Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics

Daniel Jaramillo-Morillo[1]* , José Ruipérez-Valiente[2], Mario F. Sarasty[1] and Gustavo Ramírez-Gonzalez[1]

*Correspondence:
dajaramillo@unicauca.edu.co
[1]Department of Telematics, University of Cauca, Popayán, Colombia
Full list of author information is available at the end of the article

Abstract

Massive Open Online Massive Open Online Courses (MOOCs) have been transitioning slowly from being completely open and without clear recognition in universities or industry, to private settings through the emergence of Small and Massive Private Online Courses (SPOCs and MPOCs). Courses in these new formats are often for credit and have clear market value through the acquisition of competencies and skills. However, the endemic issue of academic dishonesty remains lingering and generating untrustworthiness regarding what students did to complete these courses. In this case study, we focus on SPOCs with academic recognition developed at the University of Cauca in Colombia and hosted in their Open edX instance called Selene Unicauca. We have developed a learning analytics algorithm to detect dishonest students based on submission time and exam responses providing as output a number of indicators that can be easily used to identify students. Our results in two SPOCs suggest that 17% of the students that interacted enough with the courses have performed academic dishonest actions, and that 100% of the students that were dishonest passed the courses, compared to 62% for the rest of students. Contrary to what other studies have found, in this study, dishonest students were similarly or even more active with the courseware than the rest, and we hypothesize that these might be working groups taking the course seriously and solving exams together to achieve a higher grade. With MOOC-based degrees and SPOCs for credit becoming the norm in distance learning, we believe that if this issue is not tackled properly, it might endanger the future of the reliability and value of online learning credentials.

Keywords: Massively open online course, Academic dishonesty, Small private online courses, MOOC, SPOC, Learning analytics, Cheating

Introduction

The evolution and increasing use of communication technologies have generated new and very popular learning modalities such as the Massive Open Online Courses (MOOCs). These courses have expanded broadly in e-learning, experiencing a rapid development and receiving considerable attention from many institutions and universities, presenting the MOOC model as a new disruptive technology within the educational ecosystem

(Chen et al. 2014; Stephen 2012). MOOCs represent an opportunity for universities to expand their coverage and attract more students into their campuses. They are presented as the new path to expanding knowledge via continuous learning, university innovation, employability and the sustainable development of mass learning scenarios.

Many universities developed a large portfolio of MOOCs that required a development of new competences at a university level (Kloos et al. 2014). After developing such skills, institutions are re-using these efforts to incorporate these courses into higher education by transitioning from MOOCs to other on-campus private formats, such as the ones known as Small Private Online Courses (SPOCs) and Massive Private Online Courses (MPOCs) (Fox 2013; Guo 2014; Zhou et al. 2016). The introduction of these new online models into universities has also generated innovation in the teaching practices, with blended, hybrid and flipped classroom methodologies becoming more and more common (Rodríguez et al. 2017; Wang et al. 2016). This transition might bring long-term benefits as some studies have shown better learning outcomes and motivation towards these blended methodologies when compared to traditional learning (Tseng and Walsh 2016).

However, despite the wide agreement among stakeholders in recognizing MOOCs as an important opportunity to improve educational practices around the world, there are still many challenges that make it difficult to incorporate these courses into the university environment and to provide a valid and trustworthy academic recognition for them (Jobe 2014; Witthaus et al. 2016). One of the challenges encountered is the constant uncertainty about students' academic honesty related to verification of identity (impersonation), verification of ownership of their work (fraud), supervision of examinations (cheating during examinations), etc. This is an endemic problem in distance learning since it became popular (Lanier 2006), but nowadays with online learning becoming more popular through the numerous available options, such as MOOC-based degrees (Reich and Ruipérez-Valiente 2019) or SPOCs for credit in universities (Kaplan and Haenlein 2016), the problem becomes more critical as there is no guarantee that students are being honest. Literature reports that in most cases they are using dishonest strategies to obtain the academic credits offered, which in turn generates a constant concern to allow MOOCs and derived courses to be recognized academically (Abramovich et al. 2013; Jobe 2014; Sandeen 2013). In addition, current learning platforms do not have the necessary technological tools that allow the identification of dishonest student behaviors in order to mitigate academic dishonesty in this type of courses (Ruipérez-Valiente 2018; Ruipérez-Valiente et al. 2017).

Therefore, in this article we work towards the development of algorithmic approaches to detect dishonesty behaviors in a SPOC environment that could improve the academic for credit trustworthiness of these courses. We apply a learning analytics approach for the implementation and evaluation of the method. The context of the research is within SPOCs with academic recognition developed at the University of Cauca (Colombia) and hosted on their own Open edX instance, namely Selene Unicauca. The algorithm performs an analysis of students' behavior with the exams (submission time and exam responses obtained through the interaction of the student with the learning platform), providing as output a number of indicators that can be easily used to identify students that denote clear dishonest behaviors and practices through the sharing of responses with other peers. Our motivation is that the solutions obtained through these data-driven approaches, could eventually be the basis of a system for academic dishonesty

deterrence (Corrigan-Gibbs et al. 2015), and in this way facilitate the incorporation of MOOCs/SPOCs for credit into the university environment.

More specifically, we have two main objectives:

1. To develop a learning analytics approach that can be used within SPOCs to detect students committing dishonest behaviors by sharing responses with peers.
2. To perform a comparison analysis of the levels of engagement with courses contents and learning outcomes between those students detected as dishonest and the rest of students.

The remainder of the paper is organized as follows. "Related work" performs a review of the related work in the area of MOOCs and academic dishonesty. "Methodology" presents the methodology of the study, describing the context, the algorithm and data collection. "Results" describes the results regarding the detection of dishonest students and their behavioral characterization. Finally, "Discussion" discusses results comparing with previous work and "Conclusions and future work" sections concludes the paper.

Related work

MOOCs emerged as a new educational paradigm where large numbers of students are able to take courses for free. One of the initial goals was to facilitate access to knowledge and therefore these courses were not targeting a university environment in order to be certified or recognized academically as part of the curriculum. However, MOOC providers struggled to find financial sustainability without a clear business model (Liyanagunawardena et al. 2015), and eventually most of the big and open MOOC providers have started pivoting towards online degrees in partnership with universities or professional training programs with companies (Reich and Ruipérez-Valiente 2019). Another one of the directions that the original MOOCs are taking is the incorporation of courses following similar pedagogical guidelines and contents into universities, where in some cases the completion of these courses might provide an academic recognition (Jobe 2014; Witthaus et al. 2016).

The adoption of the MOOC model in higher education institutions and its incorporation into educational programs represents a technological and pedagogical challenge, which is being resolved thanks to strategies like SPOCs (Fox 2013; Kloos et al. 2014) and MPOCs (Guo 2014; Mutawa 2016). SPOCs and MPOCs are variants of MOOCs that are characterized by being limited in access (private courses) and therefore also in size, but they still have a wider scope of participation than any conventional online course (Cabero et al. 2014; Kloos et al. 2014), allowing universities the opportunity to expand their coverage and reach more students at the same time. These courses in new format are presented as a new path for the expansion of knowledge, university innovation, employability and sustainable development scenarios of massive learning, so many universities are strongly working on their incorporation into their curriculums and academic offer (Aguaded and Medina-Salguero 2016; Arturo Amaya and Alvarez 2015; Zirger et al. 2014).

