Modelling Student Behavior in Synchronous Online Learning during the COVID-19 Pandemic

Gianni Fenu, Roberta Galici

Department of Mathematics and Computer Science, University of Cagliari, V. Ospedale 72, 09124 Cagliari, Italy

Abstract

The outbreak of COVID-19 has significantly impacted on education, training, and mobility opportunities provided to learners, teachers, and educators. In response to this situation, synchronous online learning has been massively adopted in universities and, therefore, the analysis of student behavior in this context is becoming essential. However, the literature has mainly tackled student behavior modelling in asynchronous online learning (e.g., interactions with pre-recorded videos), making how students learn in synchronous online learning so far under-explored. Grounding on the experience on online learning at the University of Cagliari, this paper proposes a preliminary study on student participation in synchronous lessons, covering more than 80 courses, and identifies patterns and strategies of engagement in classes over the semester. Then, by means of clustering techniques applied to data from more than 25,000 students, we model fine-grained behavioral strategies and discuss how they are related with the student's experience across courses and whether planning elements (e.g., the hour of the day a lesson is delivered) influence their level of participation. We expect that this study will support the educational stakeholders with preliminary data-driven informed decisions and pose the basis for data-driven personalization for students and teachers involved in online synchronous learning.

Keywords

E-Learning, Engagement, Clustering, Learning Behavior, Student Modelling, Synchronous Learning.

1. Introduction

The Covid-19 pandemic has forced schools and universities to move face-to-face learning and teaching in classrooms to online environments. The term online learning refers to educational strategies that rely on the use of multimedia and Internet technologies for learning and teaching. For instance, this strategy has been originally adopted to support students who cannot physically attend lessons in classroom. Recent literature has focused on providing guidelines on and discussing benefits and limitations of online learning, mainly identifying asynchronous and synchronous strategies and addressing questions such as when, why and how to use these two educational delivering modes [1]. Specifically, synchronous learning refers to learning experiences where a group of participants is engaged at the same time, either in the same physical location (e.g., a classroom) or the same online environment (e.g., a web conference room), and can interact among each other in real time. Conversely, asynchronous learning refers to the

L2D'21: First International Workshop on Enabling Data-Driven Decisions from Learning on the Web, March 12, 2021, Jerusalem, IL

fenu@unica.it (G. Fenu); roberta.galici@unica.it (R. Galici)

0000-0003-4668-2476 (G. Fenu); 0000-0002-1492-3514 (R. Galici)

strategies where instructors and students are not engaged in the learning process at the same time (e.g., interactions with pre-recorded videos or on-demand online exams) [2].

Given the increasing attention on synchronous and asynchronous online learning during the last years [3, 4], analyzing how students behave in these contexts becomes of paramount importance for several aspects, such as improving teaching strategies, sizing technological infrastructures, assessing and predicting student engagement, preventing students from losing interest and possibly dropping out, and finally improving the student's learning outcomes. It has been proved that asynchronous learning benefits online students by providing them with the flexibility they need and giving them more time to reflect on the course content. However, all those characteristics that distinguish a face-to-face lesson are missing in asynchronous contexts (e.g., the opportunity of asking questions and the feeling of being part of a class). Therefore, synchronous lessons are usually preferred while moving these face-to-face online.

Tracking students' behavior in both delivering modes is producing a huge amount of data that can open new ways for better understanding how students learn and, therefore, analyze data coming these contexts is becoming an unprecedented opportunity for the research and educational research communities. A vast amount of studies on asynchronous online learning has been carried out in literature [5, 6, 7, 8, 9, 10], mainly dealing with the analysis of clickstreams to predict whether a student will drop out of an online course [11, 12] or the final grade a student will receive [13, 14]. Synchronous online learning is conversely an under-explored domain and, given its growing adoption, there is an urgent need of analyzing how students learn and how their behavior can be modelled under this educational modality as well [15].

In this paper, we analyze how students interact with a synchronous online learning platform in all the courses delivered by a university over a semester. For this purpose, we first collect and pre-process student's interactions in the form of entry-exit records from lesson rooms. Then, we apply clustering techniques to model behavioral patterns shared between students at faculty and course level. Finally, we identify and interpret the clusters, depicting their main characteristics. More precisely, we will answer to the following research questions:

1. How does the level of student's participation in courses change over a semester?
2. Which are the principal participation patterns at faculty and course level?
3. Does the hour of the day a course is delivered influence the level of participation?

The rest of this paper is organized as follows: Section 2 describes related work. Section 3 introduces the educational context and the dataset considered in this paper, while Section 4 provides a preliminary analysis of the students' participation level, the participation strategy modelling, and their analysis over the semester. 5, we represent the discussion, implications and limitations of our work. Finally, Section 6 provides insights for future work.

