2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #38248

Board 63: Work in progress: Uncovering engineering students' sentiments from weekly reflections using natural language processing

Mr. Ahmed Ashraf Butt, Purdue University at West Lafayette (COE)
Dr. Saira Anwar, Texas A&M University

Saira Anwar is an Assistant Professor at Department of Multidisciplinary Engineering, Texas A &M University. Dr. Anwar has over 13 years of teaching experience, primarily in the disciplines of engineering education, computer science and software engineering. Her research focuses on studying the unique contribution of different instructional strategies on students' learning and motivation. Also, she is interested in designing interventions that help in understanding conceptually hard concepts in STEM courses. Dr. Anwar is the recipient of the 2020 outstanding researcher award by the School of Engineering Education, Purdue University. Also, she was the recipient of the "President of Pakistan Merit and Talent Scholarship" for her undergraduate studies.

Dr. Muhsin Menekse, Campbell University

Muhsin Menekse is an Assistant Professor at Purdue University with a joint appointment in the School of Engineering Education and the Department of Curriculum & Instruction. Dr. Menekse's primary research focus is on exploring K-16 students' engagement and learning of engineering and science concepts by creating innovative instructional resources and conducting interdisciplinary quasi-experimental research studies in and out of classroom environments. Dr. Menekse is the recipient of the 2014 William Elgin Wickenden Award by the American Society for Engineering Education. Dr. Menekse also received three Seed-for-Success Awards (in 2017, 2018, and 2019) from Purdue University's Excellence in Research Awards programs in recognition of obtaining three external grants of $1 million or more during each year. His research has been generously funded by grants from the Institute of Education Sciences (IES), Purdue Research Foundation (PRF), and National Science Foundation (NSF).

©American Society for Engineering Education, 2023

Work-in-progress: Uncovering engineering students' sentiments from weekly reflections using Natural Language Processing

Introduction

In recent years, STEM education literature has increasingly emphasized students' emotional states, sentiments, and attitudes, and their relationship with learning outcomes. Such emphasis has led to a growing interest in examining the role of students' feelings or sentiments in their learning process [1]. Traditionally, it has not been easy to manually collect students' subjective feelings through questionnaires [2] or class observations [3]. Therefore, different attempts have been made to develop automated approaches to evaluate students' sentiments, particularly in STEM education, where data-driven approaches are common [4].

Sentiment Analysis (SA) is a computational method that automatically uncovers the sentiments and feelings expressed in the text into positive, negative, or neutral sentiments [5]. Recently, SA has been applied in education research to better understand students' internal expression towards different aspects of their learning experience, such as course content, teaching methods, and classroom environment [6]. For example, Sun and colleagues [7] used SA to analyze students' reflections in a STEM course. The study highlighted the importance of SA in unveiling students' subjective feelings (sentiments) for learning experiences. The authors suggest that this information could be used to improve course design and teaching practices and provide personalized feedback to students.

In this paper, we used a lexicon-based SA approach, specifically the VADER (Valence Aware Dictionary and Sentiment Reasoner) algorithm [4], to understand First-Year Engineering (FYE) students' sentiments from their course reflections. The VADER algorithm has been widely used in SA tasks and has been shown to perform well for short and informal text [8]. To collect students' reflections, we used an educational learning system called CourseMIRROR [9] that prompts students to reflect on each lecture's confusing or interesting aspects throughout the semester and provides a rich data source for SA. More specifically, this study is guided by the following two research questions: RQ1) What kind of student sentiments occur when students reflect on the interesting and confusing aspect of the lecture? And RQ2) How do students' sentiments change over the semester?

Research Methods

This study follows a correlation research design where data is analyzed using quantitative approaches.

Site and participants: In this study, we recruited students from three sections of FYE enrolled in an introductory engineering course at a large public university in the Midwest, United States. The goal of the course was to provide students with a foundation in programming skills using MATLAB and to develop their critical thinking and problem-solving abilities using mathematical models in addressing engineering problems. Three hundred eighty-two students voluntarily participated in the study, and their reflections are used as a rich data set for this study.

Data collection: We used the CourseMIRROR educational application to gather students' reflections. The application prompted students with two open-ended questions to reflect on the 1) confusing aspect of the course (Confusion) or 2) interesting aspects of the course (Interesting) at the end of each lecture. A total of 7116 reflections were voluntarily submitted by students throughout 24 lectures, with 3558 reflections for each question.

Procedure and Data Analysis: The reflections data were analyzed using Lexicon-based sentiment assessment, specifically VADER (Valence Aware Dictionary and Sentiment Reasoner) algorithm, a commonly used SA technique in NLP. The algorithm calculates the sentiment polarity of text by analyzing the words used and assigns a score that ranges from -1 to 1. The reflections with a score greater than 0.05 were considered positive and those with a score less than -0.05 were considered negative. The remaining reflections were considered neutral. The threshold [10], commonly used in the literature, was used to assign sentiment scores and inform our research questions.

