



Implementation of Learning Analytics Indicators for Increasing Learners' Final Performance

Denizer Yildirim¹ · Yasemin Gülbahar²

Accepted: 17 November 2021 / Published online: 5 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

This study aims to determine indicators that affect students' final performance in an online learning environment using predictive learning analytics in an ICT course and Turkey context. The study takes place within a large state university in an online computer literacy course (14 weeks in one semester) delivered to freshmen students (n=1209). The researcher gathered data from Moodle engagement analytics (time spent in course, number of clicks, exam, content, discussion), assessment grades (pre-test for prior knowledge, final grade), and various scales (technical skills and "motivation and attitude" dimensions of the readiness, and self-regulated learning skills). Data analysis used multi regression and classification. Multiple regression showed that prior knowledge and technical skills predict the final performance in the context of the course (ICT 101). According to the best probability, the Decision Tree algorithm classified 67.8% of the high final performance based on learners' characteristics and Moodle engagement analytics. The high level of total system interactions of learners with low-level prior knowledge increases their probability of high performance (from 40.4 to 60.2%). This study discussed the course structure and learning design, appropriate actions to improve performance, and suggestions for future research based on the findings.

Keywords Online learning · Learning analytics · Final performance · Self-regulated learning · Technical skills · Motivation and attitude

1 Introduction

Along with the digitalization of learning environments, a significant amount of data has been collected long-term, including educational programs, courses, and learner details. Learning Analytics (LA) emerged to explore and provide insight from the data in an

✉ Denizer Yildirim
dyildirim@ankara.edu.tr
Yasemin Gülbahar
gulbahar@ankara.edu.tr

¹ Faculty of Open and Distance Education, Ankara University, 06100 Beşevler, Yenimahalle/Ankara, Turkey

² Faculty of Educational Science, Ankara University, 06590 Cebeci, Çankaya/Ankara, Turkey

educational context due to the various levels of granularity of the collected data. LA aims to monitor learners' progress, predict their performance, dropout/retention rates, provide feedback to the learners, provide advice, and facilitate the self-regulation of online learners (Chatti et al., 2012; Papamitsiou & Economides, 2014). However, analytics alone are not enough to improve learning processes (Wong et al., 2019). For example, the analytics may show that a student has limited interactions across the entire system and is, therefore, likely to achieve a poor level of performance. However, human intervention is also required to improve the learner's system interactions or level of academic performance. LA implementation is a prerequisite to designing these interventions (Chatti et al., 2012; Clow, 2013; Omedes, 2018).

When starting LA, it is also necessary to draw the boundaries ("what purpose," "for whom," "what data," and "how to analyze") and to reveal objectives due to the broad scope of LA (Chatti et al., 2012). In this context, it is remarkable that by focusing on academic success, most researchers predict performance with LMS data (Conijn et al., 2017; Iglesias-Pradas et al., 2015; Mwalumbwe & Mtebe, 2017; Saqr et al., 2017; Strang, 2016; Zacharis, 2015), compare various techniques to increase the predictive power (Cui et al., 2020; Hung et al., 2019; Miranda & Vegliante, 2019; You, 2016), and predict using individual characteristics and LMS data (Ramirez-Arellano et al., 2019; Strang, 2017). However, there is still no consensus on designing interventions to increase learning outcomes.

LA studies showed that researchers should not use "learning analytics as a one-size-fits-all approach" (Gašević et al., 2016; Ifenthaler & Yau, 2020). Results of analytics (e.g., predictive analytic) may show extraordinary results according to different contexts. Ifenthaler and Yau (2020) found that the positive effect of LA on learning outcomes is in small-scale studies. Therefore, the small-scale data collected, the analysis made, and the results obtained are limited only to their context. For example, some LA research indicates the difficulty of generating generalizable models due to the differences in course structures in the online learning environment (Hung et al., 2019; Olive et al., 2019).

