COPUSense: A web application that evaluates STEM classroom practices using Edusense

Yu Fang

December 4, 2023

Abstract

With the recent development of automated classroom analytics using machine learning - Edusense [1], there is still little evidence on the impact of how this type of education technology can be useful to improve personalized learning. There are, however, studies supporting the benefits of student-centered learning over traditional lecturing [2–4]. Interestingly, students' did not perceive active learning as positive though it led to better learning outcome [4]. Beyond their self-evaluation, students tend to negatively review instructors of harder courses on RateMyProfessor [5] and can be biased against larger class size [6]. Finally, there exist multiple forms of student engagement [7], and the ability to capture these will better characterize learning. Therefore, we will characterize the classes with Classroom Observation Protocol for Undergraduate STEM (COPUS) [8].

To address these nuances, our study employs a classifier trained to predict Classroom Observation Protocol for Undergraduate STEM (COPUS) [8]. Our goal is to move beyond the simple dichotomy of active versus passive learning and gain a deeper understanding of classroom dynamics. First, we improved the model generalizability by incorporating a few sessions from the new classroom. Second, we found that the synchrony across COPUS activities were different in the new classroom 05391A. Lastly, based on the 4 main categories: Presenting, Guiding, Administration and Other [8], we easily identify the 3 sessions that contained an increased ratio of guiding. Through this analysis, we described teaching style differences through the lens of a machine learning model.

Body of proposal

Objectives and Contribution of the Research

Traditional lecturing is increasingly being replaced by student-centered methods due to its effectiveness in engaging students to regurgitate information. Additionally, there exist multiple forms of student engagement [7], and the ability to capture these will better characterize learning. The advent of machine learning technologies has ushered in a new era in education, particularly in classroom analytics. Preliminary analyses showed that classroom sensing data collected via EduSense [1] correctly classified classroom activity. By transforming sensing data into classroom activity predictions, we can begin to characterize teaching styles. Moreover, a longitudinal classroom study could help reveal teaching trends that evolve from beginning to the end of a semester. Overall, we have explored the utility of using a machine learning model to monitor/summarize teaching styles both across courses and within a course. The research question we want to ask is how learning can be defined using a more comprehensive description of classroom activity, i.e. time in each COPUS. In addition, we want to probe how personality traits [3] impacts the receptiveness to each form of engagement. We hypothesize that if various forms of student-centered learning contribute differently to student's learning, certain student activities could better predict exam/quiz scores than others. However, we also believed that this effect is different for people with different levels of communication apprehension, self-esteem and/or willingness to participate. The first hypothesis will address the effectiveness of the different forms of student engagement for learning, whereas the second hypothesis answers the individuality of these impact. Overall, filling both knowledge gaps provides insight into how to dynamically tailor a course for individual students.

Methodology

Our research project investigate the impact of automated classroom analytics on personalized learning. To achieve this, we will employ a multifaceted methodology involving model building and validation, data analysis, and continuous application development. This comprehensive approach allows us to explore the subject thoroughly. Previous model was trained on predominantly lecture-based sessions, such as courses numbered 36-200 Reasoning with Data, 15-251 Great Theoretical Ideas in Computer Science, 21-122 Integration and Approximation, and 73-102 Principles of Microeconomics. This approach, however, differed from the teaching style of Professor Harrison in the course '05-391 Designing Human-Centered Software,' which is known for its interactive and dynamic classroom environment. Upon incorporating data, we found that all activity improved significantly using the updated model(Fig. 1). This improvement shows the potential of the model to be continuously generalize to other classrooms.