2. Related Work

Our research lies at the intersection between studies on the analysis of synchronous learning from an educational science perspective and those on educational data mining in the context of online learning. To better contextualize our work, therefore, we discuss representative prior work in these two fields and identify how we differ from them.

2.1. Synchronous Online Learning

Synchronous online learning is one of the most widely adopted methods by teachers due to several reasons, defined as a real-time, instructor-led online learning experience where participants are connected at the same time and communicate with each other [16]. For instance, students can ask questions and get answers in real-time, as the lesson is live, and the instructor can assess and shape students' understanding in real-time, adjusting his/her spoken words accordingly. Moreover, students feel an increased sense of the instructor actually being there and instructors can facilitate workshop-style classes and run breakout group activities, live chats or office hours allow for real-time interaction (e.g., a conversation and synchronous sessions provide a schedule to help students who struggle with task initiation to stay on track). To better understand the benefits of synchronous online learning, for instance, Francescucci et al. [17, 18] investigated the effects of a novel online synchronous course format, based on virtual, interactive, real-time, instructor-led rooms, on students' learning outcomes and their level of engagement, compared to a fully face-to-face format in classroom. Similarly, Kohnke and Moorhouse [19] analyzed whether students' perceptions about communities of inquiry constructs reflect their behaviours in the synchronous online learning settings and in what ways the interrelationships between constructs differ in terms of students' perceptions and behaviours. The widespread adoption of synchronous online courses has become more and more evident during the last year, as demonstrated by the recent studies conducted by [20, 21], whose goal was to evaluate learner's behavior in synchronous online continuing education lectures during the COVID-19 pandemic. Some high schools found themselves in crisis to manage online learning [22]. Indeed, it has become clearer that the education system is susceptible to external dangers [23]. Feldman [24] addresses student assessment during this pandemic on how districts can legislate unbiased and evenhanded grading policies based on these recommendations; (i) pandemic related anxiety will have negative effects on student academic performance, (ii) academic performance of students might be affect by racial, economic and resource differences, and (iii) the larger parts of instructors were not effectively ready to deliver high-quality instruction remotely.

2.2. Student Behavior Modelling

Educational data mining deals with the development of methods for exploring the increasingly large-scale data obtained in educational settings [25, 26, 27, 28, 29]. Recent studies modelled student's behavioral patterns in terms of their participation in online initiatives (e.g., forums, courses) and investigated the existence of relationships between students' behavioral patterns and their learning performance, assuming that different habits of online students will lead to different levels of learning achievement in students [30]. Some studies focused their attention on analyzing server logs, to construct knowledge on typical patterns of online learning behaviors, to further discover the unique advantages of data mining techniques to support dynamic online instruction, and to build a predictive model for online learning [31]. In [32], data mining techniques were applied to extracted LMS server logs to analyze and compare student online learning behaviors between peer-moderated and teacher-moderated groups in an undergraduate online collaborative project-based learning course in Taiwan. The authors in [33] compared the performance of various clustering and classification algorithms applied on the same educational dataset. In [34], the authors provided an unexplored review of educational data mining based on clustering with respect to teaching and learning processes. In [35, 36], a clustering approach to partition students into different groups based on their learning behavior has been presented. Furthermore, the authors presented the personalized e-learning system architecture used to detect and adapt teaching contents according to the students' level of knowledge. Others, such as [37, 38], used the K-Means clustering algorithm to model behavioral patterns of passing and failing students and others built effective indicators of students' performance in degree programmes. Our work differs from prior work in terms of educational context, given that we tackle synchronous and not asynchronous online learning, while we rely on similar techniques to model behavior in this emerging context (e.g., K-Means). We argue that behavioral patterns are highly dependent from the educational context, so our study brings novel observations and contributions to better understand the synchronous context.

3. Educational Context and Dataset

Our work considers data coming from synchronous learning delivered by an Italian university. Specifically, the University of Cagliari has chosen to provide its lesson-related services through Adobe Connect[1], an e-learning platform where students can access to virtual rooms by means of computers, tablets and smartphones. For each student and each day of a given university lesson, we analyzed their participation level based on the amount of lesson time attended by the student (i.e., remained connected to virtual room). Before describing our approach, we present the context, the university structure, and the data collected from Adobe Connect.