LA studies showed that "researchers should not use learning analytics as a one-size-fits-all approach" (Ifenthaler & Yau, 2020). Results of analytics (e.g., predictive analytic) may show extraordinary results according to different contexts. Ifenthaler and Yau (2020) found that the positive effect of LA on learning outcomes is in small-scale studies. Therefore, the small-scale data collected, the analysis made, and the results obtained are limited only to their context. For example, some LA research indicates the difficulty of generating generalizable models due to the differences in course structures in the online learning environment (Hung et al., 2019; Olive et al., 2019).

Another example is the way learning outcomes are handled. Learning outcomes are a new paradigm referencing learner achievement for various levels of education (Macayan, 2017). This paradigm refers to the knowledge, skills, and values students acquire at graduation or the end of a course (Premalatha, 2019). In other words, learning outcomes represent a more comprehensive experience than learners' assessment grades. However, although this paradigm reflects the ideal situation, adaptation towards different countries has not fully achieved this goal. For example, in distance education programs in Turkey, the effect of process assessment, such as performance, project, homework, thesis, and portfolio, and unsupervised exam and assessment activities on overall success cannot be more than 40%. Until Sep 2020, this ratio was also 20% (Council of Higher Education, 2020). Therefore, this proportional power (60%-80%) given to the supervised exam (e.g., final exam) in the assessment process significantly affects the teacher's behavior while structuring the lesson

and the student's behavior while continuing the learning activity. In this study, the learning outcome is considered only final performance due to the Turkish context.

Namoun and Alshantiti (2021) systematically examined the studies that predict learning outcomes. Researchers found that the dominant factors in most of the studies were learning and activity behavior (e.g., time and number of online sessions), assessment data (e.g., assignment, exam grade), emotions (e.g., motivation), and previous academic performance (e.g., prior knowledge). Another study (Yau & Ifenthaler, 2020) determined LA indicators of study success collected in three groups as student profile data (e.g., prior knowledge, motivation), learning profile data (e.g., LMS engagement data, assessment grade), and curriculum data (e.g., course characteristics, course structure). Therefore, both studies showed that prior knowledge, motivation, LMS engagement data, assessment data have a predictive effect on learning outcomes.

Predictive analytics enable the design of meaningful, evidence-based interventions to collect and store data for the various factors mentioned above and combine them, thereby improving learning outcomes/study success (Ifenthaler & Yau, 2020). Therefore, expanding the data profile collected and analyzed (e.g., self-regulation, e-readiness) can yield more evidence-based results to design these interventions. This study tested multiple regression and classification models by considering self-regulation and online readiness variables and the data profiles (e.g., online activities, prior knowledge, motivation) to predict the final performance.

In an ICT course and Turkey context, this study aims to determine indicators that affect the final exam performance of students within an online learning environment by using predictive learning analytics. Research questions are as follows.

1. Does prior knowledge, e-readiness, self-regulation skills predict final grade? If so, to what extent?
2. Do LMS engagement analytics have a positive relationship with final grades and with other predictors?
3. According to the classification model generated using learners' Moodle engagement and other predictors, what variables come to the fore in learners' final performance?

2 Related Literature

2.1 Prediction Analytics Using Data in Learning Management System

The log data produced in Learning Management System (LMS) constitute the primary reference source of LA research based on data. Naturally, there have been many studies investigating the prediction of academic success with LMS data. In some studies (Mwalumbwe & Mtebe, 2017; Saqr et al., 2017; Zacharis, 2015), the classification power of LMS data on academic achievement is worth considering, while in some studies (Conijn et al., 2017; Iglesias-Pradas et al., 2015; Strang, 2016) LMS data partially contributed. For example, Saqr et al. (2017) found that engagement parameters showed significant positive correlations with student performance, especially those reflecting motivation and self-regulation. The researchers were able to classify performance with 63.5% accuracy and identify 53.9% of at-risk students. Another study (Mwalumbwe & Mtebe, 2017) found that peer interaction (beta value = 19.6%) and forum posts (beta value = 77.1%) significantly affected

students' performance in Applied Biology. But, in Service and Installation IIT, forum posts (beta value = 48.5%) and exercises (beta value = 51.5%) impacted students' performance.