Academic recognition of courses has a significant impact on student behavior. The act of signing up and paying to access a verified certificate track, that would allow a student to receive a certificate, can have an effect on students' behavior acting as a commitment device (Littenberg-Tobias et al. 2020). Several studies have shown that those courses where students have the chance to obtain academic recognition, reduce the high

dropout rates of MOOCs (Halawa et al. 2014). In fact, in all MITx and HarvardX MOOC portfolio, the general completion rates were about 5%, but they would rise up to 50% for verified learners (Reich and Ruipérez-Valiente 2019). Academic recognition encourages students to pay more attention to their progress in the course, encourages students to self-regulated their learning better while being more participatory and engaged, etc., showing overall indications that the recognition is a strong motivator to complete the course (Jaramillo-Morillo et al. 2017; Jaramillo-Morillo et al. 2017).

The introduction of an extrinsic motivator like credit recognition, can shift students' behavior towards fulfilling the requirements to achieve the pass the course while losing interest in the actual learning (Lei 2010). Therefore, we would see that the behavior of students change in different aspects and how they interact with the contents. For example, previous work identified that in courses with academic recognition, students invested higher levels of activity in the platform prior to the summative evaluations that were scheduled as part of the course (Jaramillo-Morillo et al. 2017). This study observed that after the introduction of the academic recognition, students' motivation and main objectives moved towards the passing grade that would allow them to get this recognition.

The combination of academic recognition and online exams can lead to dishonest behaviors in order to obtain better grades with less effort (Bao 2017; McGee 2013). Students will often try to game the system in these interactive learning environments in order to pass a course without actually learning the contents (Baker et al. 2008). Over the last few years, researchers have reported a number of studies in the literature exploring the different dishonest strategies or dishonest behaviors in MOOCs. The most frequent methods include searching for answers shared by other participants, impersonation, conducting exams in the company of experts in the subject, obtaining correct answers by registering several accounts, asking questions or searching forums for questions and answers related to the subject, among others (Bao 2017). In Watson and Sottile (2010), the authors identified anonymity, poor teacher-student interaction, lack of time, or class difficulty as triggers for an increased likelihood of cheating in online courses. Another study also found that some of the main motivations to cheat in an online course are the lack of barriers to copying, wanting to pass the course and getting better grades (Backman 2019). In Watson and Sottile (2010), they conducted a comparative study of cheating in face-to-face courses and online courses through student surveys. This study found that cheating in online courses is more unrestrained than cheating in live classes, the data showed that students were significantly more likely to get answers from others during an online test or exam. Students were found to have a higher rate of dishonesty in online courses related to getting answers from someone during a test (23.3% to 18.1%) and using instant messaging during a test (4.2% to 3.0%). Interestingly, students reported that it is easier to share their answers in online courses and that they were more likely to be caught cheating on a face-to-face test (4.9% to 2.1%).

Academic dishonesty causes numerous issues within the context of online learning. The first one is that it generates uncertainty regarding if students are passing the course applying any of the aforementioned dishonest methods or are actually learning the contents, hence making very complicated that educational institutions and industry can recognize in the same way a MOOC certificate when compared to traditional face to face courses (Abramovich et al. 2013; Jobe 2014). This has limited or delayed the incorporation of MOOC-based courseware into the university environment. The second issue is that

several studies report that cheating behaviors lead to poor learning (Palazzo et al. 2010). One last issue, and perhaps less obvious, is that since cheating represents an outlying behavior, these data can be prone to bias when learning analytics models are built, which can be systematically impacting all MOOC research reported during the last decade (Alexandron et al. 2019).

Most MOOC providers and other online learning platforms have not made strong efforts to counter this issue, with very few initiatives incorporating mechanisms for identifying dishonest student behavior. The most noteworthy examples were developed in Coursera and edX, which have implemented mechanisms to proctor the exam of the student through webcams (Coursera 2013). However, the setup is not straightforward representing a complex process for the student and it also has a series of minimum requirements in terms of hardware and Internet connection bandwidth, which might represent solid barriers in rural areas or less affluent countries like Colombia. In addition, these processes are not yet fully automatic and therefore not scalable.

Previous work presented (Bao 2017; Northcutt et al. 2015; Ruipérez-Valiente 2018) algorithms for the detection of CAMEO (Copying Answers using Multiple Existences Online). CAMEO is one of the reported methods of cheating in online environments, where harvester (fake) accounts are used to get correct answers using the feedback of the system, that are then used by a master account to achieve the minimum grade that allows the student to get a certificate. The authors present an algorithm to identify and tag the submissions that were cheated using the CAMEO method; these algorithms are based on several heuristics and also use the IP of the submissions. In Ruipérez-Valiente et al. (2017) they present a machine learning algorithm that detects CAMEO without using the IP, which implemented a supervised machine learning method using several characteristics of the submission, student and design of the problem to predict the likelihood of a submission being completed using CAMEO. However, CAMEO is only one specific method among the ones reported in the literature and used by students to commit course fraud. Fortunately, CAMEO is controlled more easily in private environments such as SPOCs and MPOCs since each student will have access to only their personal account, and they cannot create more.

Another previous work also implemented a method that does not rely on the IP of the submission (Ruipérez-Valiente et al. 2017). In this case, the study implements and algorithm that detects "invisible" collaborative links between students in online learning environments. Specifically the work presents a method developed to detect links between students based on the temporal closeness when submitting their quizzes (one of the criteria used by us in this study). The results show that most students were grouped in pairs, although some larger communities submitting their responses together were also detected. The study also found that close submitters needed significantly less activity with the contents of the course to get a certificate of completion in two MOOCs from Coursera. Their results confirmed that detected close submitters were carrying out some dishonest collaboration or even engaging in other unethical behaviors like CAMEO, which facilitated their access to a certificate without interacting with the courseware. However, the authors finished the paper indicating that more work is needed in the future to characterize students' behaviors based on the interaction data with the platform to determine the specific inappropriate behaviors that students are committing.

All the previous methods have been implemented and tested in open MOOCs, in this study we want to continue this line of research in a by implementing a fraud detection method that in the context of a private online courses with academic recognition. To the best of our knowledge, is the first study reporting a case study applying a data-driven method to detect academic dishonesty in a SPOC or MPOC for academic credit.

Methodology

This section describes the methodology in three subsections. The first subsection presents the specificities of the context and the case study, the second the data collection and the last one the method that is implemented and applied for the identification of students suspected of academic dishonesty.

Description of the context and case study

As a case study we have used the following two courses: "Introduction to Lean Startup Entrepreneurship" (*Lean Startup Course*) and "Comprehension of Argumentative Texts" (*Texts Course*) of the University of Cauca in Colombia, offered via the Selene learning platform, an instance of Open edX that is maintained by the university. Since its inception, these courses have been considered internally as SPOCs as these are private courses, these have a number of students four times more students than a traditional classroom course but are not massive as the tutors has managed to provide support to each of the students in the course. In addition, uses contents, courseware and schedule similar to the ones frequently found in MOOCs. For this study, we selected the last iterations of the courses that was offered during the first semester of 2019.