University Structure Description. More than 25,000 students have been covered by our study, across the lessons of each degree programme. Six faculties have been involved, namely *Biology and Pharmacy*, *Engineering and Architecture*, *Medicine and Surgery*, *Science*, *Economic, Law and Political Sciences*, and *Humanities*. Figure 1 shows the number of degree programmes per faculty, for a total of 89 degree courses. Each faculty nominates a person for coordinating the lessons planning across programmes, booking then the virtual rooms in Adobe Connect.

Online Educational Infrastructure and Delivering. Each lesson has been delivered in an Adobe Connect virtual room, which provides tools and services for teleconference, e-learning sessions, and collaborative content creation and delivering. Each virtual room is a permanent online space an instructor can use to deliver lessons to students. In the context of the university we consider, more than *50* virtual rooms have been created, and each room was assigned to a lesson of a specific course at a given time slot, similarly to classrooms assignments in a face-to-face context. Then, each instructor is assigned to a virtual room on some fixed time slots across weeks. Students enrolled in a specific course can login using their Adobe Connect account and enter the virtual room associated to the desired lesson. As soon as the instructor starts the lesson, students can take part in it. During the lesson, they can virtually raise the hand, write messages in a chat, and ask the teacher to activate the video and audio in case they want to answer to some questions or ask questions and then participate in the discussion. Figure 2

[1]https://www.adobe.com/products/adobeconnect.html

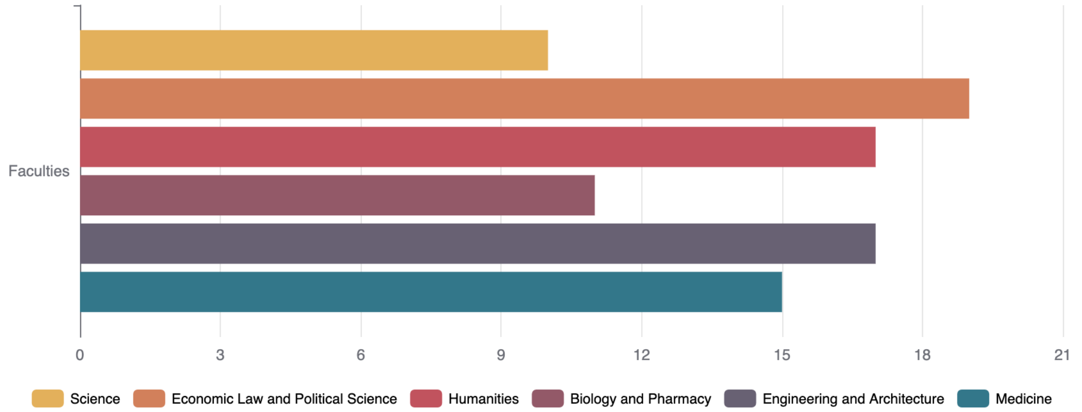


Figure 1: Bachelor's and Master's degree programmes per faculty covered by our study (89 in total).

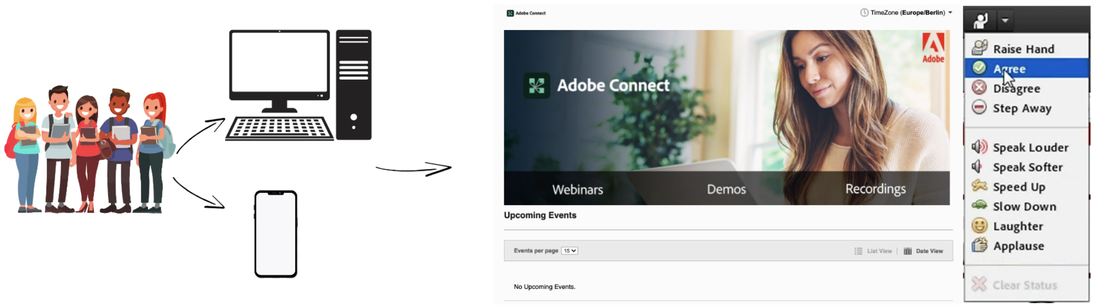


Figure 2: Schema of the online synchronous environment considered in this study.

provides a schematic representation of the Adobe Connect environment.

Data Collection Processes and Format. For each lesson, the date and time of entry and exit of students from the virtual room, the content of the chat, and the reactions and raised hands have been traced. However, for privacy reasons, this paper only deals with the analysis of entry-exit data. This data is structured in folders, following a hierarchical structure aligned with the faculty–degree-course structure of a university. Therefore, the main folder includes six sub-folders, one per faculty. Inside each sub-folder, one Excel file with multiple sheets pertaining to the entry-exit student's records is stored for each course in that faculty. Each sheet includes entry-exit records for a specific lesson of that course. Table 1 provides an example of how a sheet is structured. We collected a total of 6,296 sheets across faculties and degree programmes, with 684 sheets for the Faculty of Biology and Pharmacy, 1,320 sheets for the Faculty of Engineering and Architecture, 947 sheets for the Faculty of Medicine, 760 sheets for the Faculty of Science, 1,362 sheets for the Faculty of Economic, Law and Political Sciences, and 1,223 sheets for the Faculty of Humanities (around 464.5 MB of data). This data covers an entire semester, from March to June 2020.