In studies where LMS data partially contributed, Conijn et al. (2017) revealed that the accuracy of the prediction models differed mainly between the courses, with between 8 and 37% explaining variance in the final grade. For early intervention or in-between assessment grades, the LMS data proved to be of little value. Another study (Strang, 2016) compared student test grades with engagement LA indicators to measure hypothesized relationships' strength and predictive nature. The researcher indicated very little correlation between student online practices and their academic outcomes. Iglesias-The study of Iglesias-Pradas et al. (2015) founded no relation between online activity indicators and either teamwork or commitment acquisition. Therefore, LMS data alone may sometimes not provide sufficient information in terms of the final performance.

2.2 Using Different Techniques to Increase Predictive Performance

LA can use a data analysis combination determined by the purpose. This combination may include simple statistical methods to very complex techniques such as deep learning (Avella et al., 2016; Leitner et al., 2017). While advanced data analysis methods are used widely in computer science, analytical techniques such as statistics, data visualization, clustering, regression, and decision trees are used primarily to support decision-making with regards to learning (Du et al., 2019). Therefore, to increase the classification accuracy/precision or predictive power of LMS data on academic achievement, many studies have been conducted in which different data conversion or classification techniques (Cui et al., 2020; Helal et al., 2018; Hung et al., 2019; Miranda & Vegliante, 2019).

These studies in the literature (Cui et al., 2020; Helal et al., 2018; Hung et al., 2019; Miranda & Vegliante, 2019) showed that some classification techniques produced better results in some situations (e.g., data conversion). For example, Cui et al. (2020) compared various machine learning classifiers (e.g., logistic regression, Naïve Bayes, neural network, ensemble model, gradient boosting machine) for the three undergraduate courses. The researchers found the mean grade of quizzes/assignments as one of the essential features for all three courses rather than time-related and frequency-related LMS data. Another study (Hung et al., 2019) investigated how the absolute frequency variables or the relative-transformed variables affect the forecast results in estimation analysis. Classification algorithms (Neural Network, Random Forest) produced better estimation results when using relative-transformed variables (e.g., scale the frequency variables from 0 to 10). However, each classification technique has different estimation results by data collection types (self-report or event-based) even under the same conditions (Moreno-Marcos et al., 2020). In this context, it is challenging to create a precise roadmap about which analysis technique to use in LA under different conditions. Therefore, in this study, the researchers used a practical way of selecting the one that produces more precise results among multiple classification techniques.

2.3 Prediction Using Individual Characteristics as Well as LMS Data

LA research focused on increasing the number of variables by combining LMS data with data collected from different sources (e.g., self-report data) to strongly predict academic success (Ramirez-Arellano et al., 2019; Strang, 2017). Ramirez-Arellano et al. (2019) investigated the relations between students' motivation, cognitive–metacognitive strategies,

behavior, and learning performance in the context of blended courses in higher education. This experimental study was carried out with 137 Mexican students. Only six (e.g., missing learning activities, self-efficacy, metacognitive self-regulation) of the 19 variables discussed explain approximately 67% of each student's overall grade variance. The model correctly classifies 96% of the risk of failing using six variables. Strang (2017) used a mixed-method approach that examined several important student attributes and online activities, which seemed to predict higher grades best. Strang (2017) collected qualitative data and analyzed it with text analytics to uncover patterns and tested Moodle engagement analytics indicators as predictors in the model. The findings revealed a significant General Linear Model with four online interaction predictors that captured 77.5% of grade variance within an undergraduate business course. In this context, the use of individual characteristics with LMS data can yield better predictive results.

Due to the difficulties of physically providing practical support to learners within online learning environments, learners need to self-regulate. Learners' self-regulation concerns individual factors that can change in many ways (Wong et al., 2019). In the online or blended environment, many variables (e.g., adopting learning strategies, prior knowledge, self-regulated strategies, motivation) that may affect academic success addressed (Azevedo et al., 2010; De Barba et al., 2016; Pardo et al., 2016; Sun et al., 2018). Pardo et al. (2016) provided robust evidence of the advantages of combining self-reported and observed data sources to gain more precise insight into learning experiences leading to more effective overall improvement. The current study address self-regulation skills and e-readiness are as examples of these individual characteristics. These variables were measured through self-reported data and added as indicators in the predictive and classification models.