These courses are offered as elective classes within the University of Cauca and are part of the Social and Human Integral Training (FISH, Formación Integral Social y Humana) component that are mandatory to take for all university students. The courses are recognized academically and are valid for two credits within the undergraduate training programs. In addition, access to the courses is controlled by the administrator of the platform, hence making the possibility of creating additional accounts and performing CAMEO not feasible.

The fundamental purpose of the *Lean Startup Course* is to introduce students to practical contents of one of the most successful methodologies for enterprise development in recent times: Lean Startup. The fundamental purpose of the *Texts Course* is to improve students' communicative and textual skills. These courses have been organized through thematic units divided into modules, with a total of five modules. The evaluation was carried out according to the guidelines of the students' regulations, with an evaluation for each course module. The evaluations are operationalized through online exams with multiple choice questions that have a single correct answer.

Each evaluation consisted of 12 to 30 questions with the possibility of submitting the exam on two different dates. Each exam was open and available for students for only a time window of 60 minutes to minimize the chances of fraud or sharing responses. In addition, there was a trial tests, which was not taken into account for the courses grade. The trial tests was carried out to help students to become acquainted with the examinations mechanics in the courses.

Data collection

In the selected courses, we had 192 students enrolled (92 *Lean Startup Course* students and 100 *Texts Course* students) and it had an overall completion rate of 76%. From the 192 enrolled students, 147 (76%) of them presented the final exam and 96 (50%) students passed the courses. We have collected all the interactions of these students with the platform. In order to identify those students that are solving the exams together and passing the responses to their friends, we follow a data-driven approach and we need then to obtain all the interaction logs with the Open edX the learning platform. We follow a previous reported technical setup as it is described in Fig. 1 (Jaramillo-Morillo et al. 2017).

Using this setup, we obtain all the tracking logs from the Open edX instance, that include all the events that students performed within the learning environment, including the answers to exams, the date and the time of the submission of the students and so on. A total of 297355 student interactions with the learning platform were obtained during the period from April to July 2019. We decided to include those students that completed more than half of the course exams (i.e., at least they submitted three exams), to avoid noisy measures; therefore, the analysis includes data from those 147 students (69 *Lean Startup* students and 78 *Texts Course*) students.

Identification of students suspected of academic dishonesty

The method that we implement contains ideas of previous work (Ruipérez-Valiente et al. 2017) and incorporate new ones. We have used two criteria to identify students who are suspected of fraud: students who have very similar exam responses and students who submit exams very close in time. In order to perform the analysis based on this criteria, the following mathematical nomenclature is used: We have N denoting the total number of students and M denoting the total number of variables. We also have N vectors, one for each of the students so that:

$$\vec{sp} = \left[\begin{array}{cccc} sp_{i,1} & sp_{i,2} & \cdots & sp_{i,M} \end{array} \right], i \in \{1 \cdots N\} \tag{1}$$

SP_i contains all M variables for student i in the course. In case the student did not submit an exam, the variable will be $sp_{ij} = N.A.$

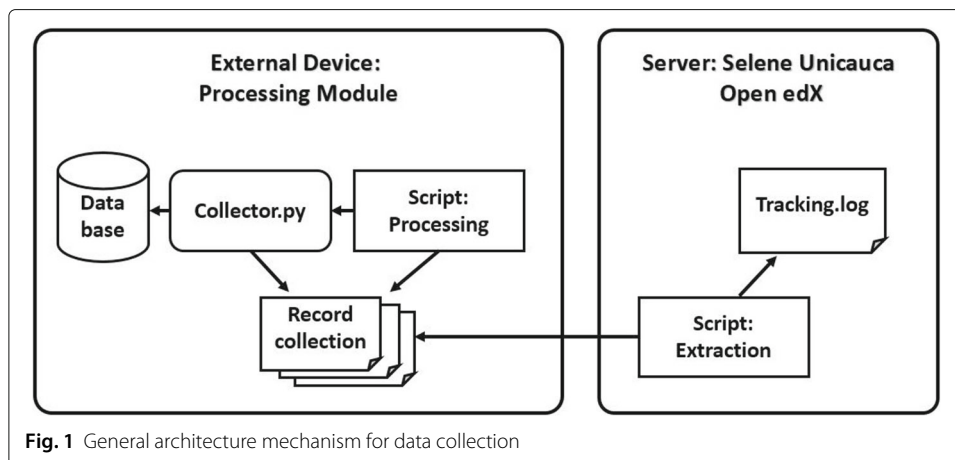**Fig. 1** General architecture mechanism for data collection

This defines the $SP \in N^{NxM}$ matrix as:

$$SP = \begin{pmatrix} \vec{sp}_1 \\ \vec{sp}_2 \\ \vdots \\ \vec{sp}_N \end{pmatrix} = \begin{pmatrix} sp_{1,1} & sp_{1,2} & sp_{1,3} & \cdots & sp_{1,M} \\ sp_{2,1} & sp_{2,2} & sp_{2,3} & \cdots & sp_{2,M} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ sp_{N,1} & sp_{N,2} & sp_{N,3} & \cdots & sp_{N,M} \end{pmatrix} \tag{2}$$

Where $SP_{i,j}$ represents the value of the variable j for the student i. That is, for the method of identifying students by similarity of responses, $SP_{i,j}$ represents the response j submitted by a student i. In the case of students identified by submitting the exams very close in time, $SP_{i,j}$ represents the timestamp of the exam j presented by student i.

Then, we define $DS \in R^{NxN}$ a dissimilarity matrix as follows:

$$DS = \begin{pmatrix} ds_{1,1} & ds_{1,2} & ds_{1,3} & \cdots & ds_{1,N} \\ ds_{2,1} & ds_{2,2} & ds_{2,3} & \cdots & ds_{2,N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ ds_{N,1} & ds_{N,2} & ds_{N,3} & \cdots & ds_{N,N} \end{pmatrix} \tag{3}$$

Where each $ds_{i,j}$ entry represents the difference between the values of students i and j based on a dissimilarity metric. We are implementing these matrices for the following two metrics:

- The responses selected for each item of every exam in the course.
- The submission timestamps for each exam in the course.

We explain next with details how is each distance metric computed.

Distance matrix based on comparison of exam responses

Each DS matrix element is calculated by a dissimilarity function $diss(sp_i, sp_j) \in R$ that operates on student response vectors. In this case the dissimilarity function is based on the Simple Matching Coefficient (SMC).

$$diss_{SMC}\left(\vec{sp}_i, \vec{sp}_j\right) = 1 - \frac{\text{Matching count betwen} sp_i \text{ y} sp_j}{M} \tag{4}$$

In this way, when sp_i and sp_j have exactly the same answers, the distance will be 0 and when sp_i and sp_j do not match in any of the answers, the distance will be 1.

Distance matrix based on exam submission times

In this case, the calculation of the DS distance matrix is carried out using the Mean Absolute Deviation (MAD). Where the distance between two vectors is calculated as the average of the absolute differences between the two vectors.

$$diss_{MAD}\left(\vec{sp}_i, \vec{sp}_j\right) = \frac{1}{M} \sum_{k=1}^{M} \left| sp_{i,k} - sp_{j,k} \right| \tag{5}$$

In this way, the distance $ds_{i,j}$ between two students i and j is the average of the time distances of each submission. The closer the submissions of two students are in the time, the smaller the distance between them.