Attribute	Description	Example of a registered user (Guest)
transcript-id	The id of the transcription and it is unique for each line	4445626093
asset-id	The id related to that session. When a teacher closes the classroom and ends the meeting, another one with a different id is generated at the next login	4445465743
sco-id	The unique id of the virtual classroom	4211809190
principal-id	The unique id of the registered user. For guest logins this field is empty	3727616656 (-)
login	It is the username and for guests it is empty	student@gmail.com (-)
session-name	It is the name field, which is also present for guests	2019_20/40/12345 John Smith (Guest Name)
sco-name	It is name of the virtual classroom. Now it has changed its name and it is called "training object"	CdL Letters - Room 3
date-created	Is the timestamp of entry into the classroom	2020-11-10 08:37:50
date-end	It is the timestamp of leaving the classroom	2020-11-10 08:38:02 (-)
participant-name	It is the registered username field and it is not present for guests	2019_20/40/12345 John Smith (-)
answered-survey	If a survey is proposed, indicate who answered	0

Table 1
Example of data recorded for a lesson of a given course.

4. Student Behavior Characterization

In this section, guided by three research questions, we will analyze student participation patterns throughout the semester, how student groups are related to the courses and if there is a correlation between student participation and class schedule.

Participation Level. In this section, we aim to investigate the participation level of students at faculty level in terms of the average number of students who attended courses in that faculty, to understand whether the trend of the participation level changes over the semester. As mentioned in Section 3, Adobe Connect collects entry-exit logs in forms of CSV files. For each CSV pertaining to a given lesson of a course, we counted the number of students present in each lesson. Cases in which someone from the university staff logged in, empty fields and duplicates were excluded (e.g., some students, due to connection problems, logged in several times, for this reason a single access for each student was considered, not counting the other attempts). This procedure was repeated for all the lessons pertaining to a given faculty, averaging the number of students present in a lesson on a given day across all the lessons delivered for that faculty on that day. Figure 3 shows the participation level over the semester for the six graphs considered faculties. It should be noted that we cannot compare the extent to which

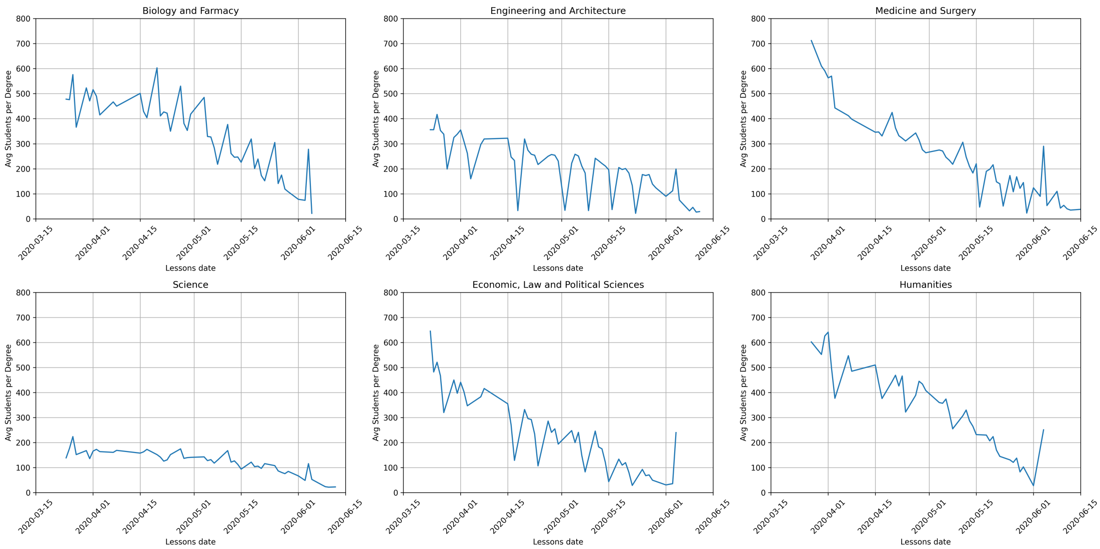


Figure 3: Average number of students per lesson for the six faculties considered in this study.