For RQ1, all reflections were considered in single group for both question types. For RQ2, sentiment in the students' reflection was counted at three equal time points (lectures 1-8, 9-16, and 17-24), and Friedman's two-way analysis of variance by ranks was used for analysis as data violated the normality assumption of One-way Repeated Measure ANOVA.

Results

To answer RQ1, we first used descriptive statistics and examined the frequency of various sentiments found in students' reflections. The SA results for the students' reflections are summarized in Table 2.

Table 2. Descriptive statistics of students' reflection sentiments

	Reflection question type	
Sentiments	Confusion	Interesting
Positive	1156	2437
Neutral	1128	990
Negative	1274	131

It is observed that while answering the confusion question, students had negative sentiments (e.g., 1274), indicating that confusion also causes negative sentiments, while 1156 reflections had positive sentiments. In contrast, with interesting questions, most reflections indicated positive sentiments (e.g., 2437), and fewer indicated negative sentiments in students. The reflections with neutral sentiment were lower for both question types, with 1128 students for the confusing question and 990 for the interesting question. To answer RQ2 and to examine the change of sentiment scores across different time points of the class for each question type, we used Friedman's two-way analysis of variance by ranks (see Table 3 for results). We conducted a separate test for each sentiment.

Table 3. Results of Friedman's two-way analysis of variance by ranks

	Confusing	Interesting

Time points	Mean rank	χ2 ANOVA (2,382)	*p* value	Mean rank	χ2 ANOVA (2,382)	*p* value
Positive						
Lectures 1-8	1.97	16.95	< 0.001**	2.05	127.41	< 0.001**
Lecture 9-16	2.13			2.33		
Lecture 17-24	1.90			1.62		
Neutral						
Lectures 1-8	1.89	11.88	0.003*	1.76	94.40	< 0.01*
Lecture 9-16	2.07			1.92		
Lecture 17-24	2.05			2.32		
Negative						
Lectures 1-8	1.92	7.82	0.020*	1.98	18.65	< 0.001**
Lecture 9-16	2.09			1.94		
Lecture 17-24	1.98			2.07		

*, ** indicates $p < 0.05$ and $p < 0.001$, respectively.

The results showed a significant difference in each sentiment, as indicated by the χ2 values and p-values in table 3 for both question types. We further carried out pairwise comparisons with a Bonferroni correction for each sentiment in the students' reflections on the interesting question. The positive sentiment in students' reflections showed a significant difference between the first time point (Mdn = 1.00) and the second (Mdn = 2.00) and third time points (Mdn = 0.00) with p < 0.001. The negative sentiment remained constant throughout the students' reflections at each time point, with no significant differences between any time points. The neutral sentiment in the students' reflection showed significant differences between the first time point (Mdn = 0.00) and the third time point (Mdn = 2.00; p < .001) and between the second (Mdn = 0.00) and third time points (p < 0.001).

For each sentiment in the students' reflections on the confusing question, the positive sentiment showed a significant difference between the second and third-time points (Mdn = 0.0; p =0.004). However, no other significant differences were found between other time points. Regarding negative sentiments, no significant differences were found between any pair of time points, suggesting that negative sentiment remained constant across all time points. For neutral sentiments, the results showed that the neutral sentiment was significantly different between the first time point (Mdn = 0.00) and the second time point (Mdn = 1.00; p < .001), but no other significant differences were found between any other pairs of time points.

Discussion and Conclusion

This research study aimed to understand the sentiments expressed by students in their reflections for an introductory FYE course. The VADER algorithm was used to evaluate the students' sentiments, and the frequency of positive, neutral, and negative sentiments was used to inform the study. Descriptive statistics and Friedman analysis were used to understand the presence of the most prevalent sentiment in the students' reflections and changes in the presence of each sentiment in their reflections over a semester. The results showed that the positive sentiment was the most prevalent in students' reflections on the interesting question. Negative sentiment was more prevalent for reflections on the confusing question, but a substantial number of positive sentiments were also expressed. The reasons for positive sentiments could be rooted in the fact that students either were excited to overcome the confusion or felt optimistic about a chance of learning.

In literature, reflection refers to a moment or process where students establish their understanding by becoming aware of their thoughts, sentiments, and past experiences [11], [12]. Positive sentiment in reflections can indicate growth, self-awareness, or resolution of conflicts [13]. Hence, we see the prevalence of positive sentiments in their reflections. Negative sentiments, a broader category of the students' sentiment, has linked to students' dissatisfaction or frustration with their class learning experience [14]. Therefore, it makes sense that negative sentiment was found in students' reflections when explaining confusing lecture aspects, highlighting the importance of instructors addressing and understanding the source of this negativity to enhance students' learning experience.

Friedman's test showed a significant difference in students' sentiments over time in their reflections on the interesting and confusing aspects of the lecture. The increase in positive and neutral sentiments and the persistence of negative sentiments could reflect the reflection activity's role in building student resilience and coping skills [15]. Furthermore, the study result must be viewed with some limitations. First, the study followed a correlational research design; thus, experimental studies are warranted to explore any causality. Second, the algorithm used for the sentiment is limited in specific areas, such as lack of context awareness, being rule-based, and inability to understand the nuances of human emotion. Therefore, there is a need to explore different SA approaches that can more effectively assess students' sentiments. Another limitation is the limited size, which may not represent the larger STEM population. Future research could address these limitations by conducting a more extensive and diverse study to better understand the intervention's impact on students' sentiments. Additionally, future studies could combine psycho-physiological and reflection data to understand the relationship between students' sentiments, learning, and well-being.