Results

This section presents the results of the study, with first subsection focusing on the identification of students suspected of being dishonest based on our method, and the second subsection characterizes the engagement and learning outcomes of those students.

Students identified as suspected of academic dishonesty

We use the method as described in previous section to identify those students that are suspected of academic dishonesty. The output of the method are two cohorts, one for each one of the proposed dissimilarity metrics. Therefore, we have one cohort for the metric regarding the similarity in responses and one for the metric submitting their exams very close in time.

In order to identify students with high similarity in their responses, we take all the responses of the 5 modules, making a total of 75 responses submitted per student in the *Lean Startup Course* and 142 responses submitted per student in the *Texts Course*. Then, we use the array with 75 responses for the *Lean Startup Course* and we use the array with 142 responses for the *Texts Course* to calculate the distance matrices. Finally, we take all pairs of students with a 90% similarity in their responses throughout the entire course. In practice, that means that we selected those pairs of students that have a distance $ds_{i,j}$ less than 0.1 in the distance matrix. Based on this metric we identify 15 students for the *Lean Startup Course* and 11 students for the *Texts Course* that are suspected of fraud.

On the other hand, we also identify other cohort of students based on the time closeness of their submissions. In this case, for the computation of the distance matrix we use the five timestamps when the exam of each module was submitted. Finally, students detected as suspects of fraud were selected based on an absolute mean distance of less than five minutes. That would mean that all exams were submitted on average between these students less than five minutes away. In this case, a cohort of 17 students for the *Lean Startup Course* and 15 students for the *Texts Course* were detected as suspected of fraud.

It stands out how the same students were identified in both metrics except for few students, despite having used completely different and unrelated metrics. In order to be conservative, we intersect the two sets of students keeping only the 15 students that are in both cohorts for the *Lean Startup Course* and 10 students that are in both cohorts for the *Texts Course*, making then both the similar responses and time closeness mandatory conditions.

We can visualize how the students suspected of academic dishonesty are organized in Fig. 2. This visualization shows a dendrogram for each one of the similarity metrics, which is a tree diagram with the cohorts formed by creating clusters of observations based on their levels of similarity. The level of similarity is measured on the vertical axis and the observations are specified on the horizontal axis.

In the case of the dendrogram of students identified by the similarity of their responses (Responses Cluster Dendrogram), the vertical axis shows the similarity of the students with a measure between 0 and 1. Where a 0 indicates an exact match of the responses and a 1 indicates a difference of 100% in their responses. Please note how the students have been grouped in couples or triads according to the similarity they have in their responses.

For the cohort of students identified by time closeness (Time Cluster Dendrogram) the vertical axis shows the proximity of the students' submissions measured in hours.
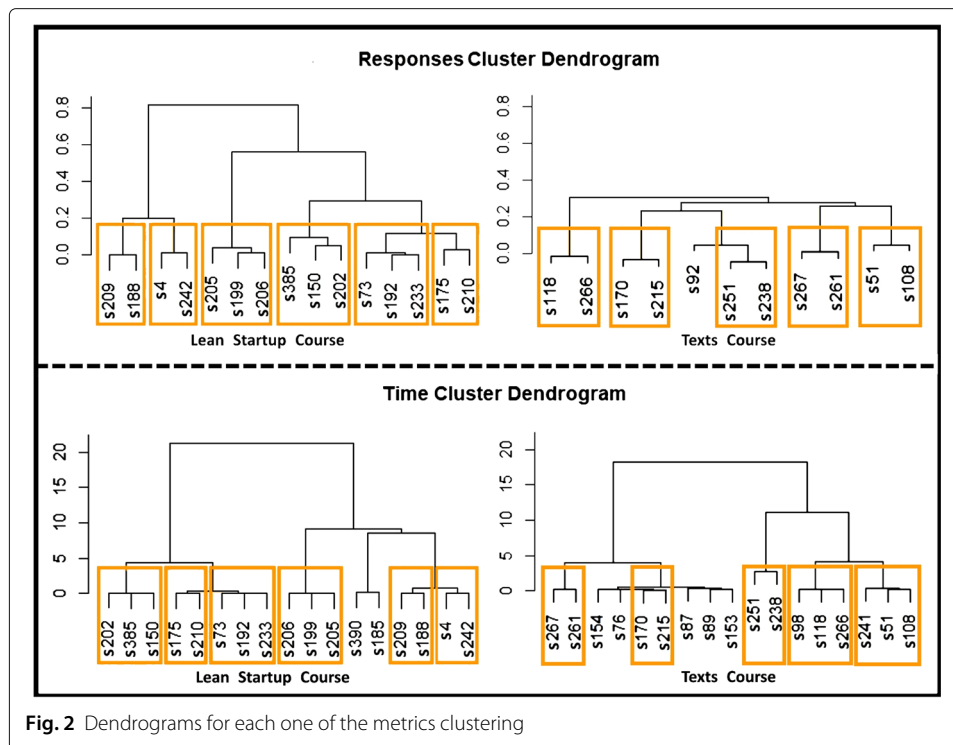
Fig. 2 Dendrograms for each one of the metrics clustering

The groups that were created correspond to exactly the same groups created by the analogous grouping in Fig. 2 based on the similarity of the responses, despite the metrics are completely different. Table 1 shows the similarities in responses and the closeness of submissions in time for each one of the groups generated. As we can see all the groups have very high similarities in their responses and a very low distance in time of their submission. We can interpret each row in the following way: The row with students s205, s199 and s206, would indicate that this group of students have 96% of their responses in common (from the total of 75 questions in the course), and that the average distance between the timestamps when they submitted each exam (from the five exams in the course) is only 11 seconds. Therefore, the results are self-evident showing values that indicate that these groups of students are working together. On the other hand, these three students, besides being identified by our algorithm, were also identified as cheaters by the course instructor, since these students individually complained to the teacher about the low grade received for the module 4 exam alluding that the platform had graded them wrong. Then, the instructor reviewed their answers and found that their wrong answers were identical. This is positive confirmation that the algorithm is working correctly.

In addition, the row with students s251 and s238, indicates that this group of students has 100% of their answers in common (out of the total of 142 questions in the course), and that the average final grade was 74 on a scale from 0 to 100. So these students made exactly the same mistakes on 37 questions with four possible answers. These are just some detailed analysis examples, but we can establish similar conclusions from the rest of cases.

Analysis of behaviour and engagement

This second subsection of results focuses on analyzing the behavior of the cohort of students identified as suspected of academic dishonesty and compare it with the behavior

Table 1 Similarities in responses and the closeness of submissions in time

Students	Similarity in responses	Proximity in time	Average grade
Lean startup course			
s209, s188	100%	0:02:05	92
s4, s242	100%	0:02:36	72
s175, s210	97%	0:00:07	96
s385, s150, s202	91%	0:02:53	89
s73, s192, s233	99%	0:03:18	92
s205, s199, s206	96%	0:00:11	55
Texts course			
s118, s266	97%	0:01:59	67
s170, s215	99%	0:00:05	71
s251, s238	100%	0:01:35	74
s261, s267	96%	0:04:56	70
s51, s108	93%	0:01:40	69

of the rest of the students. In order to perform such comparison, we divide all the students that were included in the analysis considering two criteria: if the student was detected as a suspect and if the student passed the course. Then, a total of three cohorts are created: Identified Suspects, Regular Students (Pass), and Regular Students (Fail).