the student participate in the lessons across faculties, given that each faculty has a different number of students. We are thus interested in the general participation trend in each faculty, individually. For each plot, the time span in which the lessons have been held is represented on the x-axis, ranging from mid-March to early June. The average of the number of students is indicated in the y-axis. All plots are characterized by a strong initial interest, but over time this interest tends to decrease. This decline is very present in three faculties in particular, namely Medicine and Surgery, Economic, Law and Political Science, and Humanities. Science is one of the few that tends to maintain a constant student average, but even here, there is a decline as the months go by[2]. Based on these patterns we understood that there is a need for an adaptive load based on the needs of the varying faculties. This kind of graphs have been generated and analyzed mainly to give support to students and teachers. In particular, if we focus our attention on trends, it is immediately clear that over time, fewer and fewer students connect to the platform. This fact can help teachers to make the lessons more interactive and perhaps increase the attention threshold, trying to limit as much as possible the drop in students over the months. While these plots point out on the possibility of adapting the technological infrastructure based on the changing participation level over time, it does not tell us to what extent these patterns vary for each course in a given faculty. Therefore, the following section analyzes the participation patterns at course level.

Group Modelling. In this section, we are interested in investigating whether there exist core participation patterns in each course and how these patterns vary across courses. To this end, due to the amount of courses included in the dataset and to better shape our analysis, we focused only on the courses of a specific faculty and a specific degree programme, namely

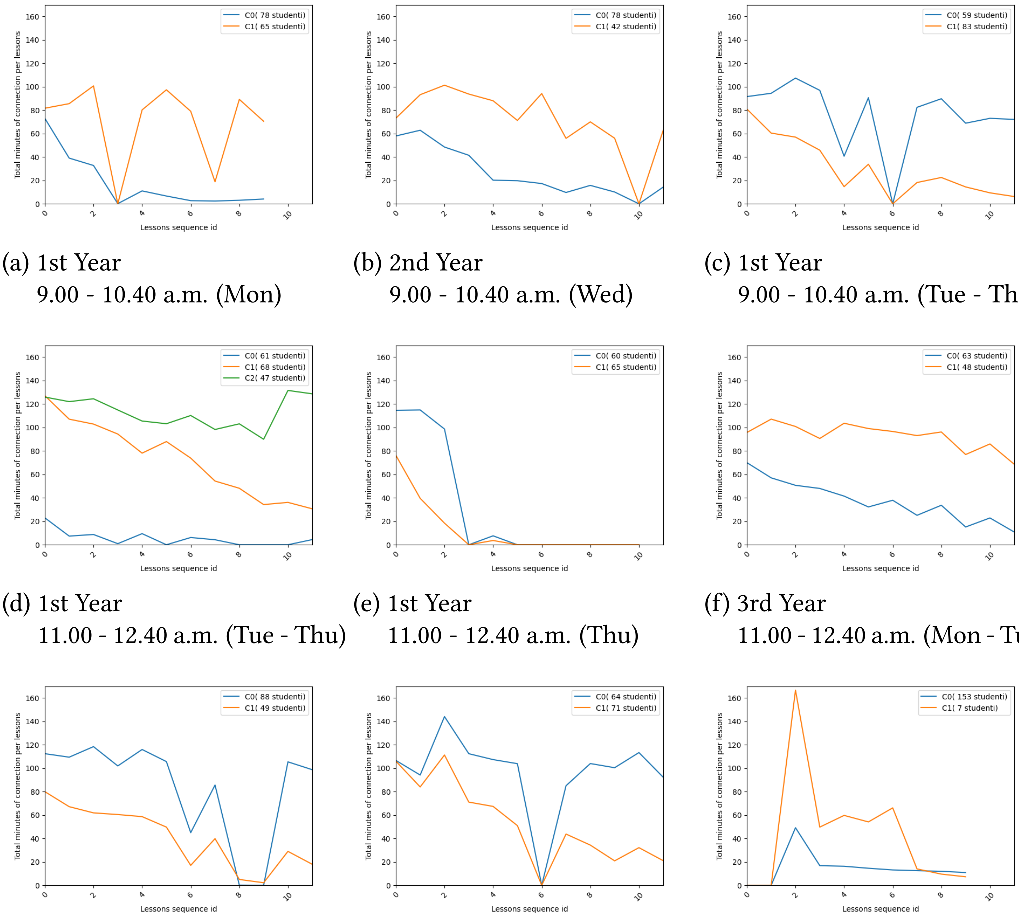
[2]Please, note that the peaks down in the plots are associated to lessons that were planned but not given then. This highlights the need for a more fine-grained analysis, we left as a future work.

