sciendo

Student engagement and academic performance in pandemic-driven online teaching: An exploratory and machine learning approach

Emilia Mioara CAMPEANU
Bucharest University of Economic Studies, Bucharest, Romania
emilia.campeanu@fin.ase.ro

Iustina Alina BOITAN
Bucharest University of Economic Studies, Bucharest, Romania

Dan Gabriel ANGHEL
Bucharest University of Economic Studies, Institute for Economic Forecasting, Romanian Academy, Bucharest, Romania

Abstract. *Fostering student engagement to acquire knowledge and achieve academic performance requires understanding how students engage in learning and its influence on academic achievement. This provides valuable insights that help improve learning experiences and outcomes. The paper relies on a mixed methods approach by expanding the traditional dimensions of student engagement and by employing a machine learning framework to identify which specific dimension of student engagement exhibits the main impact on student academic achievement. A questionnaire-based survey is conducted for the period 2020-2021 among a cohort of Romanian students. The outcomes of this preliminary exploratory analysis are further embedded into a machine learning framework by performing a LASSO regression. The findings reveal that the most relevant dimensions of student engagement, during remote education, that contribute the most to outcomes were represented by the behavioural, social, cognitive, and emotional engagement dimensions. Furthermore, the switch to online education appeared to have inverted the positive relationship between social and cognitive engagement and academic achievement. Despite the inherent challenges, the student's interest in class participation and homework completion was stimulated, and they managed to adapt without difficulty to study independently.*

Keywords: student engagement, academic achievement, remote education, linear regression, machine learning.

Please cite the article as follows: Campeanu, E. M., Boitan, I.A. and Anghel, D.G. (2023), "Student engagement and academic performance in pandemic-driven online teaching: An exploratory and machine learning approach", Management & Marketing, Vol. 18, No. SI, pp. 315 - 339, DOI: 10.2478/mmcks-2023-0017.

Introduction

The context of the pandemic has prompted behavioural changes in view of adaptation to the new rules governing our lives and economic activities. In terms of educational processes, it has become a challenge to identify adequate ways to transfer and consolidate knowledge in a remote manner, based on the joint involvement of teachers and students in this process. Engagement in the educational process differs between transmitters and receivers, especially during turbulent times when remote education is spreading and becoming, in many cases, the only possibility to continue the educational process. Against this challenging background, it is meaningful to investigate students' degree of engagement in educational processes and the dynamics of their academic achievement to highlight the possibility of improving their long-term accumulation of knowledge.

We consider that a bottom-up process of improving educational outcomes provides useful information on how to stimulate student engagement in learning and how to develop educational policy measures in terms of objectives, instruments, and allocation of funds. In this bottom-up approach, teachers play an important role in student engagement for higher academic performance. Through mastery goals, students attain higher engagement based on an induced feeling that success can be achieved by each student with persevering efforts towards continuous development. Teaching practices for student engagement and academic achievement appear to be useful in this regard (Cooper, 2014). Therefore, the findings of this research provide new insight into what drives students to engage in learning to accumulate knowledge.

The aim of the paper is consistent with the main goal of improving student engagement in the learning process, by identifying and understanding the specific dimensions of engagement that exhibit the main impact on student academic performance. In particular, our paper follows three complementary research objectives:

(1) Defining a comprehensive list of student engagement components (dimensions) and conducting a questionnaire analysis to find out how students resonated with the various types of engagement in the context of the switch to online education due to the pandemic occurrence;

(2) Investigating the statistical relationship between the various types of student engagement and the academic performance of the respondents, to uncover those components that are associated with obtaining high academic performance;

(3) Identifying the level of importance or relevance exhibited by each dimension of engagement in relation to academic performance.

Our research contributes to the existing literature in this field in several novel ways. First, we expand the traditional dimensions considered to measure student engagement in the educational process, by adding new ones related to student subjective perceptions (such as driving forces for learning, awareness of the individual engagement in learning, emotional engagement) or with changes in the teaching environment during the pandemics (student engagement with learning during traditional instruction, use of technology in learning).

Second, our analytical approach highly delineates from the mainstream literature, since it allows the identification of the key student engagement dimensions that lead to the achieved academic performance, by employing a machine learning framework which is novel for this strand of literature. Existing studies evaluate the interaction between academic

performance and various measures of student engagement in the learning process (learner engagement – Kuzminykh et al., 2021; cognitive, affective, and behavioural engagement – Delfino, 2019; Maamin et al., 2022) by generally using simple, straightforward analytical approaches, such as Pearson's correlation (Kuzminykh et al., 2021), linear regression (Maamin et al., 2022), or a mix of correlation and linear regression (Delfino, 2019).

Third, the machine learning approach is complemented by a questionnaire-based survey, data collected from Romanian students with the specialisation of Finance, from Bachelor's and Master's degree programmes.

By summing up the results of the baseline model and the various robustness checks, we find that the most relevant dimensions of student engagement during pandemic times were behavioural engagement, social engagement, cognitive engagement, and emotional engagement.

The paper is structured into six sections which present the responses to the above research questions. After the introduction, Section 2 performs a literature review on the meaning and different aspects of student engagement in educational processes that must be correlated with their academic performance. The survey instrument, detailed in Section 3, is applied in this research to capture the many facets of student engagement in the educational process and how this impacts their future academic performance. Data and statistical methodologies are indicated in the fourth section, while the research findings are presented in Section five. The conclusions are detailed in a dedicated section at the end of the paper.

Literature review

The educational process is more than 'an act of depositing' as Freire (1970) described through a metaphor. It involves achievement, skill acquisition, self-esteem, personal development, well-being, and much more for the whole society. In this research, the complexity of the educational process is considered focusing on student engagement and their relationship with academic achievements.

Student engagement in the educational process is constrained by students' likelihood of involvement in learning and their achievements, based on the variety of learning environments, from face-to-face education to e-Learning and even exclusively online learning due to the COVID-19 pandemic. Both from a theoretical and practical point of view, there is a misconception of the pillars of learning engagement that require the actions of learners and teachers for a productive educational process. The following part will clarify this issue.

The literature review reveals a variety of approaches in terms of the definition of student engagement. Finn (1993) used involvement in school as being connected with student achievement and having participation and identification as components.

Student engagement is presented as: i) energy invested by students (Astin, 1999; Fredricks et al., 2019); ii) commitment (Coates, 2006; Skinner et al., 2009; Henrie et al., 2015; Boulton et al., 2019); iii) goals achieving (Krause & Coates, 2008; Christenson et al., 2012); iv) involvement (Connell, 1990; Ben-Eliyahu et al., 2018); v) quality of efforts (Kuh, 2001; Fredricks et al., 2004; Ribeiro et al., 2019). In addition, it is described based on its effects in terms of qualitative indicators of the learning process (D'Errico et al., 2016) that lead to progression and achievement (Boulton et al., 2019).

Student engagement is perceived as having the potential to predict student academic achievements (Coates, 2006; Collie et al., 2017; Glapaththi et al., 2019; Abdelrahman, 2020;

Reeve et al., 2020; Madigan & Kim, 2021). Additionally, it can be seen as a mediator or outcome (Green et al., 2012; Barak et al., 2016).

This research is based on the hypothesis that student engagement is a driver of academic achievement. Therefore, understanding student motivation to learn is a key element in implementing deep learning strategies, especially in the case of a large number of students participating in teaching activities. This becomes a challenging issue during remote education when the teacher's ability to observe the student is significantly reduced.