As part of the first analysis, we analyze differences in grade by cohort, Fig. 3 shows a boxplot visualization with the grade distribution by cohort.

Students can pass the course with a grade higher than 60 on a scale from 0 to 100. The cohort of identified suspects had a 100% pass rate, while the rest of the students had a 72% success rate for the *Lean Startup Course* and 51% success rate for the *Texts Course*. Note that the best grades are achieved by the cohort of students who were identified as being suspected of academic dishonesty. With an average course grade of 84 compared to 72 for the rest of students that passed and of 40 for students that failed for the *Lean Startup Course* and with an average course grade of 68 compared to 63 for the rest of students that passed and of 41 for students that failed for the *Texts Course*

Next, we analyze the interaction and behavior of students with the courses in order to find differences between cohorts. Figure 4 shows the number of engagement and interaction metrics that students had with course contents.

Overall, we can observe that in both courses the cohort of students that did not pass, had slightly lower levels of activity than the rest, but that the differences between students
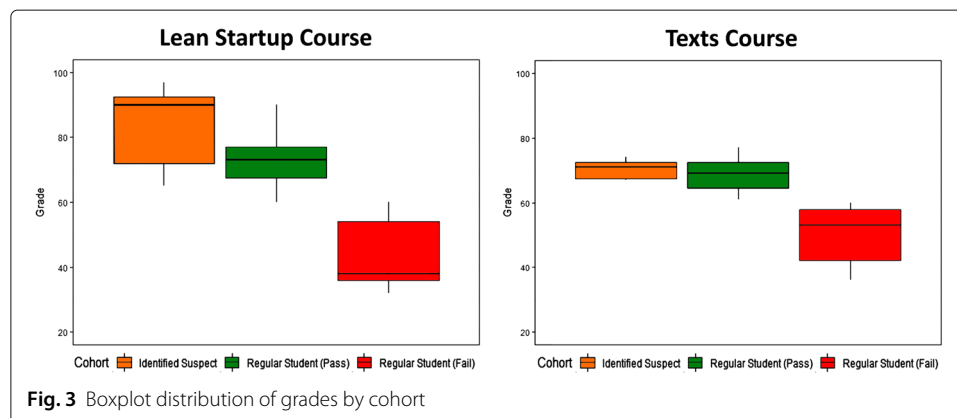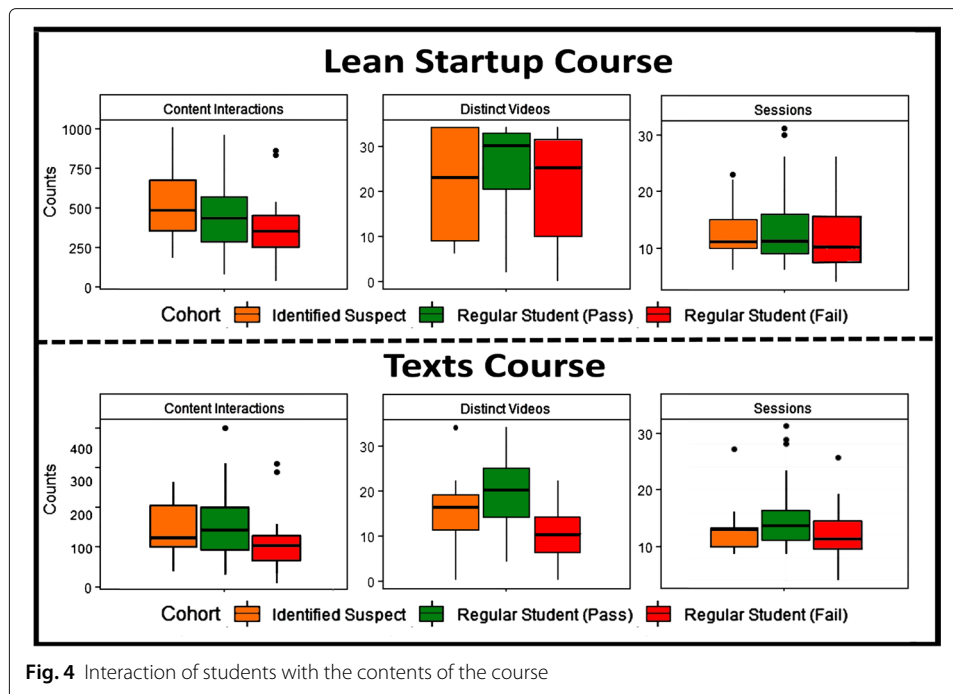
Fig. 3 Boxplot distribution of grades by cohort

Fig. 4 Interaction of students with the contents of the course

that passed the course legitimately and those who are suspected of academic dishonesty are not that noteworthy. Also, we found that the number of interactions with the content of the *Lean Startup Course* is higher for students suspected of fraud in contrast to the *Texts Course*, where the students with the highest number of interactions with content are the students who legitimately passed the course. One explanation may be that students identified as suspects of fraud in *Lean Startup Course*, interact much more with content in search of test answers because it is a course heavy on contents. This does not happen in the *Texts Course* because it is a more practice-oriented course. On the other hand, we noticed that the different videos viewed are higher for regular students who passed the course. We believe that these students have a higher interested in learning and not only in getting a good grade.

We ran *t*-tests for each one of the metrics comparing the cohorts of identified suspects with the regular students that passed, and we did not find any statistically significant differences between the cohorts. Therefore, even though we do observe some differences in the means of some indicators when comparing both cohorts, these differences are not statistically significant (probably due to the sample size).