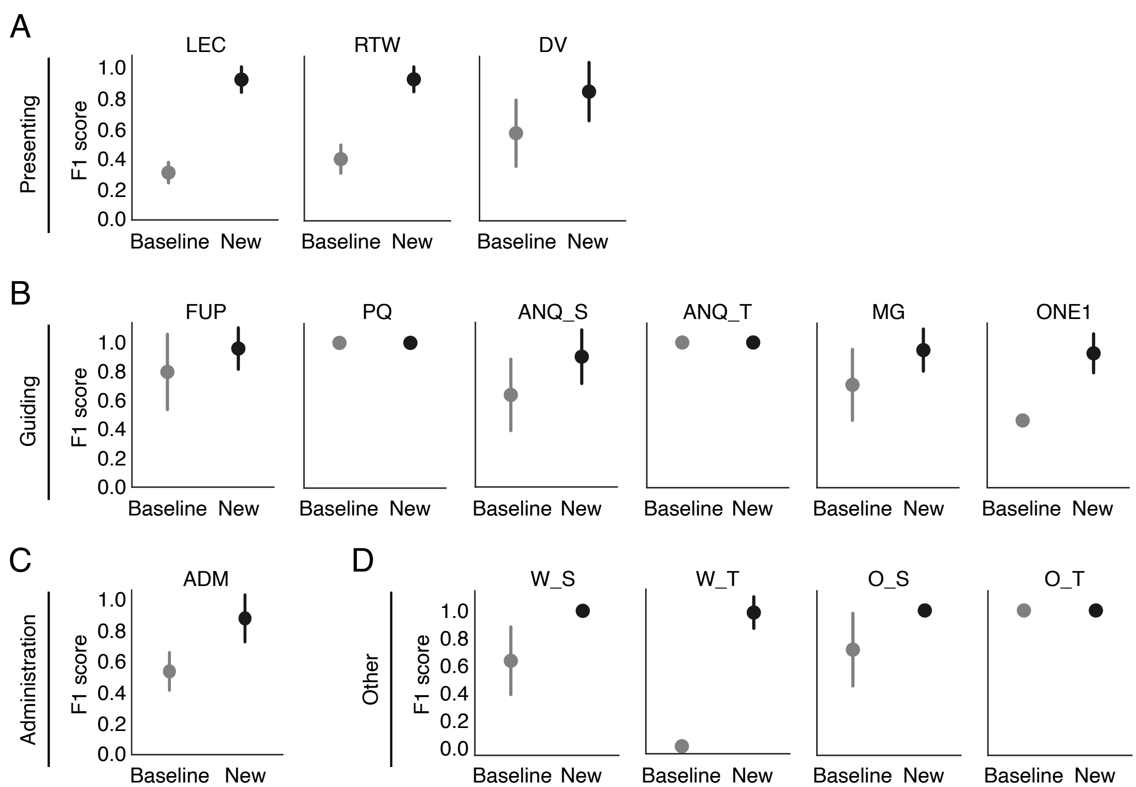

Figure 1: COPUSense performance on a new classroom. Mean ± standard deviation of performance on 20 randomly partitioned held-out new classroom data for each of the four main categories (A: Presenting; B: Guiding; C: Administration; D: Other). New is after incorporating 5 sessions from the new classroom - 407, whereas baseline is before.

Our research will be conducted in a classroom equipped with pre-installed cameras to capture classroom activities. We will also process video data directly without human intervention. With each session, we can apply the updated machine learning model to gain insights into instructor's teaching style. An example can be found when applying the updated model on lecture-based and the new classroom (05391A). Preliminary analysis revealed interesting differences in pairs of COPUS activities. By definition, most of lecture time (LEC) points will coincide with instructor writing on board (RTW) (Fig. 2A), but almost never when the instructor is waiting (W_T) (Fig. 2B). We can compute such pairwise correlation with all the COPUS activity to characterize overall styles. We found that in lecture-based classes (05681A, 15251C, 19603A, 21122S, 36200S, and 73102S), there appears to be higher correlation amongst individual guiding exercises, the likes of follow-up questions (FUP), posting clicker questions (PQ), students asking questions (ANQ_S), and instructor asking questions (ANQ_T) (Fig. 2C). On the other hand, that correlation structure is minimal and only appears between PQ and ANQ_S (Fig. 2D) in 05391A. Moreover, in 05391A the presenting COPUSes (lecturing (LEC) and instructor writing on board (RTW)) were negatively correlated with the occurrences of a couple guiding exercises (moving through class guiding students during active tasks (MG) and one-on-one discussions (ONE1)). However, that does not be the case for the other classes, aligning with the lecture-based format.