Acknowledgments

This research was partly supported by a grant from the Institute of Education Sciences (R305A180477), and the views expressed do not represent those of the U.S. Department of Education.

References:

[1] M. J. Price, N. V. Mudrick, M. Taub, and R. Azevedo, "The role of negative emotions and emotion regulation on self-regulated learning with MetaTutor," in *Intelligent Tutoring Systems*, R. Nkambou, R. Azevedo, and J. Vassileva, Eds., Cham: Springer International Publishing, 2018, pp. 170–179.

[2] C. R. Jackson, "Validating and adapting the motivated strategies for learning questionnaire (MSLQ) for STEM Courses at an HBCU," *AERA Open*, vol. 4, no. 4, p. 2332858418809346, Oct. 2018, doi: 10.1177/2332858418809346.

[3] L. Michaluk, R. Stoiko, G. Stewart, and J. Stewart, "Beliefs and attitudes about science and mathematics in pre-service elementary teachers, STEM, and Non-STEM majors in

undergraduate physics courses," *J. Sci. Educ. Technol.*, vol. 27, no. 2, pp. 99–113, Apr. 2018, doi: 10.1007/s10956-017-9711-3.

[4] Y. Wu, Z. Ming, J. K. Allen, and F. Mistree, "Evaluation of students' learning through reflection on doing based on Sentiment Analysis," *J. Mech. Des.*, vol. 145, no. 3, Jan. 2023, doi: 10.1115/1.4056495.

[5] A. P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, Jul. 2016, pp. 628–632. doi: 10.1109/ICATCCT.2016.7912076.

[6] K. Mite-Baidal, C. Delgado-Vera, E. Solís-Avilés, A. H. Espinoza, J. Ortiz-Zambrano, and E. Varela-Tapia, "Sentiment analysis in education Domain: A systematic literature review," in *Technologies and Innovation*, R. Valencia-García, G. Alcaraz-Mármol, J. Del Cioppo-Morstadt, N. Vera-Lucio, and M. Bucaram-Leverone, Eds., Cham: Springer International Publishing, 2018, pp. 285–297.

[7] Y. Sun, Z. Ming, Z. Ball, S. Peng, J. K. Allen, and F. Mistree, "Assessment of Student Learning Through Reflection on Doing Using the Latent Dirichlet Algorithm," *J. Mech. Des.*, vol. 144, no. 12, Sep. 2022, doi: 10.1115/1.4055376.

[8] U. Naseem, I. Razzak, K. Musial, and M. Imran, "Transformer based Deep Intelligent Contextual Embedding for Twitter sentiment analysis," *Future Gener. Comput. Syst.*, vol. 113, pp. 58–69, Dec. 2020, doi: 10.1016/j.future.2020.06.050.

[9] X. Fan, W. Luo, M. Menekse, D. Litman, and J. Wang, "CourseMIRROR: Enhancing Large Classroom Instructor-Student Interactions via Mobile Interfaces and Natural Language Processing," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, in CHI EA '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 1473–1478. doi: 10.1145/2702613.2732853.

[10] M. Singh, H. K. Dhillon, P. Ichhpujani, S. Iyengar, and R. Kaur, "Twitter sentiment analysis for COVID-19 associated mucormycosis," *Indian J. Ophthalmol.*, vol. 70, no. 5, p. 1773, 2022.

[11] Q. D. Nguyen, N. Fernandez, T. Karsenti, and B. Charlin, "What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model," *Med. Educ.*, vol. 48, no. 12, pp. 1176–1189, Dec. 2014, doi: 10.1111/medu.12583.

[12] A. A. Butt, S. Anwar, and M. Menekse, "WIP: Investigating the relationship between FYE students' reflections and academic performance across gender," in *2022 First-Year Engineering Experience*, 2022.

[13] M. A. Cohn, B. L. Fredrickson, S. L. Brown, J. A. Mikels, and A. M. Conway, "Happiness unpacked: Positive emotions increase life satisfaction by building resilience.," *Emotion*, vol. 9, pp. 361–368, 2009, doi: 10.1037/a0015952.

[14] K. Trigwell, R. A. Ellis, and F. Han, "Relations between students' approaches to learning, experienced emotions and outcomes of learning," *Stud. High. Educ.*, vol. 37, no. 7, pp. 811–824, Nov. 2012, doi: 10.1080/03075079.2010.549220.

[15] T. Trivate, A. A. Dennis, S. Sholl, and T. Wilkinson, "Learning and coping through reflection: exploring patient death experiences of medical students," *BMC Med. Educ.*, vol. 19, no. 1, p. 451, Dec. 2019, doi: 10.1186/s12909-019-1871-9.