2.4 Self-Regulation and e-Readiness

Self-regulated learning in educational research offers a means to understand why students from various perspectives are more successful than others. For example, academic performance can decrease when learners do not apply self-regulated learning strategies (Pardo et al., 2016). Self-regulated learning (SRL) strategies are considered from a broad perspective as forethought, performance, and reflection (Lu & Yu, 2019), and more complexly as goal setting, strategic planning, self-evaluation, task strategies, elaboration, and help-seeking (Kizilcec et al., 2017; Papamitsiou & Economides, 2019). Kizilcec et al. (2017) examined the relations between SRL strategies, learner behavior, and goal achievement. The results of their research showed that learners with high goal-setting and strategic planning scores are more likely to achieve their personal course goals. However, other SRL strategies, except for help-seeking, have been associated with the frequency of re-interacting with course material. Papamitsiou and Economides (2019) showed that goal-setting and time management have strong positive effects on autonomous control. Effort regulation moderately positively affects learner autonomy, while help-seeking can have a strong negative impact (Papamitsiou & Economides, 2019).

Online learning readiness appears to be an essential variable for instructional design processes, with a high potential to influence learners' academic performance. For example, Joosten and Cusatis (2020) stated that student success might increase by evaluating learners' preparedness and readiness. For example, a student who has the skills (e.g., technical skills) to learn online and has motivation and expectations for online learning can succeed. The literature has stated that readiness is positively related to academic success and satisfaction (Horzum et al., 2015; Yilmaz, 2017). For instance, in a study with

236 undergraduate students in the context of flipped learning, Yılmaz (2017) found that e-learning readiness positively affected student satisfaction ($\beta=0.61$; $R^2=0.43$). Joosten and Cusatis (2020) found that online learning efficacy (the belief that online learning can be as effective as traditional classroom learning) was significant in predicting the academic performance or course grades of learners ($\beta=0.38$, $p < 0.0001$).

3 Methods

The current study followed learning analytics to answer the research questions (see Fig. 1).

3.1 Study Group

The study was conducted in an online computer literacy course (14 weeks in one semester) delivered to freshmen students of all faculty and schools of a large state university. The study group consists of 3765 registered users from 17 different faculty and 75 different departments. A total of 1209 students participated in the research. Of the participant students, 382 (31.6%) are female and 827 (68.4%) are male; they are aged between 19 and 75 (mean = 20.9; median = 19), and a total of 1106 (91.5%) are unemployed and 103 (3.5%) are employed.

3.2 Teaching–Learning Process

Information and Communication Technologies 101 (ICT 101) is a beginner-level course as a fully asynchronous online activity in Moodle. The course content was organized linearly, with students advised to study the relevant topic based on their pre-assessment scores and forced to be 70% successful or more before starting the next topic. In other words, students were expected to perform a level of at least 70% in their post-assessment tests to be labeled as "successful" and thereby ready to start the next topic.

Based on these facts, learning goals were defined for each topic which varied in number and difficulty. Interactions to be investigated in an LMS were limited by reading course handouts, watching instructional videos, solving interactive questions, and completing