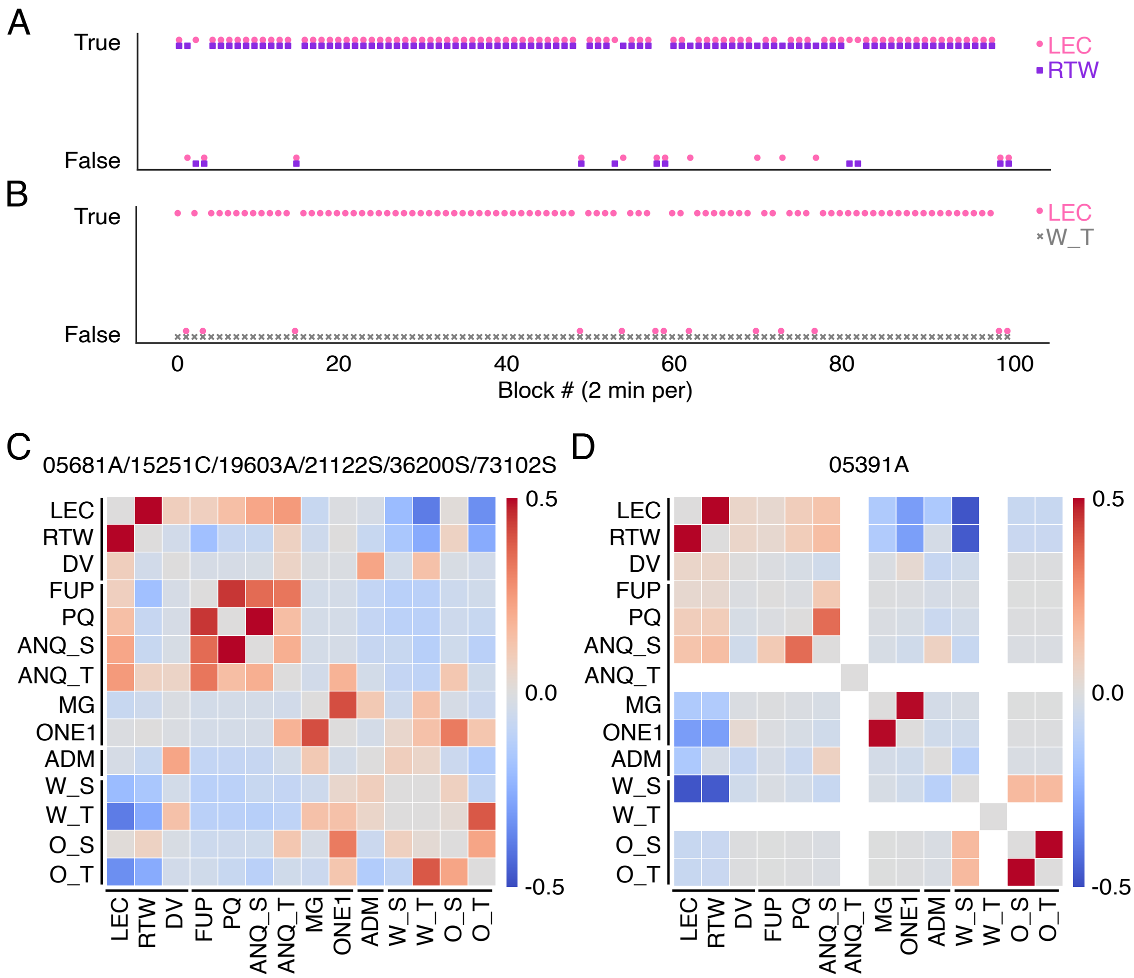

Figure 2: Teaching style differences demonstrated by inter-COPUS activity correlation. A) An 100-block example of predicted occurrences of LEC (lecturing) and RTW (real-time writing). B) The same 100-block example of predicted occurences of LEC (lecturing) and W_T (intructor waiting). C) Heatmap of pair-wise inter-COPUS activity correlation for classrooms 05681A, 15251C, 19603A, 21122S, 36200S, and 73102S. D) Heatmap of pairwise inter-COPUS activity correlation for classroom 05391A.

In terms of course sampling, we will record all sessions. Our technology can also study how classes that focus on different concepts may represent on a COPUS level. It is believed that each session will include a varying amount of presenting versus guiding simply due to the nature of the materials. We examined 05391A throughout the entire 2019 Spring semester. For each of the four main COPUS categories (presenting, guiding, administration, and other), we compute the percentage of occurrence averaged across all activity that belonged to that category (LEC, RTW, DV belong to presenting, for example). To better understand the relative amount, we divide each category's mean percentage of occurrence to the total (summation of all four categories). As expected, the ratio between these four main COPUS categories are dynamic, and easily identified sessions 5, 16, and 17 contain more guiding activities than the rest (Fig. 3A). Additionally, when we grouped by beginning, middle, and late semester (February 2019, March 2019, and April 2019), we found an interesting progression. During the middle part of the semester, the instructor seems to spend more time presenting, and less time guiding the students through active exercises. It can be due to a plethora of reasons, including, but not limited to, instructor becoming less engaged, students attendance dropped, or simply due to course material being less practical to be guided. Overall, we gained valuable insights into instructor teaching styles through our machine learning model.

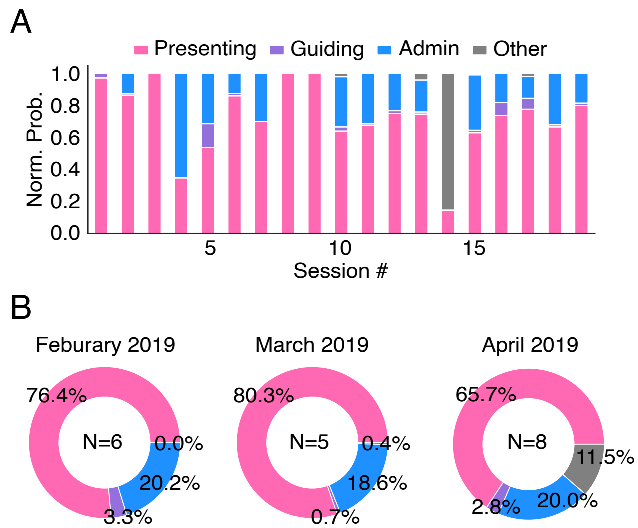

Figure 3: Utility of COPUSense in a longitudinal 05391A classroom study. A) Stacked bar plot with normalized activity ratio for each session. B) Pie chart of normalized activity ratio binned by each month.