Finally, Fig. 5 shows the number of interactions over time with the number of events per day during the development of the course. This visualization can help to know if the activity of each cohort was more or less spaced over time, and the behavior of each cohort with respect to the scheduled exam dates. The overall patterns seem to indicate that students that were identified suspected of academic dishonesty concentrate their activity more on the exam dates than the rest of the cohorts, which have an activity a bit more spaced over time in most cases. We can match the peaks with the exam dates (see module below the *x*-axis), and as a reminder we would like to clarify that most exams had two possible dates. For the *Lean Startup Course* the two dates were separated by one week and for the *Texts Course* the second date was the day after the module exam.
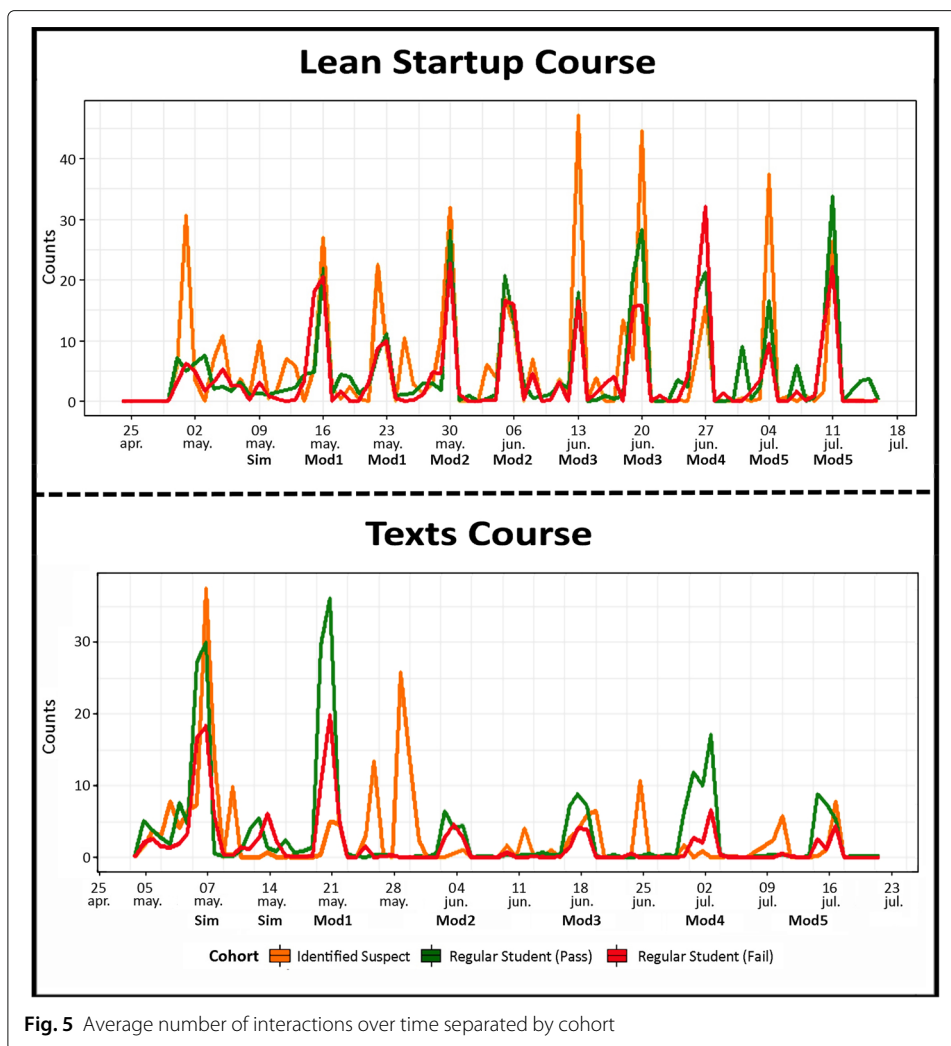
Fig. 5 Average number of interactions over time separated by cohort

For the *Lean Startup Course*, the average number of interactions per day for the cohort suspected of fraud is 6.05 with a variance of 108.9, while the average number of actions for the regular students that passed the course is of 4.69 with a variance of 51.9. Therefore, the students suspected of fraud were more active with the platform, and their activity was concentrated in fewer days, more specifically on the exam dates, whereas the students that passed the course without being academically dishonest were a bit less active, and their activity was more spaced during the entire course timeline.

In contrast, for the *Texts Course*, the average number of interactions per day for the cohort suspected of fraud is 2.71 with a variance of 33.76, while the average number of interactions for the regular students that passed the course is 3.48 with a variance of 49.93. We see that in this case the regular students that passed the course are more active than the students that are suspected fraud and also that their activity is less concentrated on the exam dates. This may be because the *Texts Course* is more practical and the answers to the exams are not explicit in the course contents as opposed to the *Lean Startup Course* which is a more theoretical course. On the other side we have found that students identified as suspects of fraud, have interactions previous to the exams, just when the tutor was releasing a necessary reading to solve the exam.

Discussion

Several previous studies have used algorithmic approaches to detect cheating in online environments (Bao 2017; Corrigan-Gibbs et al. 2015; Northcutt et al. 2015; Ruipérez-Valiente et al. 2017; Ruipérez-Valiente et al. 2017). However, all of these studies were focused on MOOCs, which have important key differences with our context. In these MOOCs, students could opt-in to receive a course certificate if the course was completed successfully, by paying the enrolment fee to access the verified track; however, these courses were not providing credit for a degree. Another key difference due to the context is that the students of MOOCs can create several accounts, and so one single physical person can be handling more than one account which could be used to exploit the feedback properties of the system. However, in this case scenario account creation is centralized as part of the University of Cauca system, hence students can only have one account, and are then they are forced to rely on working along their friends. Furthermore, the anonymity of MOOC environments is helpful to undertake such unethical activities, however, students of Selene Unicauca platform know that their accounts are linked to their academic record. While the algorithmic detection methodology is similar, to the best of our knowledge this is the first time that this kind of data-driven methodologies relying on tracking data logs have been used for courses with academic recognition.

Our algorithmic approach relying on the similarity between the solutions and submission timestamps have detected a total of 15 students for the *Lean Startup Course*, six of them organized in three dyads, and nine of them organized in three triads and 10 students for the *Texts Course*, organized in 5 dyads. The dynamics of these different communities of students, can represent different behaviors, in some cases they might have a similar workshare and really perform a joint effort, in others one student might lead the workload and pass the responses to their friends. Previous work (Alexandron et al. 2019) indicated different profiles of CAMEO students, in some cases performing more deliberate cheating while in others using it as a backup or help plan. While we have not delved into analyzing the different behavioral dynamics, we consider it a promising line that should be explored in depth. The 25 students detected, represent a 17% of the students that have been included in the study, and so this is a high percentage of students committing academic dishonesty to pass a for credit course, even more since our algorithmic design lies on the conservative side and we would expect more students to be performing academic dishonesty. These overall percentages are above the order of magnitude when compared with other studies, such as the 13% of students in Alexandron et al. (2019) or 1.3% of students in Northcutt et al. (2015), despite these studies also indicated that their estimations were conservative. We believe that in our case study we are detecting a higher percentage of students due to these courses being recognized for credit as part of a degree. It should be noted that this high prevalence of academic dishonesty is these courses is also influenced by the assessment and evaluation models based on quizzes. These exams are easy to cheat both face-to-face and online, but is specially straightforward in online exams if good design practices (such as randomization, large pools of questions, etc) are not taken into account. There are other ways to implement course evaluations that are less prone to present dishonesty issues.

Regarding the interaction levels and behavioral characteristics, all previous studies (Alexandron et al. 2019; Northcutt et al. 2015; Ruipérez-Valiente et al. 2017) agree that the students performing these academic dishonest behaviors were able to pass the course

with significantly less effort than the rest of students. However, in this study, we did not find the same results with the *Lean Startup Course*, since students that are detected by the algorithm interacted similar amounts or even more with the course (but the difference was not statistically significant). One possible interpretation is that students get together to solve assignments jointly because they want to achieve good grades or they might usually work with these teams, hence they discuss exercises, agree to a solution and submit together, and they interact a lot with the contents by searching for the answers to the exams. This behavior might have different motivational roots than the more deliberate ones found in CAMEO papers. The temporal patterns clearly show more abrupt peaks of activity close to the exam in the case of the students detected performing academic dishonesty, which might be explained by working group meetings the same days to study and solve the exam together. In fact, the variance of actions per day of the detected students is 109 compared to 52 for the rest, denoting a clearly more spaced activity for the cohort of not detected students.

