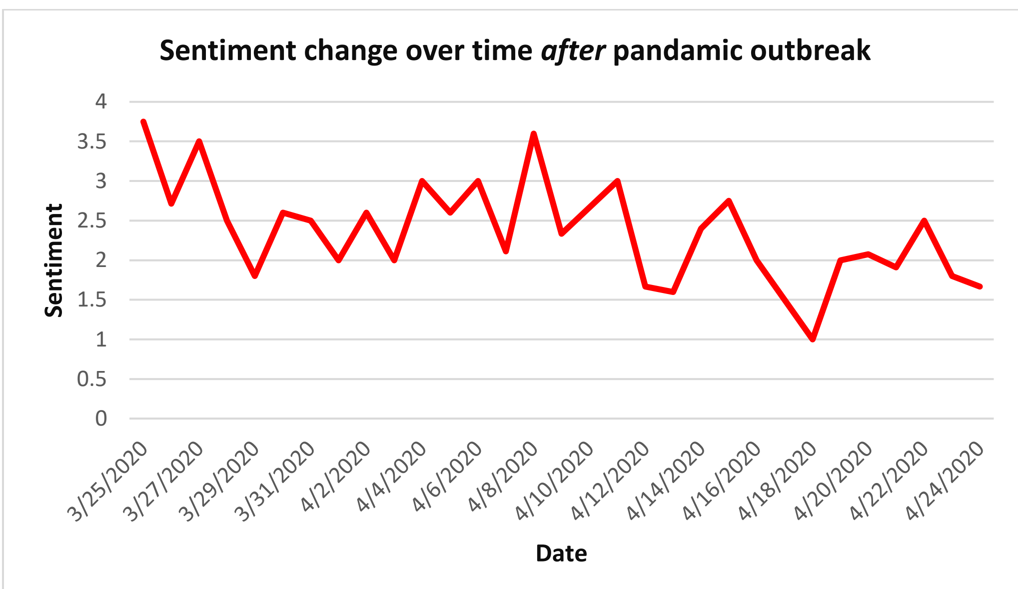

Fig 4. The graph shows the sentiment changes over time *after* the spread of the pandemic

EXPLORATORY DATA ANALYSIS:

To have unique insight into the two datasets, we created the Wordcloud figures.

Fig 5. Wordcloud plot of pre-pandemic dataset

Fig 6. Wordcloud plot of after the spread of the COVID-19 Pandemic

As it may be noticed, there are common and also different words between the two Wordcloud plots. Common words are project, assignments, completed, etc. Examples of the new, used words in the second data late, sorry, problem grades

CONCLUSION:

This study compared student engagement in an online course before and after the COVID-19 outbreak by analyzing email sentiment using deep learning techniques. The slight decline in average sentiment suggests that the pandemic, along with other contextual factors, influenced students' perceptions and engagement. Future research will aim to separately evaluate the impact of academic uncertainty, environmental conditions, and health concerns on student sentiment.

REFERENCES:

[1] Y. Ismaili, "Evaluation of students' attitude toward distance learning during the pandemic (Covid-19): a case study of ELTE university," *Horiz.*, vol. 29, no. 1, pp. 17–30, Jan. 2021, doi: 10.1108/OTH-09-2020-0032.

[2] K. A. A. Gamage, D. I. Wijesuriya, S. Y. Ekanayake, A. E. W. Rennie, C. G. Lambert, and N. Gunawardhana, "Online Delivery of Teaching and Laboratory Practices: Continuity of University Programmes during COVID-19 Pandemic," *Educ. Sci.*, vol. 10, no. 10, Art. no. 10, Oct. 2020, doi: 10.3390/educsci10100291.

[3] I. Kamarianos, A. Adamopoulou, H. Lambropoulos, and G. Stamelos, "TOWARDS AN UNDERSTANDING OF UNIVERSITY STUDENTS' RESPONSE IN TIMES OF PANDEMIC CRISIS (COVID-19)," *Eur. J. Educ. Stud.*, vol. 7, no. 7, Art. no. 7, Jun. 2020, doi: 10.46827/ejes.v7i7.3149.

[4] A. López-Valenciano, D. Suárez-Iglesias, M. A. Sanchez-Lastra, and C. Ayán, “Impact of COVID-19 Pandemic on University Students’ Physical Activity Levels: An Early Systematic Review,” *Front. Psychol.*, vol. 11, 2021, Accessed: March 11, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2020.624567

[5] K. Kedraka and C. Kaltsidis, "EFFECTS OF THE COVID-19 PANDEMIC ON UNIVERSITY PEDAGOGY: STUDENTS' EXPERIENCES AND CONSIDERATIONS," *Eur. J. Educ. Stud.*, vol. 7, no. 8, Art. no. 8, Jul. 2020, doi: 10.46827/ejes.v7i8.3176.

[6] S. Sujarwo, S. Sukmawati, A. Akhiruddin, R. Ridwan, and S. S. S. Siradjuddin, “An Analysis of University Students’ Perspective on Online Learning in the Midst of Covid-19 Pandemic,” *J. Pendidik. Dan Pengajaran*, vol. 53, no. 2, Art. no. 2, Jun. 2020, doi: 10.23887/jpp.v53i2.24964.

[7] M. Mansoor, K. Gurumurthy, A. R. U, and V. R. B. Prasad, "Global Sentiment Analysis Of COVID-19 Tweets Over Time," *ArXiv201014234 Cs*, Nov. 2020, Accessed: February 23, 2021. [Online]. Available: http://arxiv.org/abs/2010.14234

[8] U. Desai, V. Ramasamy, and J. Kiper, "Evaluation of student collaboration on canvas LMS using educational data mining techniques," in *Proceedings of the 2021 ACM Southeast Conference*, New York, NY, USA, Apr. 2021, pp. 55–62. doi: 10.1145/3409334.3452042.

[9] D. Vargo, L. Zhu, B. Benwell, and Z. Yan, "Digital technology use during COVID-19 pandemic: A rapid review," *Hum. Behav. Emerg. Technol.*, vol. 3, no. 1, pp. 13–24, 2021, doi: 10.1002/hbe2.242.

[10] L. Mishra, T. Gupta, and A. Shree, "Online teaching-learning in higher education during lockdown period of COVID-19 pandemic," *Int. J. Educ. Res. Open*, vol. 1, p. 100012, Jan. 2020, doi: 10.1016/j.ijedro.2020.100012.

[11] A.-S. G. Abdel-Salam, K. Hazaa, and E. A. Abu-Shanab, "Evaluating the Online Learning Experience at Higher Education Institutions," *Int. J. Cyber Behav. Psychol. Learn. IJCBPL*, vol. 12, no. 1, pp. 1–23, Jan. 2022, doi: 10.4018/IJCBPL.298692.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *ArXiv181004805 Cs*, May 2019, Accessed: March 11, 2022. [Online]. Available: http://arxiv.org/abs/1810.04805

[13] S. Ravichandiran, *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*. Packt Publishing Ltd, 2021.

[14] C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to Fine-Tune BERT for Text Classification?," in *Chinese Computational Linguistics*, Cham, 2019, pp. 194–206. doi: 10.1007/978-3-030-32381-3_16.

[15] Z. Gao, A. Feng, X. Song, and X. Wu, "Target-Dependent Sentiment Classification With BERT," *IEEE Access*, vol. 7, pp. 154290–154299, 2019, doi: 10.1109/ACCESS.2019.2946594.