Moreover, we will follow 10 selected students throughout the term. To incentivize participation, we will provide each student with a budget of $50 to secure permission for facial recordings and post-test surveys. To track the students, we will use the VGG face recognition model within DeepFace `https://github.com/serengil/deepface` [9–11]. Beyond tracking the identity of students and exporting their position, movement, and posture using EduSense [1], the video will not be used for further analysis. Our lab has obtained an IRB: STUDY2019_00000142: Instructor Notifications in University Setting. These surveys and recordings will provide valuable insights into students' reactions and actions and allow us to track students during lectures and the impact of active learning on their academic performance.

Our analysis will involve the use of regression analysis to assess multi-variate correlations between teaching styles and students' average exam/quiz scores. We will test the hypothesis that student-centered activities positively correlate with improved exam scores. Our methodological choices aim to ensure a comprehensive understanding of the impact of automated classroom analytics on personalized learning. By combining video analysis, surveys, and regression analysis, we create a well-rounded perspective. The research project is organized within an 4-month timeline, with distinct phases that include model retrieval, data collection, data analysis, application development, post-study reporting, refinement, and considerations for future directions. This timeline provides us with ample time to execute each phase effectively while maintaining strict adherence to ethical guidelines and privacy regulations throughout the project.

Background

I spearheaded the development of web application to visualize classroom activity. In my academic journey, I excelled in courses like **15-112 Fundamentals of Programming**, **15-121 Introduction to Data Structures**, and **05-499 Special Topics in HCI** where I developed a recognized game and a mobile therapy application. I also gained professional experience as a Data Analytics intern at Siemens and served as a Teaching Assistant for "15-110 Principles of Computing." Additionally, I have taken on leadership roles as Student Body Vice President and currently as the Vice President for Outreach and Finance at the User Experience Association. My achievements reflect my proficiency in programming, data analysis, professionalism, and leadership.

Budget

An allocation of 500 dollars is set aside to incentivize participants. This fund will be used to obtain their consent for portrait recognition, enabling accurate tracking of their in-class activity using the DeepFace technology and the camera system in Room 407. Additionally, this budgetary provision will motivate participants to share their exam scores post-assessment.

References

1. Ahuja, K. *et al.* EduSense: Practical Classroom Sensing at Scale. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol* **3,** 71. `https://doi.org/10.1145/3351229` (2019).
2. Freeman, S. *et al.* Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 8410–8415. ISSN: 10916490. `https://www.pnas.org/doi/abs/10.1073/pnas.1319030111` (June 2014).
3. Rocca, K. A. Student Participation in the College Classroom: An Extended Multidisciplinary Literature Review. **59,** 185–213. ISSN: 1479-5795 (2010).
4. Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K. & Kestin, G. Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences of the United States of America* **116,** 19251–19257. ISSN: 10916490. `https://www.pnas.org/doi/abs/10.1073/pnas.1821936116` (Sept. 2019).
5. Gao, B. J. & Katrompas, A. A Preliminary Experimental Analysis on RateMyProfessors. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020,* 5679–5681 (Dec. 2020).
6. Liaw, S. H. & Goh, K. L. Evidence and control of biases in student evaluations of teaching. *International Journal of Educational Management* **17,** 37–43. ISSN: 0951354X (2003).
7. Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research* **74,** 59–109. ISSN: 00346543 (2004).
8. Smith, M. K., Jones, F. H., Gilbert, S. L. & Wieman, C. E. The classroom observation protocol for undergraduate stem (COPUS): A new instrument to characterize university STEM classroom practices. *CBE Life Sciences Education* **12,** 618–627. ISSN: 19317913. `https://www.lifescied.org/doi/10.1187/cbe.13-08-0154` (Dec. 2013).
9. Parkhi, O. M., Vedaldi, A. & Zisserman, A. Deep face recognition. *BMVC 2015 - Proceedings of the British Machine Vision Conference 2015* (2015).
10. Serengil, S. I. & Ozpinar, A. LightFace: A Hybrid Deep Face Recognition Framework. *Proceedings - 2020 Innovations in Intelligent Systems and Applications Conference, ASYU 2020* (Oct. 2020).
11. Serengil, S. I. & Ozpinar, A. HyperExtended LightFace: A Facial Attribute Analysis Framework. *7th International Conference on Engineering and Emerging Technologies, ICEET 2021* (2021).