Remote education has accelerated the transition from traditional campus-based higher education settings with teacher-centred lessons and student hand-raising to self-regulated learning that relies on student autonomy in terms of learning. The COVID-19 pandemic has brought a transformation of the paradigm from instructor-led to learner-centred learning strategies (Ituma, 2011; Olelewe & Agomuo, 2016; Dima et al., 2022) and imposed a lot of pressure on students and teachers to adapt to the new context of knowledge transfer. Students faced certain challenges during e-Learning processes and especially online education, such as lack of social engagement, students' emotions in a different learning context, and variety of technical tools, information, tasks, and many disturbing activities that occurred within the educational process. Kim et al. (2019) revealed the dynamics of student achievement in online learning related to students' mixed perceptions of e-learning. The lack of social interaction had negatively influenced the cognitive engagement of students (Xie et al., 2017) and their academic performance (Xie et al., 2013).

Student engagement in the educational process is the driving force behind the promotion of student knowledge in terms of understanding, using, and thinking skills, and long-term self-development. Given the complexity of the learning outcomes, it is difficult to capture it in a synthesised form using some variables. The variables most used to quantify the outcomes of the educational process of students are related to the academic achievement reflected by their results during tests. This allows the investigation of a particular student cohort for a specific medium/long–term period of time to verify the existence of an improvement in student grades.

In terms of concepts used to reflect the student outcome of the educational process, the literature review offers a variety of approaches. The most widely used constructs are learning outcome (Morgan-Thomas & Dudau, 2019), academic achievement (Boheim et al., 2020), academic performances (Boulton et al., 2019), and academic success (Fredricks et al., 2019). Several papers frequently employ constructs such as outcome (Green et al., 2012; Barak et al., 2016; Reeve et al., 2020; Snijders et al., 2020) and achievement (Collie et al., 2017; Palos et al., 2019; Abdelrahman, 2020; Madigan & Kim, 2021) related to the educational process. Additionally, Green et al. (2012, p. 1112) expressed the learning outcome as an academic self-concept referring to "students' evaluations of their academic ability".

Academic achievement is the result of high-quality active learning that involves both teachers and learners. It is more than simply focusing on student grades because it can foster self-esteem, which is different from academic self-concept (Marsh, 2007). This generates successful academic performance reflected in the final grades of the students (Boulton et al., 2019) as a consequence of retention and completion of the educational process. Understanding how student engagement in learning is linked to success requires a holistic view and a proper assessment of results.

Good academic preparation is the key to academic success achieved through the deep learning approach of students (Ribeiro et al., 2019). Taking a specific approach to learning is the way students react to various factors (such as subjects, teaching strategies, and requirements). A surface approach is detrimental to students' learning performance due to their lack of focus or compliance with learning demands while relying on memory to reproduce the learning content in the short term. However, there is general consensus on the link between academic achievement and a deeper approach to learning. This is more geared toward achieving successful outcomes through the acquisition of knowledge, understanding, and combining information as investment from students (Korhonen et al., 2017). Sometimes other factors can affect the learning approach such as motivation, learning strategies, sociocultural status (for example, parental cultural capital as De Graaf et al. (2000) or the first-generation status as Choi and Rhee (2014).

Quantitative data reflecting student academic achievement is based on grade level. Klem and Connell (2004) obtained a decrease in academic performance as students get less engaged while passing from primary to secondary and from secondary to tertiary education units. Furthermore, student grades can be a precursor of their participation in the educational process when guided by grade motivation for self-efficacy (Glynn et al., 2011).

Some authors use the grade point average (GPA) as an expression of student learning achievement (Gonzalez & Padilla, 1997; Collie et al., 2017; Kim et al., 2019; Xie et al., 2020). The GPA is determined on the basis of the academic scores of the students from the database of universities or school units. Ribeiro et al. (2019, p. 5) relies on 'the mean of students' high school grades across the three grade levels (60%) and the marks from national exams at the end of high school (40%)'. Course achievements for students can be quantified by equally weighted scores for summative exam grades, such as midterm and final evaluation, without considering effort-based activities to verify student knowledge accumulation for a specific curriculum (Reeve et al., 2020). Furthermore, student grades can be collected through student surveys (Boulton et al., 2019) and used to determine GPA as an average credit-weighted grade of the evaluated term.

Other ways of highlighting student academic achievements imply students' cumulative grade point average (CGPA) used to assess their metacognitive knowledge (Abdelrahman, 2020). Intrinsic motivation plays an important role in both educational processes geared towards achieving academic excellence and in daily life activities. This implies that students are more than just information recipients. Learners should be active participants in the educational process, with a deeper engagement. But 'optimum learning outcomes are achieved when learners possess the intrinsic motivation and true interest in the subject they learn' (Abdelrahman, 2020, p. 1).

Analysing student engagement in the educational process in relation to academic achievement involves a comprehensive investigation of the general cohorts, followed by a person-centred approach to reveal the unique profile pattern (Xie et al., 2017) and to establish dedicated active learning techniques and engaging teaching practices for successful outcomes.

Student engagement in learning and academic achievement is perceived as a vicious cycle (Anderson et al., 2019) that can weaken student success due to the erosion of student productivity. In addition, lack of motivation and involvement causes a worsening of student grades, especially for those who are disconnected from learning over time. This implies

educational risk factors that can predict academic performance (Finn, 1993). The feed-forward effects for learning become also very important.

Academic achievement is under pressure in the virtual learning environment, where students could be tempted to disengage in remote learning due to the many ways of distraction. This is the consequence of a lack of student self-direction. In this case, teachers can guide them through feedback, offering learning support, encouraging critical thinking, and collaborative learning.

Toth (2021) highlights the passive attitude of students, a lower sense of social interaction, and disengagement in learning after COVID-19 remote education based on students' and teachers' surveys from the EdWeek Research Centre. Almost 50% of the students felt unmotivated. According to the teachers' views, almost 87% of the students looked disengaged in education. As a consequence, students lack the connection to learning through a specific flow of joy, creativity, and involvement. Csikszentmihalyi (2008) explains that high skills, active student role, responsibilities, high challenges, and rigorous academic tasks are necessary conditions to create higher academic achievements.

The COVID-19 pandemic challenged students, teachers, and universities to quickly adapt to remote education with differences related to attitudes, contexts, well-being, and academic outcomes. Some studies have estimated achievement gaps and even expressed concerns about a so-called lost generation (see, for example, Dorn et al., 2020; Hanushek & Woessmann, 2020). The United Nations report refers to 'the largest disruption to education in history' because almost 1.6 billion learners from more than 190 countries were affected (United Nations, 2020). Based on research from the UK University of Southampton, it seems that almost a full year would be needed to catch up to the learning outcome gap in the case of students coming from disadvantaged families. Pensiero et al. (2020) estimate that six months is the time frame to restore lost learning of students from advantaged families. The gap is expected to increase in case the virtual learning environment prolongs in the educational process. The adverse consequences may consist of long-term economic impacts such as lower skills, lower earnings, and business collapse.

Our paper is strongly connected with the research goals of the above-mentioned literature review, by exploring the multiple dimensions of student engagement in educational processes and establishing a connection between involvement and academic achievement, considering as a case study the cohorts of students from the Faculty of Finance and Banking within a Romanian university, enroled in both bachelor's and master's degree programs.

Analysis of questionnaire

The mainstream of literature addresses student engagement mainly from the perspective of pedagogy and effectiveness of online learning as predictors of outcomes. Therefore, fostering student engagement in online education is the implicit aim of active learning. Reflection on the literature review reveals the framework for defining and studying student engagement in the learning process both cognitively and affectively. The particular case to be investigated in this paper features students from Bachelor's and Master's degree programmes in economics from a Romanian university. This provides valuable insights by investigating student engagement as a multidimensional concept and its impact on academic achievements.