Science and its *Bachelor's Degree in Computer Science*. Specifically, we have analyzed the second semester, a period between March and June 2020, including 9 courses (4 courses of the first year, 4 courses of the second year and 1 course of the first year). Given the large number of faculties and consequently of courses, we have selected a degree course with significant patterns that can also be found in other courses. We leave pattern analysis in other courses besides computer science courses to future work. For all the courses in this Bachelor's Degree, all the Excel files with the corresponding entry-exit records have been examined. For each course and lesson in that course, we measure the amount of minutes a student was connected to that lesson during that day, i.e. how long that particular student was connected to the platform. Therefore, for each student in a given course, we computed a N-dimensional vector (where N is the total number of lessons in that course) and each cell contains the amount of minutes that student was connected into the platform for that lesson. Then, we computed the pairwise distances across vectors and feed the pairwise-distance vectors (one per student) of all the students in that course to a *K-Means* clustering algorithm. In order to identify the right number of clusters, we applied the *Elbow* method, using the Silhouette score as a measure of the quality of the clusters. Figure 4 shows the centroids of the clusters identified through the applied methodology for each considered course. It can be observed that, for almost all the courses, two main clusters of students have been identified and these two clusters are well-balanced in terms of the number of students they include. One cluster refers to the students who constantly follow the course over the semester, while the second cluster identifies students whose level of participation decreases over time. This finding is particularly important as a first point for shaping adaptive interventions to students who would need to be motivated along the course. Interestingly, the course in the plot (d) was characterized by three main clusters. For this course, there is a clear subset of students who do not engage with the course at all, from its beginning. Observing this also highlights the need of multiple level of adaptive interventions[3]. Considering the participation patterns with respect to the hour of the day a course is delivered, we observed that there is a relation between the slope of the curves for the cluster who lose engagement and the time of the day, for the courses delivered on the same year of study. This would be important for the future planning of the lessons schedule.

5. Implications and Limitations

There are some limitations on our work, but these do not compromise reliability of this preliminary study. We only used the entry and exit data of the students' connection, without considering the intermediate intervals, so we do not manage cases where a student may have been connected for a short time and have made many intermediate accesses, without perhaps following the lesson for a certain period. Furthermore, due to privacy constraints, no additional information from the students was considered, e.g., all student interventions via chat or

[3]It should be noted that the line associated to each cluster has sometimes a peak towards zero, e.g., in plot (a) in Figure 4 for the lesson with ID 3. By carefully observing the data, it is possible to notice that for that lesson only two students were present. For the peak with the lesson with ID 7, the students connected for a maximum of 30 minutes. This could mean that the platform had problems and the connection was interrupted; another alternative would be that the teacher assigned an exercise and so students no longer needed to stay online. Similar circumstances were present for the other courses.



(a) 1st Year
9.00 - 10.40 a.m. (Mon)

(b) 2nd Year
9.00 - 10.40 a.m. (Wed)

(c) 1st Year
9.00 - 10.40 a.m. (Tue - Thu)

(d) 1st Year
11.00 - 12.40 a.m. (Tue - Thu)

(e) 1st Year
11.00 - 12.40 a.m. (Thu)

(f) 3rd Year
11.00 - 12.40 a.m. (Mon - Tue)

(g) 2nd Year
3.00 - 4.40 p.m. (Mon - Wed)

(h) 2nd Year
3.00 - 4.40 p.m. (Thu - Fri)

(i) 2nd Year
3.00 - 4.40 p.m. (Tue)

Figure 4: Centroids of the clusters identified for the nine considered courses.

system interactions (raised hand) were not included in this study. A limited period was taken into consideration, given that the data refers to three months of lessons (i.e., a semester). This means that, taking into account a longer period of time, different findings could emerge, so it is our task to extend the work and analyze longer periods of time. The process has not considered issues associated with connection problems. This can cause, as we have seen previously, some problems on the analysis, in particular on the entry-exit records. It is not clear to us if users who connected for short periods, less than 20 minutes, decided to leave the lesson or if they had connection problems and were unable to reconnect shortly. Lastly, we used the K-Means algorithm, but other techniques can be adopted to model the data, so it will be our future task to try to model the data using other techniques and make comparisons.