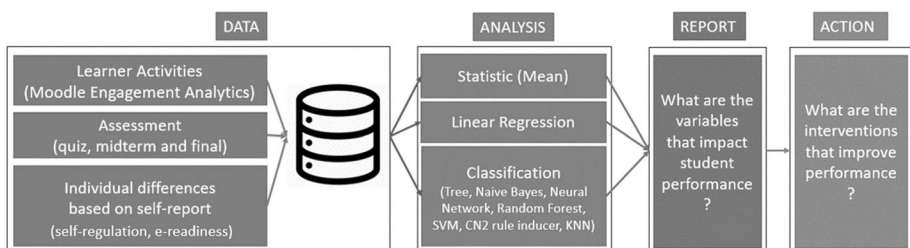


Fig. 1 Application of learning analytics in this study

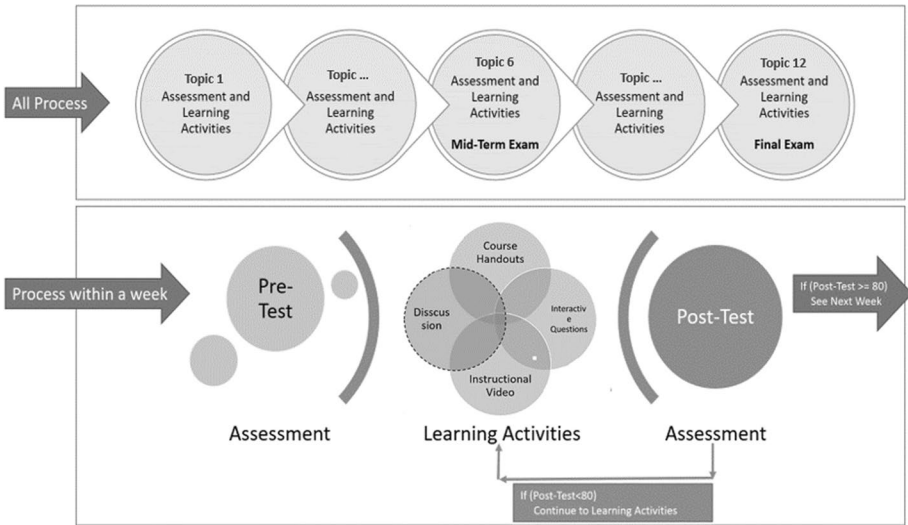


Fig. 2 Teaching–learning process

achievement tests. The students discussed (at 8th and 12th weeks) the topics before the midterm and the final exam (see Fig. 2).

3.3 Data Collection

Data were collected from five tools: Self-Regulation Survey, e-Readiness Scale, Moodle Analytics, Assessment Tests (Table 1). The scales developed in the native language of the participants were used.

3.3.1 Self-Regulation Survey

For revealing students’ self-regulated learning skills (SRS), the questionnaire named "Self-regulated Learning Skills for Self-managed Courses," as developed by Kocdar et al. (2018) was used to collect data. The questionnaire contains 30 items as 5-point, Likert-type

Table 1 Data collection process

Data collection tools		When
Assessment tests	Pre-test	At the beginning of each topic
	Interactive questions	At the end of the instructional video
	Posttest	At the end of each topic
Self-regulation survey		Week 1
e-Readiness scale		Week 2
Assessment test	Midterm	Week 8
Moodle analytics		Week 10
Assessment test	Final	Week 14

questions. Kocdar et al. (2018) found the total variance of the scale to be 58.204% and the Cronbach's alpha coefficient for the reliability of the scale to be 0.918. As a result of the tests made with our study's data, the scale's total variance was 67.73% (Kaiser–Meyer–Olkin (KMO)=0.968, $p=0.000$), and the Cronbach's alpha coefficient of the scale was 0.96.

3.3.2 e-Readiness Scale

"e-Readiness Scale" was originally developed by Gülbahar (2012). The KMO value of the scale was found to be 0.941, and the value of the Bartlett test was found to be significant ($p < 0.001$) by Gülbahar (2012). The researcher found the reliability of the scale to be 0.94 (Cronbach's alpha). For the current study, questions from two factors of the e-Readiness Scale, "Technical Skills (TS: $\alpha=0.79$)" and "Motivation and Attitude (MaA: $\alpha=0.79$)," which both consist of six questions, were taken into consideration. Thus, the scale version employed in the current study is composed of 5-point, Likert-type questions. In our study, the total variance of the Technical Skills scale ($\alpha=0.94$) was 71.66% (KMO=0.993, $p=0.000$) and the total variance of the "Motivation and Attitude ($\alpha=0.88$)" was 74.12% (KMO=0.812, $p=0.000$).

3.3.3 Moodle Analytics

Moodle Engagement Analytics refers to the system interactions of learners in an online course. Eleven variables were determined regarding the system interaction (actions for all components: creating, viewing, submitting). These variables were chosen according to the activities (Table 2) included in Teaching–Learning Process (Sect. 3.2).