The diversity of analyses involving student engagement in the educational process is usually based on a multitude of approaches that include surveys and interviews with students.

Student engagement surveys are implemented on a national scale, with the participation of groups of universities and research centres. An example of good practice in the assessment of student engagement is the case of Australia and New Zealand (Australasian Survey of Student Engagement). The results offer the necessary support to indicate student participation in learning and the pattern of academic outcomes to improve it and design better services and programmes that provide adequate resources.

Universities can investigate student engagement or success following different approaches. The preferred approach involves student surveys to assess learning engagement and outcomes. This is a common practice that can be found in the case of UK or US universities. In the Romanian case, universities implement student surveys to identify only general satisfaction considering their experiences with academic activities and facilities. For example, Bucharest University of Economic Studies (BUES) evaluates student satisfaction based on an annual survey. In terms of student engagement, as far as we know, only one research is dedicated to this topic. Stefenel and Neagos (2020) investigated how student participation was affected during COVID-19 pandemic measures based on 227 respondents (undergraduate students from several Romanian public universities). They replicate the Utrecht Work Engagement Scale for Students (UWES-9S) (Schaufeli et al., 2006).

Against this background of scarce evidence for the Romanian higher education environment, our paper fills a literature gap by considering a broader view on learning engagement to reveal the relationship with academic achievements obtained by BUES students, the Faculty of Finance and Banking (first and second-year bachelor's degree program and second year master's degree programme in Taxation).

The research aims to investigate the relationship between Romanian students' engagement in the educational process and their academic achievements looking on a specific specialisation and the student cohort who responds in a due time. The recommendations exhibited in the existing literature have been followed when defining the multidimensions of student engagement. Therefore, we rely on a mixed method approach to collect and analyse data, considering that these address the complexity of the paper's aim and reinforce the findings. Data are collected using a student questionnaire-based survey and faculty data that involve academic achievements.

The characteristics of student engagement in the educational process are investigated based on a questionnaire randomly distributed at the end of the second semester of the academic year 2020-2021, through a Google Form to students enroled in the Faculty of Finance and Banking within the Bucharest University of Economic Studies. We collected data from 73 students (38 first-year Bachelor's programme students, 22 second-year Bachelor's programme students, 13 second-year Master's programme students). The response rate is lower than the expectation, which could offer a first sign regarding the participation of students in learning. This raises a lot of concern especially for first-year Bachelor's students, due to the response rate of 43.7 %, even though we used a reward for those students who participated in this survey.

We applied the student reward because we wanted to see if this can be an incentive to comply with a learning request and to test the conclusion of some research papers considering the importance of the reward for student engagement with learning (Black &

Deci, 2000; Glynn et al., 2011; Hofferber et al., 2016). However, the response rate contradicts this hypothesis. Therefore, the engagement of students in the first-year Bachelor's programme in learning is a real problem that needs to be recognised in order to establish and apply an appropriate educational strategy in this area. This result may be a consequence of online education due to the COVID-19 pandemic because this generation of students faced only remote teaching and were not connected in person to faculty issues. Also, changing the learning environment from college to university has put additional pressure on first-year students because they need to quickly adapt to a new curriculum, to learning content and demands, to teaching practices, social networks, and educational platforms. These require an increase in personal autonomy and responsibility in terms of learning and acquiring new skills. Similar findings were obtained by Collie et al. (2017).

The questionnaire is constructed to cover multiple dimensions of student engagement in the educational process, their perception of learning engagement, and their view of the determinants of academic achievement and learning engagement. The categories of questions are described in Table 1. We use the Likert scale to allow for neutral responses.

Table 1. Content of the questionnaire

Categories	**Details**	
Driving forces for learning & knowledge retention (*DFLKR*)	• Intrinsic motivation • External incentives • Personal relevance	
Behavioural engagement (*BE*)	• Class participation	• Asking questions during your classes • Contributing to classes discussion • Preparation for classes • Taking notes • Reviewing notes prior to the following class
	Absenteeism	
	• Homework completion	• Attending classes without having completed readings or assignments • Preparing homework with a classmate • Attending a review session, tutorial or help session
Social engagement (*SOE*)	• Working with other students on projects • Assembling ideas or concepts from different courses when completing assignments or during class discussions • Work harder than you thought to meet your instructor's standards or expectations	
Emotional engagement (*EE*)	• Experiences during courses or seminars this semester • Student burnout • Perceived relationship between student emotional engagement in learning and academic achievements	
Cognitive engagement (*CE*)	• Cognitive engagement practices during classes • Types of exams to challenge you to do the best • Reading outside materials relevant to classes	
Traditional instruction and engagement in the educational process (*TI*)		
Technology enhanced classroom and engagement in the educational process (*TH*)		
Awareness of student engagement in learning (*AW*)	• Technology • Lectures • Teachers' s attitude	

Source: Authors' own research.

The questions are defined following a praxeological approach because we consider that the answers of the respondents are the result of their conscious perception, conduct, and actions without asking for the reasons behind them.

In our view, student engagement in the educational process shows the following dimensions, explained in Table 2.

Table 2. Multidimension of student engagement

Dimensions	**Motivation**
Driving forces for learning & Knowledge retention	This dimension has to do with the core personal value related to education. Therefore, we can obtain different profiles of this issue.
Behavioural engagement	Defined based on the literature review.
Social engagement	
Cognitive engagement	
Emotional engagement	This facet of engagement reflects the emotional aspects manifested within courses and seminars. The novelty of this research is represented by the fact that we include student burnout due to its effects on the educational process and on the entire behaviour of students under overwhelming learning circumstances, contents, and tasks.
Traditional instruction and student engagement with learning	This is another issue that influences learning engagement, according to the literature review, but has not been considered so far by existing studies. Therefore, we include it as a dimension of student engagement with learning that ensures the novelty of this research.
Technology and student engagement with learning	
Awareness of student engagement in learning	This dimension is connected to the personal point of view on how students see themselves acting when it comes to education. We consider this an important issue.

Source: Authors' own research.

The academic achievements of the students are reflected in the credit-weighted average grade (CAG) for all disciplines from the first semester of the academic year 2020-2021.

Analysing the students' responses to the questionnaire, we establish that the key engines of student engagement in the educational process are behaviour and cognitive dimensions, with an explanatory power of almost 50.6%, indicating that students are willing and have the ability to invest time in learning activities with more persistence. Students recognise their need to better connect with the educational process to improve academic achievement in a timely and consistent manner (71.23% of the respondents).

The responses of the students to the questionnaire revealed the driving forces for student engagement in learning, such as: i) content of learning, as a key engine for students from the Bachelor programme with an increasing tendency, while its importance is lower for students from the second-year master programme; ii) the role of teachers through their attitude toward learning, interactions with students, fairness, and support provided to students, as being the second determinant of student engagement with learning; iii) the

student motivation and interest for professional development. As the university years go by, students become more concerned (after the first-year Bachelor program) with the necessary time dedicated to learning, including individual study.

Student engagement in the educational process has an increasing pattern across university programmes with a higher score at the end of the master programme (second year). The findings contradict the literature review, revealing cases for students from United States, UK or other countries (Morgan-Thomas & Dudau, 2019). Therefore, students of the Faculty of Finance and Banking are more engaged in learning as university years pass with low volatility among the levels of engagement. We identify the specificity of learning engagement considering the fact that full participation is attained in the master programme due to the maturity of the student.