On the other hand, the behavior of the students in the *Texts Course* depicts slightly different patterns. In this case the students detected by the algorithm have interacted a bit less with the contents in contrast with the students detected in the *Lean Startup Course*. This means that, although the courses are designed following a similar structure and assessment methodology, we observe differences in the behavior of the students we detected. In general, the amount of interactions in the *Texts Course* is lower than in the *Lean Startup Course*. There might be multiple variables that can affect these findings; such as the different subject matter, and hence the type of contents and its difficulty, can definitely have an influence on student behavior. For example, in our case study the *Texts Course* has less video content and that most of its content is readings. This may explain why students detected as suspected of fraud in the *Texts Course* review the contents prior to a test while the students suspected in the *Lean Startup Course* have a lot of interactions during the test, as in this case the students can more easily use search functionality on the contents for answers. We believe that these findings highlight the importance of the existing balance between pedagogical design, learner behavior and academic dishonesty.

Taking into account that the MOOC ecosystem is heavily collaborating with universities towards a transition towards MOOC-based degrees and with companies for professional training programs (Reich and Ruipérez-Valiente 2019), and that online courses for credit are becoming more and more common, we anticipate that this issue will become more prevalent in the near future, and if not tackled properly, it might jeopardize the validity and stability of the online learning courses and programs for credit.

Conclusions and future work

In this study we have implemented a data-driven method for the detection of cheating in online learning that was based on previous work but has also introduced new features for a more reliable detection. We have applied this method in a for credit course taught in Selene Unicauca platform and found that 17% of the students have performed academic dishonest actions, based on current conservative thresholds. A 100% percent of the students that were detected performing academic dishonesty passed the course, while in the case of students that were not detected a 62% passed the course and we also reported significantly different behavioral characteristics between these two cohorts.

The study has some limitations, the first one and most obvious is that we have no hard proof (like video feed) that students are performing such academic dishonesty together, however the evidence shown in Table 1 is quite self-explanatory. We have only tested the algorithm in two courses, and thus we cannot argue that this could generalize to other courses, however, based on the findings of other similar studies, we have no reason to believe this to be the case. In fact, most probably this pattern in happening in most Selene Unicauca courses that have auto-graded test evaluations for credit, and more work is required to explore those courses. Furthermore, based on our findings, we believe that the severity and academic dishonesty behaviors may be importantly influenced by multiple variables and cannot be generalized from one course to another, specially when there are changes in the pedagogical design, contents and subject matter. For example, in practical courses with well-designed continuous evaluations and without auto-graded exams, the prevalence of academic dishonesty would drop significantly. One clear weakness of the detection method would be to perform adversarial attacks, this means that if students get to know how the detection algorithm works, they could just submit at different times and select some different responses to add noise that would make them hard to detect.

As part of future work, we plan to use lessons learned from our case study and others regarding the influence of course design to propose guidelines to instructors that can reduce the amount of cheating performed by students. We also plan to keep working on the robustness of the method combining more features to increase their reliability. The current method was implemented using the whole dataset of the course, and thus can only be applied retrospectively once a course is finished. Therefore, the next generation of algorithms should be able to work with less data and while the course is running, so that we could provide information to the teacher indicating which students are likely performing academic dishonest behaviors, empowering the teacher to intervene accordingly.

With MOOC-based degrees and programs becoming an important trend, and many universities experimenting with completely online and blended methodologies for credit, our findings and the ones presented in other studies clearly indicate the severity of academic dishonesty in these environments, and call for more studies, intervention experiments and alignment of teaching practices and platform functionalities with the research findings, so that the whole community can orchestrate improvements to the issue. Otherwise, this situation might endanger the future of the reliability and trustworthiness of online learning credentials.

Authors' contributions
D J-M: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing. JR-V: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing. MS: Formal analysis, Resources, Writing - Review & Editing. GR-G: Resources, Writing - Review & Editing. The author(s) read and approved the final manuscript.

Author details
¹Department of Telematics, University of Cauca, Popayán, Colombia. ²Faculty of Computer Science in the University of Murcia, Murcia, Spain.

References

Abramovich, S., Schunn, C., & Higashi, R.M. (2013). Are badges useful in education?: It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, *61*(2), 217–232. https://doi.org/10.1007/s1142301392892.

Aguaded, I., & Medina-Salguero, R. (2016). Certificación de los MOOC y su reconocimiento en créditos universitarios. *International Studies on Law and Education, 23 mai-ago*, 39–50.

Alexandron, G., Yoo, L.Y., Valiente, J.A.R., Lee, S., & Pritchard, D.E. (2019). Are MOOC learning analytics results trustworthy? with fake learners, they might not be!. *International journal of artificial intelligence in education*, *29*(14), 484–506. https://doi.org/10.1007/s40593-019-00183-1.

Arturo Amaya, A., & Alvarez, M.V. (2015). Beneficios de los MOOC en la educación superior. *Memorias del encuentro internacional de educación a distancia*, *1*(4), 1–13.

Backman, j. (2019). Students' experiences of cheating in the online exam environment. phdthesis, Laurea University of Applied Sciences.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research*, *19*(2), 185–224.

Bao, Y. (2017). Detecting multipleaccounts cheating in MOOCs. phdthesis Delft University of Technology. http://resolver.tudelft.nl/uuid:64ee55268c9e40139019c63a63413ca2. Accessed 27 Apr 2018.

Cabero, J., Llorente, C., & Vázquez, A. (2014). MOOC's typologies. *Design and educational implications*, *18*, 13–26.

Chen, X., Barnett, D., & Stephens, C. (2014). Fad or future: The advantages and challenges of massive open online courses (MOOCs).

Corrigan-Gibbs, H., Gupta, N., Northcutt, C., Cutrell, E., & Thies, W. (2015). Deterring cheating in online environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *22*(6), 28–12823. https://doi.org/10.1145/2810239. Accessed 01 Dec 2019.

Coursera (2013). Introducing Signature Track. *Coursera Blog*. https://blog.coursera.org/signaturetrack/. Accessed 1 June 2019.

Fox, A. (2013). FromMOOCs to SPOCs. *Communications of the ACM*, *56*(12), 38–40. https://doi.org/10.1145/2535918. Accessed 26 Oct 2016.

Guo, W. (2014). From SPOC to MPOC – the effective practice of peking university online teacher training. In *2014 International Conference of Educational Innovation Through Technology (EITT' 14), IEEE Computer Society* (pp. 258–264). https://doi.org/10.1109/EITT.2014.48.

Halawa, S., Greene, D., & Mitchell, J. (2014). Dropout prediction in moocs using learner activity features. *Proceedings of the second European MOOC stakeholder summit*, *37*, 58–65. Accessed 28 July 2016.

Jaramillo-Morillo, D., Sarasty, M.S., González-Ramírez, G., & Pérez-Sanagustín, M. (2017). Estrategia de seguimiento a las actividades de aprendizaje de los estudiantes en cursos en línea masivos y privados (MPOC) con reconocimiento académico en la Universidad del Cauca. Séptima Conferencia de Directores de Tecnología de Información TICAL 2017 (pp. 277–296). Costa Rica.