6. Conclusions and Future Work

In this paper we analyze how students interact through the use of Adobe Connect, a synchronous online learning platform, by means of data extracted from the platform logs. Then, we modelled the students' behavior, measuring the participation of students for each lesson across an entire semester on all the courses and faculties of a university. Based on the results:

- The workload associated to the technological infrastructure varies across the semester, with the first period subjected to higher workloads. This finding can be used to adapt the computational resources associated to the platform across time.
- Half of the students often exhibited an strong decrease in attention along the semester, with the majority of them following regularly at beginning, then decreasing considerably. In some degree programs, this behavior is more evident.
- The patterns exhibited in the analyzed data open up to adaptive interventions strategies, such as notifying the teacher in case there are students who are loosing engagement or who have stopped following, in order to understand the reasons and try to keep them engaged.
- There is a relation between the participation level and the time of the day the courses are delivered, with a more rapid decrease in participation for courses delivered earlier in the day. This would be an important element to consider for the future planning of the lessons schedule.

Our work opens up a wide range of future research directions. We plan to enrich the amount of data analyzed, gradually inserting all the logs relating to the lessons that are done at the moment and those done in the first half of 2020 (September-December). Furthermore, we will extend our analysis to other approaches, taking into account chats, student reactions, raised hands. It would be also interesting to make a comparison with students' participation in face-to-face courses in the pre-pandemic period and explore whether the behavioral patterns are similar to those we showed in this paper. Finally, our findings can be of help for teachers to understand the students' participation level during their lessons, given that it is important that a teacher is aware of how many students are present and how they follow the lessons.

Acknowledgments

Roberta Galici gratefully acknowledges the University of Cagliari for the financial support of her PhD scholarship.

References

[1] S. Hrastinski, Asynchronous and synchronous e-learning, Educause quarterly 31 (2008) 51–55.

[2] A. T. Peterson, P. N. Beymer, R. T. Putnam, Synchronous and asynchronous discussions: Effects on cooperation, belonging, and affect., Online Learning 22 (2018) 7–25.

[3] G. Fenu, M. Marras, Leveraging continuous multi-modal authentication for access control in mobile cloud environments, in: New Trends in Image Analysis and Processing - ICIAP 2017 - ICIAP International Workshops, volume 10590 of *Lecture Notes in Computer Science*, Springer, 2017, pp. 331–342.

[4] S. Barra, M. Marras, G. Fenu, Continuous authentication on smartphone by means of periocular and virtual keystroke, in: Network and System Security - 12th International Conference, NSS 2018, volume 11058 of *Lecture Notes in Computer Science*, Springer, 2018, pp. 212–220.

[5] F. Yang, F. W. Li, Study on student performance estimation, student progress analysis, and student potential prediction based on data mining, Computers & Education 123 (2018) 97–108.

[6] M. Cantabella, R. Martínez-España, B. Ayuso, J. A. Yáñez, A. Muñoz, Analysis of student behavior in learning management systems through a big data framework, Future Generation Computer Systems 90 (2019) 262–272.

[7] T. Lubicz-Nawrocka, K. Bunting, Student perceptions of teaching excellence: an analysis of student-led teaching award nomination data, Teaching in Higher Education 24 (2019) 63–80.

[8] Y. Wang, W. M. White, E. Andersen, Pathviewer: Visualizing pathways through student data, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017, pp. 960–964.

[9] P. Prinsloo, S. Slade, M. Khalil, Student data privacy in moocs: A sentiment analysis, Distance Education 40 (2019) 395–413.

[10] L. R. Amir, I. Tanti, D. A. Maharani, Y. S. Wimardhani, V. Julia, B. Sulijaya, R. Puspitawati, Student perspective of classroom and distance learning during covid-19 pandemic in the undergraduate dental study program universitas indonesia, BMC medical education 20 (2020) 1–8.

[11] C. G. Brinton, S. Buccapatnam, M. Chiang, H. V. Poor, Mining mooc clickstreams: Video-watching behavior vs. in-video quiz performance, IEEE Transactions on Signal Processing 64 (2016) 3677–3692.

[12] G. Kőrösi, R. Farkas, Mooc performance prediction by deep learning from raw clickstream data, in: International Conference on Advances in Computing and Data Sciences, Springer, 2020, pp. 474–485.

[13] J. L. Harvey, S. A. Kumar, A practical model for educators to predict student performance in k-12 education using machine learning, in: 2019 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2019, pp. 3004–3011.

[14] R. I. Rashu, N. Haq, R. M. Rahman, Data mining approaches to predict final grade by overcoming class imbalance problem, in: 2014 17th International Conference on Computer and Information Technology (ICCIT), IEEE, 2014, pp. 14–19.

[15] Q. Cao, T. E. Griffin, X. Bai, The importance of synchronous interaction for student satisfaction with course web sites, Journal of Information Systems Education 20 (2009) 331.

[16] D. Leeder, H. Wharrad, T. BChir, Beyond institutional boundaries: reusable learning objects for multi- professional education (2002).