The questionnaire offers powerful information on other aspects of learning engagement and academic achievement. In terms of type of classroom, students express their preference for both face-to-face learning (on campus) and a technology-enhanced classroom because more than 72% of the responses indicate it. Online education is more useful for master students, who indicate it as their preference (almost 28%).

Regarding the determinants of academic achievement, students indicate that teacher pedagogy is the most important factor due to its ability to provide specific content of disciplines and contribute to the accumulation of new skills. Additionally, the second is considered the substance of the curriculum, indicating that students are attracted to learning by the teaching practices and learning content. Course design is the last element that students consider.

The student self-report mixed with university data provides useful information to identify students' needs and expectations in terms of enhancing teaching engagement, the determinants of learning engagement and academic achievement, and the relationship between them. We describe the variety of engagement patterns that include other variables in its dimensions to better highlight all its complexity. Furthermore, the results support future learning engagement strategies developed by teachers to reinforce students' performances in terms of learning engagement and outcome. Therefore, learning engagement needs both teachers and learners' efforts to achieve higher-order thinking and to acknowledge the nature of educational risk.

Data and statistical methodologies

Data presentation and preliminary analyses

To accomplish our second and third research objectives and to provide an in-depth understanding of both the influence and importance of the various dimensions of student engagement in relation to their academic performance, we rely on data gathered through the aforementioned questionnaire analysis and end-of-year examination grades. Consequently, the list of independent variables is represented by the indicators defined previously in Table 1.

Table 3. Descriptive statistics

Variables	Mean	Std. dev.	Min.	Max.
Credit weighted average grade (CAG)	7.87	1.06	5.6	10
Girls enrolled in a study programme (GSP)	0.78	0.42	0	1
Students enrolled in a master programme (SMP)	0.16	0.37	0	1

Variables	Mean	Std. dev.	Min.	Max.
Driving forces for learning & Knowledge retention (DFLKR)	13.52	1.02	10	14
Behavioural engagement (BE)	30.45	5.07	17	41
Social engagement (SOE)	10.35	2.21	5	15
Emotional engagement (EE)	23.85	2.52	14	29
Cognitive engagement (CE)	37.98	4.68	30	53
Traditional instruction (TI)	4.28	0.82	2	5
Technology (TH)	3.36	0.96	1	5
Awareness of student's involvement (AW)	11.53	1.16	6	12

Source: Authors' own research.

The statistical characteristics of the data set are summarised in Table 3. On average, the overall performance of a student, indicated by the credit-weighted average grade, is around 7.87. The sample average (mean) is higher for cognitive and behavioural engagement and lower for the use of technology for learning purposes and traditional teaching. We may conclude that the self-perceived level of student engagement is higher for the former dimensions and lower for the latter.

The standard deviation provides additional information on how dispersed the values of a given variable around the central tendency are. When comparing the various dimensions of academic performance, the highest spread is around the mean, and hence the highest heterogeneity of the data is obtained for the behavioural engagement (5.07), followed by cognitive engagement (4.68). Therefore, the variability of the responses to the questionnaire is greater for these two dimensions, moderate for the social and emotional engagement (around 2.21-2.52), the lowest deviation being recorded for the use of technology for learning purposes and traditional teaching (0.82 – 0.96).

As previously discussed, one of the main reasons for using machine learning-based regression models is to better control for potential multicollinearity between the variables, especially given that the data sample is relatively small. Consequently, we tested for the presence of multicollinearity by using two measuring tools: the correlation coefficient and the VIF. Indeed, our preliminary analysis shows that we are in the presence of strong multicollinearity that would severely bias traditional OLS estimates.

Table 4. Correlation matrix

	CAG	GSP	SMP	DFLKR	BE	SOE	EE	CE	TI	TH	AW
CAG	1.00										
GSP	-0.06	1.00									
SMP	0.41	-0.03	1.00								
DFLKR	0.08	0.01	0.03	1.00							
BE	0.24	0.37	0.11	0.07	1.00						
SOE	-0.03	0.01	0.23	-0.05	0.48	1.00					
EE	0.01	0.26	-0.08	-0.03	0.27	0.33	1.00				
CE	-0.02	0.04	-0.04	-0.06	0.40	0.25	0.20	1.00			
TI	-0.27	0.06	-0.20	-0.08	-0.11	0.00	0.00	0.10	1.00		
TH	0.11	0.27	0.02	0.01	0.37	0.26	0.32	0.26	-0.35	1.00	
AW	0.08	0.10	0.18	-0.02	0.24	0.16	0.10	0.21	0.00	0.14	1.00

Source: Authors' own research.

On the one hand, Table 4 shows that the correlation coefficients between some pairs of variables are nontrivial, e.g., SOE and BE are strongly positively correlated with a coefficient of 0.48, while TH and TI are strongly negatively correlated with a coefficient of -0.35. On the other hand, we also estimate and report the Variance Inflation Factor (VIF) for each independent variable in Table 5. Technically speaking, the variance inflation factor provides an estimate of how much the variance of a regression coefficient is inflated due to multicollinearity. This determines whether an independent variable has a strong collinear relationship with the others. The formula for VIF is the following:

$$VIF_i = {}^{1}/_{1 - R_i^2}$$

where: R_i^2 represents the unadjusted coefficient of determination for regressing the i*th* independent variable on the remaining ones.

The rule of thumb is that a value exceeding 10 indicates high multicollinearity (Melkumova & Shatskikh, 2017; Maamin et al., 2022), and hence lower reliability of the regression results in case these highly correlated variables would be included in the model.

Table 5. Variance inflation factor

Variable	VIF
GSP	6.4300
SMP	1.4183
DFLKR	96.7479
BE	73.3921
SOE	38.1590
EE	103.0580
CE	88.9674
TI	35.4314
TH	20.6720
AW	95.7141

Source: Authors' own research.

In our case, the VIFs are much higher than the threshold in all but two cases, the latter being related to the characteristics of the respondents in terms of gender and enrolment in a certain study programme. Overall, the particular features of the dataset support our choice to employ the Least Absolute Shrinkage Selection Operator (LASSO).

Methodological insights

The second research approach, intended to investigate the statistical relationship between the various types of student engagement and the academic performance of the respondents, in order to reveal those components of student engagement that are associated with obtaining high academic performance, makes use of the Least Absolute Shrinkage Selection Operator (LASSO) method.

It is a supervised statistical learning method that belongs to the field of machine learning. Our choice for this particular method relies on its many advantages over the classical regression methods: robustness and good performance (Chan-Lau, 2017), more

accurate predictions by identifying pertinent predictors and excluding irrelevant ones (Lee et al., 2021; OECD, 2021), alleviates the model overfitting problem, and facilitates the selection of model features (Chan-Lau, 2017; Sneiderman, 2020; OECD, 2021). The LASSO regression model represents an optimised type of linear regression (Friedman et al., 2010) that solves a minimisation problem and produces coefficient estimates close to zero (or equal zero) for those variables that do not contribute to model improvements (Melkumova & Shatskikh, 2017).

Similarly, Alshanqiti & Namoun (2020) explain the overall accuracy of LASSO regressions in two complementary ways: (i) model simplification, by performing a feature selection which neglects the useless features showing no significant impact on the dependent variable; (ii) estimation of sparse coefficients whilst adding a penalty as a regularizer, to prevent the model from being overfitted.

This method generates simpler, sparser models (with fewer parameters) and is well suited for high-dimensional models with many variables (OECD, 2021), for datasets exhibiting high levels of multicollinearity, or for purposes of variable selection or parameter elimination (Chan-Lau, 2017). Especially for cross-sectional regressions, as is the case of our study, LASSO ensures variable selection consistency, which means that the estimated null coefficients exactly coincide with the true zero coefficients when the sample size is large enough (Lee et al., 2021).