Jaramillo-Morillo, D., Solarte, M., & Ramírez, G. (2017). *Estrategia de seguimiento a las actividades de aprendizaje de los estudiantes en cursos en línea masivos y privados (MPOC) con reconocimiento académico en la universidad del cauca. Séptima Conferencia de Directores de Tecnología de Información, TICAL 2017*, (pp. 277–296). Costa Rica.
   Estrategia de seguimiento a las actividades de aprendizaje de los estudiantes en cursos en línea masivos y privados (MPOC) con reconocimiento académico en la Universidad del Cauca. Séptima Conferencia de Directores de Tecnología de Información, TICAL 2017, Costa Rica, 277-296.

Jobe, W. (2014). No university credit, no problem? exploring recognition of nonformal learning. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1–7). Spain: IEEE. https://doi.org/10.1109/FIE.2014.7044389.

Kaplan, A.M., & Haenlein, M. (2016). Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the cookie monster. *Business Horizons*, *59*(4), 441–450. https://doi.org/10.1016/j.bushor.2016.03.008. Accessed 01 Dec 2019.

Kloos, C.D., Muñoz-Merino, P.J., Muñoz-Organero, M., Alario-Hoyos, C., Pérez-Sanagustín, M., Ruipérez, J.A., …Sanz, J.L. (2014). Experiences of running MOOCs and SPOCs at UC3m. In *14 IEEE Global Engineering Education Conference (EDUCON)* (pp. 884–891). Istanbul: IEEE.

Lanier, M.M. (2006). Academic integrity and distance learning. *Journal of criminal justice education*, *17*(2), 244–261. https://doi.org/10.1080/10511250600866166. Accessed 01 Dec 2019.

Lei, S.A. (2010). Intrinsic and extrinsic motivation: Evaluating benefits and drawbacks from college instructors' perspectives. *Journal of Instructional psychology*, *37*(2), 153–160.

Littenberg-Tobias, J., Ruipérez-Valiente, J.A., & Reich, J. (2020). Studying learner behavior in online courses with free-certificate coupons: Results from two case studies. *The International Review of Research in Open and Distributed Learning*, *21*(1), 1–22.

Liyanagunawardena, T.R., Lundqvist, K.O., & Williams, S.A. (2015). Massive open online courses and economic sustainability. *European Journal of Open, Distance and e-Learning*, *18*(2), 95–111. https://doi.org/10.1515/eurodl20150015. Accessed 01 Dec 2019.

McGee, P. (2013). Supporting academic honesty in online courses. *Journal of Educators Online*, *10*(1), 1–31. Accessed 26 Apr 2018.

Mutawa, A.M. (2016). It is time to MOOC and SPOC in the gulf region. *Education and information technologies*, *22*(4), 1651-71. https://doi.org/10.1007/s10639-016-9502-0.

Northcutt, C.G., Ho, A.D., & Chuang, I.L. (2015). Detecting and preventing "multipleaccount" cheating in massive open online courses. http://arxiv.org/abs/1508.05699. Accessed 23 May 2018.

Palazzo, D.J., Lee, Y., & Warnakulasooriya, R. (2010). Patterns, correlates, and reduction of homework copying. https://doi.org/10.1103/PhysRevSTPER.6.010104. Accessed 01 Dec 2019.

Reich, J., & Ruipérez-Valiente, J.A. (2019). The MOOC pivot. *Science*, *363*(6423), 130–131. https://doi.org/10.1126/science.aav7958. Accessed 01 Dec 2019.

Rodríguez, M.F., Hernández Correa, J., Pérez-Sanagustín, M., Pertuze, J.A., & AlarioHoyos, C. (2017). A MOOCbased flipped class: Lessons learned from the orchestration perspective. In C. Delgado Kloos, P. Jermann, M. Pérez-Sanagustín, D.T. Seaton, S. White (Eds.) *Digital Education: Out to the World and Back to the Campus. Lecture Notes in Computer Science, Cham* (pp. 102–112): Springer. https://doi.org/10.1007/9783319590448khar12.

Ruipérez-Valiente, J.A. (2018). Analyzing the behavior of students regarding learning activities, badges, and academic dishonesty in MOOC environment. phdthesis, Universidad Carlos III de Madrid. https://earchivo.uc3m.es/handle/10016/25297. Accessed 27 Apr 2018.

Ruipérez-Valiente, J.A., Joksimović, S., Kovanović, V., Gašević, D., MuñozMerino, P.J., & Delgado Kloos, C. (2017). A datadriven method for the detection of close submitters in online learning environments. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 361–368). WWW '17 Companion: International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/3041021.3054161. Accessed 23 May 2018.

Ruipérez-Valiente, J.A., Muñoz-Merino, P.J., Alexandron, G., & Pritchard, D.E. (2017). Using machine learning to detect multiple-account; cheating and analyze the influence of student and problem features. *IEEE transactions on learning technologies*, 112–122. https://doi.org/10.1109/TLT.2017.2784420. Accessed 25 June 2019.

Sandeen, C. (2013). Integrating MOOCS into traditional higher education: The emerging "MOOC 3.0" era. *Change: The magazine of higher learning*, *45*(6), 34–39. https://doi.org/10.1080/00091383.2013.842103.

Stephen, D. (2012). *Connectivism and Connective Knowledge: Essays on Meaning and Learning Networks*. Canada: National Research Council. http://www.downes.ca/files/books/ConnectivekharKnowledge19May2012.

Tseng, H., & Walsh, E.J. (2016). Blended versus traditional course delivery: Comparing students' motivation, learning outcomes, and preferences. *Quarterly Review of Distance Education*, *17*(1), 43–52.

Wang, X.H., Wang, J.P., Wen, F.J., Wang, J., & Tao, J.Q. (2016). Exploration and practice of blended teaching model based flipped classroom and SPOC in higher university. *Journal of Education and Practice*, *7*(10), 99–104.

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online journal of distance learning administration*, *13*, 1–9.

Witthaus, G., Santos, A.I.d., Childs, M., Tannhauser, A., Conole, G., Nkuyubwatsi, B., … Punie, Y. (2016). Validation of nonformal MOOCbased learning: an analysis of assessment and recognition practices in europe (OpenCred). Accessed 18 Feb 2020.

Zhou, J., Yu, H., Chen, B., Mai, C., & Yu, L. (2016). The construction of teaching interaction platform and teaching practice based on SPOC mode. In *2016 11th International Conference on Computer Science Education (ICCSE)* (pp. 293–298). Nagoya: IEEE. https://doi.org/10.1109/ICCSE.2016.7581596.

Zirger, B.J., Rutz, E., Boyd, D., Tappel, J., & Subbian, V. (2014). Creating pathways to higher education: A crossdisciplinary MOOC with graduate credit. In *2014 IEEE Integrated STEM Education Conference* (pp. 1–5). USA: IEEE. https://doi.org/10.1109/ISECon.2014.6891056 http://ieeexplore.ieee.org/document/6891056/. Accessed 08 Aug 2017.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.