3.3.4 Assessment Tests

In this study, assessment tools were used to increase learning performance during the application process. These tools were organized as; (1) multiple-choice pre-test measuring prior knowledge and (2) multiple-choice final exam held face-to-face with paper-and-pencil. Since the final exam has high proportional power (70%) for learners' assessment, the final grade was determined as a dependent variable. Other variables were used as independent variables to predict the learners' final performance.

3.4 Data Analysis

Data analysis includes regression, correlation, and classification analysis. Regression analysis and correlation were used to reveal linear relationships between indicators and final grade using *SPSS 20*. Classification analysis was used to disclose nonlinear relationships using *Orange 3.24.1*.

Before classification analysis, *dimension reduction* was performed by applying *Principal Component Analysis (PCA)* to compute principal components from system interaction variables. Accordingly, nine variables were collected in three dimensions; "exam (E)", "content (C)", and "discussion (D)" (KMO: 0.743). The total explained variance value was found to be 72.42%. Other variables (total time spent, total action) were

Table 2 The variables according to the activities

Assessment	Learning activities	Other
Total views of exam components	Total views of the scorm content component	Total time spent on the course page
Total number of completing exam components	Total views of instructional videos	Total number of actions (e.g., view, update, submit) performed on the course page
Total number of re-exam of attempts completed	Total views of external links	
	Total views of discussion	
	Total number of posts added to the discussion	
	Total views of user profiles	

Table 3 Principal component analysis (rotated matrix)

Variables	Exam (E)	Content (C)	Discussion (D)
Total views of exam components	0.967	0.109	0.076
Total number of completing exam components	0.966	0.102	0.077
Total number of re-exam of attempts completed	0.867	0.412	0.226
Total views of the scorm content component	-0.082	0.788	0.027
Total views of instructional videos	0.268	0.753	0.134
Total views of external links	-0.045	0.709	0.045
Total views of discussion	0.015	-0.089	0.824
Total number of posts added to discussion	0.141	0.226	0.751
Total views of user profiles	0.118	0.256	0.684
Total time spent on the course page	0.612	0.587	0.234
Total number of actions (e.g. view, update, submit, etc.) performed on the course page	0.667	0.412	0.226