As a regularisation method, LASSO regression imposes a penalty on the less important features of a dataset and makes their respective coefficients shrink to zero or near zero, thereby eliminating the least contributing variables. The larger the penalties, the higher the coefficient values that are closer to zero (and hence the prospects for obtaining simple, parsimonious models). The implicit regularization technique employed in LASSO proves its advantages when the dataset exhibits an increased number of features (independent variables), because it automatically performs a feature selection. Once the best coefficient vector is estimated, we can identify the relevant independent variables for our model that are the highest associated with developments in the dependent variable.

The simplified mathematical equation of the LASSO regression is:

Residual Sum of Squares (RSS) + λ * (Sum of the absolute value of the magnitude of coefficients)

RSS is derived from running the standard linear regression model, which assumes a linear relationship between the independent variables and the dependent variable. The sum of the absolute value of the magnitude of coefficients, also called L1 regularisation term, is computed based on the formula:

$$L1 = \lambda * \sum_{i}^{n} |\beta i|$$

where: λ represents the amount of shrinkage and controls for the strength of the regularisation penalty.

If λ = 0, it is equivalent to the linear regression where only the residual sum of squares (RSS) is considered to build a predictive model. The bias increases with increase in λ while the variance increases with decrease in λ. The coefficients (parameters) to be estimated are symbolised with β.

The third research objective consists of identifying the level of importance or relevance exhibited by the various student engagement variables in relation to academic performance, by applying another machine learning approach, called Random Forest. This

method brings additional information for our analysis since it models the data by giving up the linearity hypothesis and estimates the importance of the characteristic (variable). Thus, it signals which variables are the most relevant and contribute the most to model improvement and the accuracy of the prediction. Findings may be further exploited by academia through increased awareness raising and empowering university boards and programme leaders to recognise the particular strengths and weaknesses of their academic programmes and thus take the necessary corrective interventions to improve student achievement. Our aim is in line with Alshanqiti and Namoun (2020), who stress the importance of such analyses.

Random forests are flexible and robust statistical learning algorithms that exhibit good behaviour on high-dimensional or noisy data, without any parameter tuning (Aria et al., 2021). They aggregate a large number of trees to solve regression or classification problems and achieve state-of-the-art performance on a wide range of problems (Bénard et al., 2021).

To evaluate the importance of the characteristic (variable), two commonly algorithms are used in the literature: i) Gini importance, also called mean decrease in impurity (MDI), which measures how much the accuracy of the model decreases when a given variable is excluded; ii) permutation importance, also known as mean decrease accuracy (MDA), which identifies the average decrease in accuracy by random permutation of the value of feature. The higher the importance of a variable's permutation, the more relevant the variable is to the accuracy of the overall prediction (Nembrini et al., 2018).

However, in both cases, a high value of these metrics suggests that the covariate is used in many important operations of the forest estimation mechanism and therefore must be maintained in the model specification (OECD, 2021).

Results obtained

Findings of the LASSO regression

Following a data-driven approach, we imposed a penalty of 0.1 and let the LASSO regression to select which of the initial indicators of student engagement actually matter for academic performance (see Table 6). Nonzero coefficients indicate the variables that contribute to the model and are associated with the outcome variable, while zero coefficients show those variables that are not significant in determining the academic performance. For comparison, we also report OLS estimates as a baseline. The estimation of the least squares corresponds to solving an unconstrained minimisation problem, while being exposed to a trade-off between bias and variance of the model. The goodness-of-fit of the model can be improved by including a large number of variables that will further minimise the bias of the model at the expense of higher variance, but the predictive power of the model and its interpretability will be negatively affected (Chan-Lau, 2017).

Linear regression identifies four proxies as statistically significant for the student engagement. The sign and statistical significance of social engagement and behavioural engagement are the same as in the LASSO regression. Additionally, the driving forces for learning & knowledge retention, and, respectively, emotional engagement are found by the OLS as exerting a positive influence on the academic performance.

Table 6. Estimation results for the OLS and LASSO models

Variable	OLS coef.	OLS t-stat	OLS p-value	Lasso coef.
GSP	-0.6902	-2.0615	0.0434	0.0000
SMP	1.2470	3.6388	0.0006	0.3891
DFLKR	0.2508	2.9631	0.0043	0.0000
BE	0.0924	2.8527	0.0059	0.0629
SOE	-0.1761	-2.5861	0.0120	-0.0691
EE	0.1011	2.0460	0.0449	0.0000
CE	-0.0075	-0.2621	0.7941	-0.0168
TI	0.0332	0.2086	0.8354	-0.1046
TH	0.0940	0.6209	0.5369	0.0000
AW	0.1066	1.0825	0.2832	0.0000
R-squared	0.9864			0.2150

Source: Authors' own research.

LASSO found a more parsimonious model that better fits the data, consisting of 5 variables out of the 10 candidate ones. Social engagement, cognitive engagement, and traditional instruction are significantly but negatively related to academic performance.

The social engagement of students (through active participation in discussions, collaboration with peers for learning purposes, and direct interaction) has been severely limited during the pandemic. Switching to online teaching has caused less interaction and team work on projects/assignments/class discussions among students and appears to have inverted the traditional positive relationship found by previous studies between SOE and academic achievements (Xie et al., 2013; Xie et al., 2020). Consequently, during pandemic times, lower social engagement has contributed to higher grades in the exam session.

The negative link between cognitive engagement and academic performance during the online learning environment suggests that the students managed to adapt without any difficulty to study by themselves and successfully passed the exams. This statistical relationship is also found by Maamin et al. (2022), who explain that student learning strategies and student dependence on their teachers have a negative impact on their achievements. Similarly, the lower the importance assigned by students to traditional face-to-face lectures in their engagement in the educational process, the better the academic results. These complementary findings point that the questionnaire respondents, although spread over various study programmes and study years, have generally adapted to the new, temporary teaching and evaluation environment.

Behavioural engagement is positively associated with good academic achievements, in line with similar previous research conducted for the traditional face-to-face teaching environment (Delfino, 2019; Kuzminykh et al., 2021; Maamin et al., 2022). Therefore, attending academic activities remotely appears to stimulate student interest in class participation and homework completion, with beneficial results in terms of improved performance.

The number of students enroled in a master's programme exhibits a positive relationship with the dependent variable, suggesting that students' awareness and interest in their academic performance are higher when enroled in upper-degree study programmes.

In contrast, variables such as the personal value related to education, the emotional engagement, the awareness of student involvement in learning, and the use of technology in the learning process exhibit a zero coefficient and therefore do not significantly determine the academic achievements.

Robustness checks

The main findings generated with the baseline LASSO approach are validated with a series of robustness checks. First, we report results for Lasso regressions performed with different (more severe) penalty terms, ranging from 0.2 to 1, with a 0.1 increase. Robust coefficients (variables) will retain their magnitude and significance when higher penalty terms are imposed. Additionally, the more resilient the coefficients to the increasing penalty, the more relevant is the variable (the more important is the feature) to explain the variability in academic performance, providing complementary information to Random Forest regression analysis (see Section 5.3). The results are reported in Table 7. We find that BE is the most resilient variable to Lasso penalty increases, which implies that it is the most important linear predictor of academic performance. BE retains its statistical significance even when the most severe penalty 1.0 is enforced. The results show that two additional linear predictors remain relevant when the Lasso penalty term is increased, namely CE and SOE. CE remains relevant up to a penalty of 0.4, while SOE remains relevant up to a penalty of 0.2. Even though SMP and TI have non-null coefficients in the baseline Lasso model, they are the least robust among the group of relevant variables, as their coefficients drop to zero when penalties above 0.1 are enforced.