[17] A. Francescucci, L. Rohani, Exclusively synchronous online (viri) learning: The impact on student performance and engagement outcomes, Journal of marketing Education 41 (2019) 60–69.
[18] A. Francescucci, M. Foster, The viri (virtual, interactive, real-time, instructor-led) classroom: The impact of blended synchronous online courses on student performance, engagement, and satisfaction., Canadian Journal of Higher Education 43 (2013) 78–91.
[19] Ü. Çakıroğlu, S. Kılıç, Understanding community in synchronous online learning: do perceptions match behaviours?, Open Learning: The Journal of Open, Distance and e-Learning 35 (2020) 105–121.
[20] X. Yang, D. Li, X. Liu, J. Tan, Learner behaviors in synchronous online prosthodontic education during the 2020 covid-19 pandemic, The Journal of prosthetic dentistry (2020).
[21] T. C. Shoepe, J. F. McManus, S. E. August, N. L. Mattos, T. C. Vollucci, P. R. Sparks, Instructor prompts and student engagement in synchronous online nutrition classes, American Journal of Distance Education 34 (2020) 194–210.
[22] O. B. Adedoyin, E. Soykan, Covid-19 pandemic and online learning: the challenges and opportunities, Interactive Learning Environments (2020) 1–13.
[23] A. Bozkurt, R. C. Sharma, Emergency remote teaching in a time of global crisis due to coronavirus pandemic, Asian Journal of Distance Education 15 (2020) i–vi.
[24] J. Feldman, To grade or not to grade, Educational Leadership 77 (2020) 43–46.
[25] O. Scheuer, B. M. McLaren, Educational Data Mining, Springer US, Boston, MA, 2012, pp. 1075–1079. URL: https://doi.org/10.1007/978-1-4419-1428-6_618. doi:10.1007/978-1-4419-1428-6_618.
[26] D. Dessì, M. Dragoni, G. Fenu, M. Marras, D. R. Recupero, Evaluating neural word embeddings created from online course reviews for sentiment analysis, in: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, SAC 2019, ACM, 2019, pp. 2124–2127.
[27] L. Boratto, G. Fenu, M. Marras, Connecting user and item perspectives in popularity debiasing for collaborative recommendation, Inf. Process. Manag. 58 (2021) 102387.
[28] D. Dessì, G. Fenu, M. Marras, D. R. Recupero, Leveraging cognitive computing for multi-class classification of e-learning videos, in: The Semantic Web: ESWC 2017 Satellite Events - ESWC 2017 Satellite Events, Portorož, volume 10577 of *Lecture Notes in Computer Science*, Springer, 2017, pp. 21–25.
[29] G. Fenu, M. Marras, M. Meles, A learning analytics tool for usability assessment in moodle environments, Journal of e-Learning and Knowledge Society 13 (2017).
[30] J. Byun, D. Pennington, J. Cardenas, S. Dutta, J. Kirwan, Understanding student behaviors in online classroom: data scientific approach, in: 2014 IEEE International Congress on Big Data, IEEE, 2014, pp. 802–803.
[31] J.-L. Hung, K. Zhang, Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching, MERLOT Journal of Online Learning and Teaching (2008).
[32] J.-L. Hung, S. M. Crooks, Examining online learning patterns with data mining techniques in peer-moderated and teacher-moderated courses, Journal of Educational Computing Research 40 (2009) 183–210.
[33] R. Ahuja, A. Jha, R. Maurya, R. Srivastava, Analysis of educational data mining, in:

Harmony Search and Nature Inspired Optimization Algorithms, Springer, 2019, pp. 897–907.
[34] M. W. Rodrigues, S. Isotani, L. E. Zarate, Educational data mining: A review of evaluation process in the e-learning, Telematics and Informatics 35 (2018) 1701–1717.
[35] S. Kausar, X. Huahu, I. Hussain, Z. Wenhao, M. Zahid, Integration of data mining clustering approach in the personalized e-learning system, IEEE Access 6 (2018) 72724–72734.
[36] D. Ding, J. Li, H. Wang, Z. Liang, Student behavior clustering method based on campus big data, in: 2017 13th International Conference on Computational Intelligence and Security (CIS), IEEE, 2017, pp. 500–503.
[37] M. Durairaj, C. Vijitha, Educational data mining for prediction of student performance using clustering algorithms, International Journal of Computer Science and Information Technologies 5 (2014) 5987–5991.
[38] R. Asif, A. Merceron, S. A. Ali, N. G. Haider, Analyzing undergraduate students' performance using educational data mining, Computers & Education 113 (2017) 177–194.