Bold provides information about which variable is related to which dimension

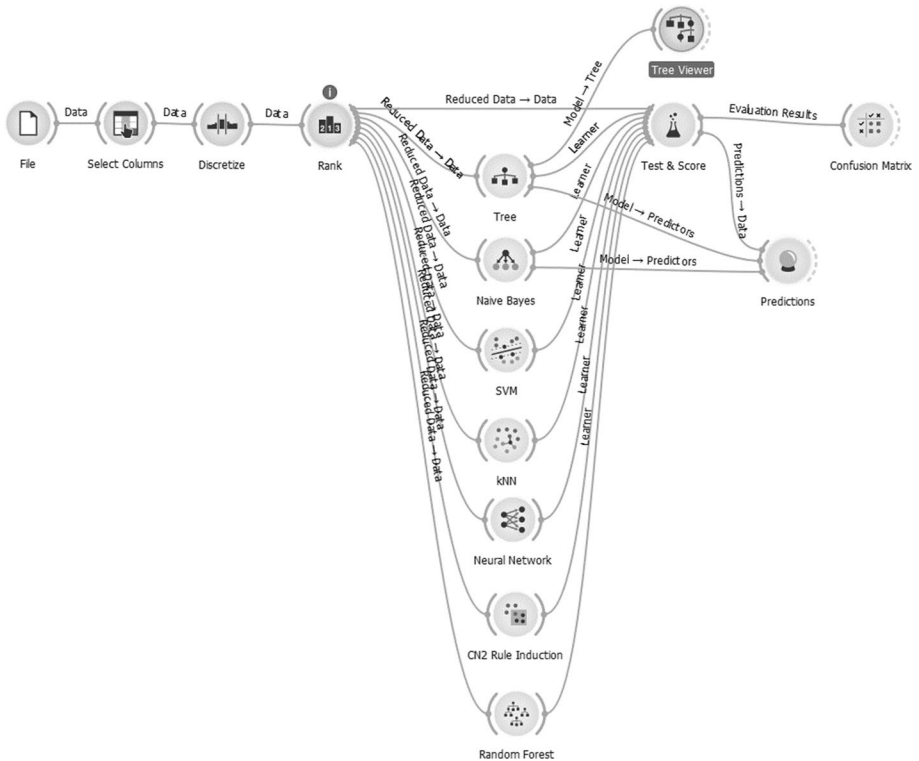


Fig. 3 Classification analysis in Orange 3.24.1

evaluated separately from exam, content, and discussion because they were not loaded according to the dimensions (Table 3).

Classification stages were shown in Fig. 3. In stage1 (Select Column), the target variable was selected as the final performance. The final performance was categorized as low performance: " < 70 " or high performance: " ≥ 70 ", as we expect students to perform at a level of 70 points for each subject (Fig. 2). Furthermore, the features were selected interaction data (Exam, Content Discussion, Total Action, and Total Time Spent), Assessment Grade (Average PreTest), individual characteristics (Technical Skills, Motivation and Attitude, and Self-Regulation). In stage2 (data discretization), continuous variables were categorized using equal frequency. Equal frequency divides the attribute into a certain number of intervals so that each interval contains approximately the same number of samples. In stage3 (Feature Selection-Rank), all features were selected because we wanted to see the predictive power of all the indicators in the classification model. In stage4 (e.g., Tree, Naïve Bayes, SVM, KNN, Neural Network, CN2), the researchers used a practical way of selecting the one that produces more precise results among multiple classification techniques. The classification model was evaluated by comparing the performance of various algorithms based on probability-based and rule-based. In stage5 (test & score), cross-validation (five-folds) sampling was used for the internal validity of the model. Cross-validation splits the data into a certain number of multiples (usually 5 or 10). The algorithm was tested by taking samples one layer at a time. Stage 6 (Tree Viewer, Confusion Matrix) was presented in the results section.

4 Results

4.1 Effect of Prior Knowledge (AvePreTest), e-Readiness (TS and MaA), and Self-Regulation Skills (SRS) on Final Grade (FG)

The effect of predictors on final grade was investigated through the Forward Regression Method. The P-P plot, Durbin Watson, and Residual Statistics (Mahalanobis Distance, Cook's Distance, and Centered Leverage Value) were examined for the assumptions. In this context, it was observed that the standardized residual distributed normally (see Fig. 4).

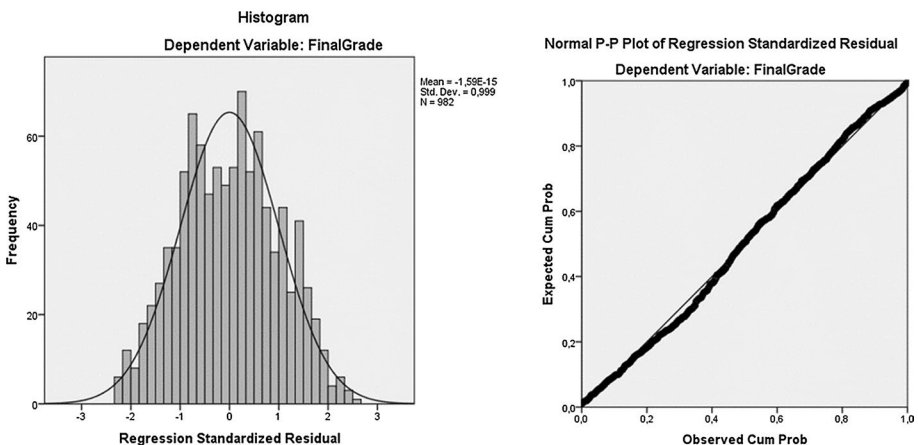


Fig. 4 Distribution of residual

Table 4 Pearson coefficients between variables

	FG	AvePreTests	TS	MOT	SumSR
FG	–	–	–	–	–
AvePreTests	0.427**	–	–	–	–
TS	0.200**	0.291**	–	–	–
MaA	0.146**	0.190**	0.571**	–	–
SRS	0.060*	0.098**	0.253**	0.373**	–