Table 7. Estimation results for the Lasso model, varying penalty terms

Penalty / Variable	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
GSP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SMP	0.3891	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DFLKR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BE	0.0629	0.0546	0.0431	0.0363	0.0314	0.0275	0.0235	0.0196	0.0156	0.0117
SOE	-0.0691	-0.0239	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CE	-0.0168	-0.0170	-0.0102	-0.0025	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TI	-0.1046	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TH	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.2150	0.0840	0.0656	0.0565	0.0507	0.0468	0.0422	0.0368	0.0308	0.0204

Source: Authors' own research.

Second, we consider the Elastic Net (EN) extension of the LASSO model, which considers two regularisation factors instead of one. Basically, the EN model combines the linear penalty term in Lasso with a quadratic penalty term, which is a characteristic of the Ridge regression. In doing so, it is capable to alleviate some limitations associated with the Lasso model, the most important in our case being a very strict selection of highly correlated variables. In particular, Lasso models tend to ignore potentially relevant variables that are

highly correlated with other important predictors. Although the penalty terms in the EN model can differ, here we choose to set the same value for λ_1 and λ_2. As for Lasso, we set penalties ranging from 0.2 to 1, with a 0.1 increment. Note that, for both the Lasso and EN models, a penalizing factor of zero corresponds to standard linear regression analysis (Moorth, 2020).

Table 8. Estimation results for the Elastic Net model

Penalty / Variable	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
GSP	-0.0956	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SMP	0.5515	0.2180	0.0043	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DFLKR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BE	0.0712	0.0628	0.0585	0.0537	0.0472	0.0424	0.0390	0.0356	0.0328	0.0308
SOE	-0.0933	-0.0602	-0.0368	-0.0221	-0.0068	0.0000	0.0000	0.0000	0.0000	0.0000
EE	0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CE	-0.0184	-0.0183	-0.0183	-0.0167	-0.0133	-0.0097	-0.0060	-0.0022	0.0000	0.0000
TI	-0.1439	-0.1040	-0.0551	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TH	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.2617	0.1793	0.1086	0.0828	0.0724	0.0650	0.0607	0.0557	0.0519	0.0501

Source: Authors' own research.

When analysing the Elastic Net with a penalty of 0.1, five engagement variables appear to be significant with a non-zero coefficient, compared to only four variables in LASSO. Emotional engagement makes the difference between the two machine learning methods, but the estimated sensibility (and hence the influence on the dependent variable) is almost negligible, at 0.0010.

When looking at the estimates corresponding to the higher penalty terms, we again find that BE is the most robust (relevant) variable, followed by CE and SOE, and TI and SMP. GSP comes last, its coefficient not withstanding penalties above 0.1. In the end, the variables identified as significant in the EN model are the same as in the Lasso model, which implies that the results are highly robust to the model specification. As we will next see, the relevant variables and their order also largely correspond to the feature importance scores identified in the Random Forest regression, showing that the linear specification is capable of extracting much of the information in the dataset. However, there are some possible differences caused by the presence of non-linear interactions.

Findings generated by the Random Forest approach

The results of LASSO regression are complemented with a classification-driven approach, the Random Forest regression. We recall that the goal of our analysis is to perform variable selection in order to identify and rank all influential variables (engagement dimensions). According to Bénard et al. (2021), for this particular aim of the analysis, even if variables exhibit strong correlation, they must be kept in the model because they may represent different data features and may have different interpretations for field experts. An additional and very important advantage of random forest regressions is that they are able to capture

any nonlinear dependencies between the variables, whereas OLS, LASSO, and EN are linear in nature.

To build Random Forest regression, we train 10,000 decision trees. The feature importance is computed by alternatively using two splitting criteria, namely the impurity importance (also called Gini Importance or Mean Decrease in Impurity - MDI) and the permutation importance (Mean Decrease in Accuracy - MDA). To compute the importance of permutation of each variable, we randomly permuted 30 times the values belonging to all students.

For the MDA algorithm, the interpretation is straightforward: the larger the permutation importance of a variable, the more relevant the variable is for the overall accuracy and goodness-of-fit of the model. The MDI algorithm measures how much the model accuracy decreases when a given variable is excluded. Therefore, the higher the loss of accuracy, the more important a variable is. The findings summarised in Table 9 allow the identification of the particular student engagement dimensions that explain most of the changes in academic performance by revealing their relative importance.

Table 9. Feature importance (contribution) to the model estimated using Random Forest (RF) regression

Variable	RF MDI	RF MDA
GSP	0.0375	0.071
SMP	0.1576	0.414
DFLKR	0.0250	0.017
BE	0.1834	0.419
SOE	0.1515	0.263
EE	0.1062	0.142
CE	0.2010	0.234
TI	0.0591	0.066
TH	0.0630	0.079
AW	0.0156	0.012
R-squared	0.8674	0.8674

Source: Authors' own research.

The most relevant dimensions of student engagement that are persistently highlighted by the various Random Forest algorithms as having a significant feature importance (although their ranking in terms of the highest importance is slightly different) are behavioural engagement, social engagement, cognitive engagement, and emotional engagement.

In addition, the significant coefficients estimated in the LASSO and Elastic net models largely correspond to the important variables identified in the Random Forest, which represents a further validation of the previous results.

Conclusion

Investigating the student engagement in the educational process is a demanding task given the multitude of dimensions that must be quantified through the definition of appropriate questions. The review of the various components of student engagement highlights the

diversity of approaches. Our reasoning is that gaining a detailed and in-depth image of these dimensions is useful to better understand the phenomenon to avoid a negative attitude of students towards learning and knowledge retention that may hinder further successful academic outcomes.

Especially during uncertainty periods such as the COVID-19 pandemic, when all educational processes were online almost all over the world, the components of student engagement in the educational process were highly challenged. Online learning imposed a rapid change from traditional face-to-face education to learner-centred learning strategies. Both teachers and students faced an almost new context of education and had to quickly adapt to a new e-Learning environment. Consequently, student engagement with education changed due to off-campus educational practices and revealed the emergence of new dimensions to be considered as components of learning engagement.

In this paper, we use a proxy for academic achievement, the credit-weighted average grade (CAG), in line with previous studies that employ student grades to better reflect learning outcomes. In our opinion, it represents a suitable indicator to assess knowledge retention and the process of understanding and using the information accumulated through courses and seminars.

The vicious cycle of student engagement in the educational process and their academic achievements is investigated at the university level using questionnaires as the main method of collecting data. Our paper proposes an extension of student engagement with learning, considering other additional variables, not only behaviour, emotional, and cognitive engagement. The learning engagement has been examined according to the views of Romanian students enroled in the Faculty of Finance and Banking of the Bucharest University of Economic Studies.

The survey results highlight an increased interest in learning engagement across successive academic years, with one exception in the case of second-year master programme students. Almost 50.6% of the respondents indicated that cognitive and behaviour dimensions are key components of student engagement in the educational process. The findings are useful in improving teaching practices and the content of learning, and they open ways to create a friendly environment designed to contribute to knowledge retention and self-development.

The questionnaire-based data collected provided support for our research directions. The second research approach identifies which components of student engagement are associated with obtaining higher academic performance by employing the Least Absolute Shrinkage Selection Operator (LASSO) method. Social engagement, cognitive engagement, behavioural engagement, and traditional face-to-face instruction proved to be significantly related to academic performance.

The last research objective aims at identifying the level of importance exhibited by the various student engagement variables in relation to academic performance and applying another machine learning approach, called Random Forest. The most relevant student engagement dimensions that contribute the most to model improvement and accuracy are represented by the dimensions of behavioural, social, cognitive, and emotional engagement.

When summing up all the research findings, it seems that the use of technology (in the form of educational platforms, online teacher-student interaction) in the educational

process that took place during the pandemics does not add too much value added to the academic achievements.

This research has some limitations due to the subjective experience and perceptions of the students that can induce some biases between the results. The moderate response rate, although students were told of a reward if they answered the questionnaire, is also a drawback encountered by similar studies in this field. In our opinion, more research in this area is necessary because findings will serve as useful tools for the decision-making process related to teaching improvements (contents, interactions with students, practices), to increasing the quality of educational services provided by each university, and to reshape the educational policy objectives and financial resources allocated.

However, the overall conclusion of the paper is in line with the ones in existing literature (Mappadang et al., 2022; Al-Tameemi et al., 2023), pointing that policymakers and university management should increase their awareness and capacity of action in order to improve those university-related factors that can lead to better academic performance. The practical implications of our findings are multiple and exhibit university-specificity in terms of implementation. For example, some universities may discover that they need to develop more their digital and technical infrastructure, or the learning facilities provided both on campus and online that lecturers and students need to support the quality of learning. In this regard, Fox et al. (2021) explain the potential of digital tools to enhance teaching and improve student outcomes, with particular focus on minoritized students (in terms of race or income level) whose study retention rate and graduation rate are lower than those of residents. A meaningful report published by OECD (2023) highlights the need to implement supportive policies for the use of digital technologies in education systems, due to their capability to transform teaching and learning practices and to create more inclusive and efficient education systems.

To sum up, although there couldn't be defined a one-size-fits-all approach to be applied by stakeholders in the field of education, the future path for supporting long-lasting, quality learning is clear: public policies and academic strategies need to be recalibrated to account for the use of innovative, high-quality digital technology across education systems.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abdelrahman, R.M. (2020). Metacognitive awareness and academic motivation and their impact on academic achievement of Ajman University students. *Heliyon,* 6(9), Article number e04192.

Alshanqiti, A., & Namoun, A. (2020). *Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification*, IEEE Access, 8, 203827 – 203844.

Al-Tameemi, R.A.N., Johnson, C., Gitay, R., Abdel-Salam, A-S.G., Al Hazaa, K., BenSaid, A., Romanowski, M.H. (2023). Determinants of poor academic performance among undergraduate students — A systematic literature review. *International Journal of Educational Research Open*, 4, Article number 100232.

Anderson, K., Ritter, G., & Zamarro, G. (2019). Understanding a vicious cycle: the relationship between student discipline and student academic outcomes. *Educational Researcher,* 48(5), 251-262.

Aria, M., Cuccurullo, C., & Gnasso A. (2021). A comparison among interpretative proposals for Random Forests. *Machine Learning with Applications*, *6,* Article number 100094.

Astin, A.W. (1999). Student involvement: a developmental theory for higher education. *Journal of College Student Development,* 40, 518–529.

Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education*, 94, 49-60.

Bénard, C., da Veiga, S., & Scornet, E. (2021). MDA for random forests: inconsistency, and a practical solution via the Sobol-MDA. Retrieved from https://arxiv.org/abs/2102.13347.

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C.D. (2018). Investigating the multidimensionality of engagement: affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology*, 53, 87-105.

Black, A.E., & Deci, E.L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: a self-determination theory perspective. *Science Education*, 84, Article number 740e756.

Boheim, R., Urdan, T., Knogler, M., & Seidel, T. (2020). Student hand-raising as an indicator of behavioral engagement and its role in classroom learning. *Contemporary Educational Psychology,* 62, Article number 101894.

Boulton, C.A, Hughes, E., Kent, C., Smith, J.R., & Williams, H.T.P. (2019). Student engagement and wellbeing over time at a higher education institution. *PLoS ONE,* 14(11), Article number e0225770.

Chan-Lau, J.A. (2017). Lasso Regressions and Forecasting Models in Applied Stress Testing, IMF Working Paper, WP/17/108. Retrieved from https://www.elibrary.imf.org/view/journals/001/2017/108/article-A001-en.xml.

Choi, B.K., & Rhee, B.S. (2014). The influences of student engagement, institutional mission, and cooperative learning climate on the generic competency development of Korean undergraduate. *Higher Education,* 67(1), 1–18.

Christenson, S.L., Reschly, A.L., & Wylie, C. (2012). *The handbook of research on student engagement*. New York, NY: Springer Science.

Coates, H. (2006). *Student engagement in campus-based and online education: University connections*. New York, NY: Routledge.

Collie, R.J., Holliman, A.J., & Martin, A.J. (2017). Adaptability, engagement and academic achievement at university. *Educational Psychology,* 37(5), 632-647.

Connell, J.P. (1990). Context, self, and action: A motivational analysis of self-system processes across the life-span, In Cicchetti, D. (Ed.). *The self in transition: Infancy to childhood* (pp. 61-97). Chicago, IL: University of Chicago.

Cooper, K.S. (2014). Eliciting engagement in the high school classroom: a mixed-methods examination of teaching practices. *American Educational Research Journal*, 51(2), 363-402.

Csikszentmihalyi, M. (2008). *Flow:* The psychology of optimal experience, HarperCollins e-books. Retrieved from https://blogs.baruch.cuny.edu/authenticityandastonishment2/files/2013/04/Mihaly-Csikszentmihalyi-Flow1.pdf.

D'Errico, F., Paciello, M., & Cerniglia, L. (2016). When emotions enhance student engagement in e-learning processes. *Journal of e-Learning and Knowledge Society*, 12(4), 9-23.

De Graaf, N.D., De Graaf, P.M., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: a refinement of the cultural capital perspective. *Sociology of Education,* 73, 92–111.

Delfino, A.P. (2019). Student Engagement and Academic Performance of Students of Partido State University. *Asian Journal of University Education*, 15(1), 42-55.

Dima, A. M., Busu, M., & Vargas, V. M. (2022). The mediating role of students' ability to adapt to online activities on the relationship between perceived university culture and academic performance. *Oeconomia Copernicana*, 13(4), 1253–1281. https://doi.org/10.24136/oc.2022.036.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and student learning in the United States: The hurt could last a lifetime. Retrieved from https://www.mckinsey.com/industries/education/our-insights/covid-19-and-student-learning-in-the-united-states-the-hurt-could-last-a-lifetime.

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: The Continuum International Publishing Group.

Finn, J.D. (1993). School Engagement & Students at Risk. National Center for Education Statistics Report NCES-93-470, Washington, DC. Retrieved from https://nces.ed.gov/pubs93/93470a.pdf.

Fox, K., Vignare, K., Yuan, L., Tesene, M., Beltran, K., Schweizer, H., Brokos, M., & Seaborn, R. (2021). Strategies for Implementing Digital Learning Infrastructure to Support Equitable Outcomes: A Case-based Guidebook for Institutional Leaders. Retrieved from https://www.everylearnereverywhere.org/resources/.

Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research,* 74(1), 59–109.

Fredricks, J.A., Parr, A.K., Amemiya, J.L., Ming-Te, W., & Brauer, S. (2019). What matters for urban adolescents' engagement and disengagement in school: a mixed-methods study. *Journal of Adolescent Research,* 34(5), 491–527.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1-22.

Glapaththi, I., Dissanayake, R., Welgama, T., Somachandara, U., Weerarathna, R., & Pathirana, G. (2019). A Study on the relationship between student engagement and their academic achievements. *Asian Social Science,* 15(11), 1-16.

Glynn, S.M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, 48(10), 1159–1176.

Gonzalez, R., & Padilla, A.M. (1997). The academic resilience of Mexican American high school students. *Hispanic Journal of Behavioral Sciences,* 19, 301–317.

Green, J., Liem, G.A.D. Martin, A.J., Colmar, S., Marsh, H.W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence,* 35, 1111–1122.

Hanushek, E., & Woessmann, L. (2020). The Economic Impacts of Learning Losses. Retrieved from http://www.oecd.org/education/The-economic-impacts-of-coronavirus-covid-19-learning-losses.pdf.

Henrie, C.R., Halverson, L.R., & Graham, C.R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education,* 90, 36–53.

Hofferber, N., Basten, M., Großmann, N., & Wilde, M. (2016). The effects of autonomy-supportive and controlling teaching behaviour in biology lessons with primary and secondary experiences on students' intrinsic motivation and flow-experience. *International Journal of Science Education,* 38(13), 2114-2132.

Ituma, A. (2011). An evaluation of students' perceptions and engagement with e-learning components in a campus based university. *Active Learning in Higher Education,* 12(1), 57–68.

Kim, H.J., Hong, A.J., & Hae-Deok, S. (2019). The roles of academic engagement and digital readiness in students' achievements in university e-learning environments. *International Journal of Educational Technology in Higher Education,* 16(21), 1-18.

Klem, A.M., & Connell, J.P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health,* 74, 262–273.

Korhonen, V., Inkinen, M., Mattsson, M., & Toom, A. (2017). Student engagement and the transition from the first to second year in higher education. In Kyndt, E., Donche, V., Trigwell, K., Lindblom-Ylänne, S. (Eds.). *Higher education transitions: Theory and research* (pp. 113–134). London: Routledge.

Krause, K.L., & Coates, H. (2008). Student engagement in first year university. *Assessment & Evaluation in Higher Education*, 33(5), 493–505.

Kuh, G.D. (2001). The national survey of student engagement: Conceptual framework and overview of psychometric properties. Bloomington: Indiana University Center for Postsecondary Research and Planning. Retrieved from https://scholarworks.iu.edu/dspace/handle/2022/24268.

Kuzminykh, I., Ghita, B., & Xiao, H. (2021). The relationship between student engagement and academic performance in online education. *ICSET 2021-5th International Conference on E-Society, E-Education and E-Technology*, 97–101.

Lee, J., Shi, Z., & Gao, Z. (2021). On LASSO for predictive regression. *Journal of Econometrics*, 229(2), 322-349.

Maamin, M., Maat, S.M., & H. Iksan, Z. (2022). The influence of student engagement on mathematical achievement among secondary school students. *Mathematics*, 10(1), Article number 41.

Madigan, D.J., & Kim, L.E. (2021). Does teacher burnout affect students? A systematic review of its association with academic achievement and student-reported outcomes. *International Journal of Educational Research, 105,* Article number 101714.

Mappadang, A., Khusaini, K., Sinaga, M., & Elizabeth, E. (2022) Academic interest determines the academic performance of undergraduate accounting students: Multinomial logit evidence, Cogent Business & Management, 9(1), Article number 2101326.

Marsh, H.W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.

Melkumova, L.E., & Shatskikh, S.Ya. (2017). Comparing Ridge and LASSO estimators for data analysis. *Procedia Engineering,* 201, 746–755.

Moorth, A. (2020). How LASSO Regression Works in Machine Learning. Retrieved from https://dataaspirant.com/lasso-regression/#t-1606404715786.

Morgan-Thomas, A., & Dudau, A. (2019). Of Possums, Hogs and Horses: Capturing Duality of Student Engagement in eLearning. *Academy of Management Learning and Education,* 18(4), 564-580.

Nembrini, S., Konig I.R., & Wright M.N. (2018). The revival of the Gini importance? *Bioinformatics,* 34(21), 3711–3718.

OECD (2021). Positive, High-achieving Students? What Schools and Teachers Can Do. TALIS, OECD Publishing, Paris. Retrieved from https://doi.org/10.1787/3b9551db-en.

OECD (2023). Shaping Digital Education: Enabling Factors for Quality, Equity and Efficiency. OECD Publishing, Paris. Retrieved from https://doi.org/10.1787/bac4dc9f-en.

Olelewe, C.J., & Agomuo, E.E. (2016). Effects of B-learning and F2F learning environments on students' achievement in QBASIC programming. *Computers & Education*, 103, 76–86.

Palos, R., Maricutoiu, L.P., & Costea, I. (2019). Relations between academic performance, student engagement and student burnout: A cross-lagged analysis of a two-wave study. *Studies in Educational Evaluation,* 60, 199–204.

Pensiero, N., Kelly, A., & Bokhove, C. (2020). *Covid-19 lockdown primary-secondary education impact*. Retrieved from https://www.southampton.ac.uk/news/2020/07/lockdown-education-impact.page.

Reeve, J., Cheon, S.H., & Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemporary Educational Psychology,* 62, Article number 101899.

Ribeiro, L., Rosário, P., Núñez, J.C., Gaeta, M., & Fuentes, S. (2019). First-Year Students Background and Academic Achievement: The Mediating Role of Student Engagement. *Frontiers in Psychology*, 10, Article number 2669.

Schaufeli, W.B., Bakker, A.B., & Salanova, M. (2006). The Measurement of Work Engagement with a Short Questionnaire: A Cross-National Study. *Educational and Psychological Measurement*, 66(4), 701-716.

Skinner, E.A., Kindermann, T.A., Connell, J.P., & Wellborn, J.G. (2009). Engagement as an organizational construct in the dynamics of motivational development. In Wentzel, K., Wigfield, A. (Eds.). *Handbook of motivation in school* (pp. 223–245). Malwah, NJ: Erlbaum.

Sneiderman, R. (2020). From Linear Regression to Ridge Regression, the Lasso, and the Elastic Net. Retrieved from https://towardsdatascience.com/from-linear-regression-to-ridge-regression-the-lasso-and-the-elastic-net-4eaecaf5f7e6.

Snijders, I., Wijnia, L., Rikers, R.M.J.P., & Loyens, S.M.M. (2020). Building bridges in higher education: Student-faculty relationship quality, student engagement, and student loyalty. *International Journal of Educational Research,* 100, 101538.

Stefenel, D., & Neagos, I. (2020). Measuring academic engagement among university students in Romania during Covid-19 pandemic. *Thesis,* 9(2), 3-29.

Toth, M.D. (2021). Why student engagement is important in a post-COVID world – and 5 strategies to improve it. Retrieved from https://www.learningsciences.com/blog/why-is-student-engagement-important/.

United Nations (2020). Policy brief: Education during COVID-19 and beyond. Retrieved from https://www.un.org/sites/un2.un.org/files/sg_policy_brief_covid-19_and_education_august_2020.pdf.

Xie, K., Lu, L., Cheng, S.L., & Izmirli, S. (2017). The interactions between facilitator identity, conflictual presence, and social presence in online collaborative learning. *Distance Education,* 38(2), 230–244.

Xie, K., Miller, N.C., & Allison, J.R. (2013). Toward a social conflict evolution model: Examining the adverse power of conflictual social interaction in online learning. *Computers & Education*, 63, 404–415.

Xie, K., Vongkulluksn, V.W., Lu, L., & Cheng, S.-L. (2020). A person-centered approach to examining high-school students' motivation, engagement and academic performance. *Contemporary Educational Psychology,* 62, 101877.