**Significant at the 0.01 level

*Significant at the 0.05 level

Table 5 Distance values

	Minimum	Maximum	Mean	Std. deviation	N
Mahal. distance	0.005	9.709	1.998	1.831	982
Cook's distance	0.000	0.004	0.001	0.001	982
Centered leverage value	0.000	0.010	0.002	0.002	982

For the multicollinearity assumption, it is considered sufficient for the coefficient values (r) to be below 0.800 and the VIF value to be 2.5 and below (Allison, 1999; Berry, Feldman, & Feldman, 1985). This study found correlation coefficients between independent variables (AvePreTest, TS, MaA, SRS) below 0.600, and VIF values were 1.092 (Table 4).

Mahalanobis, Cook, and Centered Leverage values were examined for outlier control, and the outliers were deleted until all values reached the desired level. After this process, it was ensured that (1) Mahalanobis's maximum distance is equal or lower than 16.27 in the chi-square table, (2) Cook's maximum distance is lower than " $4/(\text{sample size} - \text{number of predictors} - 1)$ " calculation and (3) Centered Leverage Value is less than " $(2 * \text{number of predictors} + 2) / \text{sample size}$ " calculation (Hair et al., 2010) (Table 5).

The regression forward method showed that MaA ($\beta=0.033$; $t=0.946$; $p=0.343$) and SRS ($\beta=-0.001$; $t=-0.023$; $p=0.982$) were excluded from the model since these variables had no significant effect on the final score. Regression model using only AvePreTest as independent variables ($F=218,001$; $p=0.000$) and the regression model using AvePreTest and TS together ($F=113,465$; $p=0.000$) were found to be significant. According to the two models, AvePreTest alone explained 18.1% ($\text{Adj. } R^2=0.181$) of the final grade and 18.7% with TS ($\text{Adj. } R^2=0.187$) (Table 6).

Prior knowledge alone positively affected the final grade ($\beta=0.427$; $p=0.000$). In the model created together with prior knowledge and technical skills, prior knowledge has a positive effect of 0.403, and technical skills positively impact 0.082 on FG ($\beta_{\text{AvePreTest}}=0.403$; $\beta_{\text{TS}}=0.82$; $p<0.01$). Accordingly, prior knowledge's significant and positive effect on the FG can be considered an ordinary situation. However, although technical skills have a significant impact, the minimal effect size concludes that there may not be a linear relationship between TS and FG (Table 7).

Table 6 Model summary and excluded variables

Models	Model 1			Model 2	
Predictors	(Constant) AvePreTests			(Constant) AvePreTests TS	
R	0.427			0.434	
R ²	0.180			0.19	
Adjusted R ²	0.181			0.187	
Std. error	8.391			8.363	
R square change	0.182			0.006	
F change	218			7.486	
df1	1			1	
df2	980			979	
Sig. F change	0.000**			0.006*	
Model	1			2	
Excluded variables	TS	MaA	SRS	MaA	SRS
<i>Excluded variables</i>					
Beta In	0.080	0.070	0.018	0.030	-0.001
t	2.740	2.290	0.620	0.950	-0.023
Sig	0.01**	0.02*	0.535	0.340	0.982
Durbin-Watson	1.783				

*Significance at the 0.05 level

**Significance at the 0.01 level

Table 7 Coefficients

Model	B	Std. Error	Beta	t	Sig	Lower	Upper	Tolerance	VIF
<i>1</i>									
(Constant)	67.290	1.004		67.038	0.000	65.320	69.260		
AvePreTests	0.306	0.021	0.427	14.765	0.000	0.265	0.347	1.000	1.000
<i>2</i>									
(Constant)	64.152	1.522		42.149	0.000	61.165	67.138		
AvePreTests	0.289	0.022	0.403	13.377	0.000	0.246	0.331	0.915	1.092
TS	0.121	0.044	0.082	2.736	0.006	0.034	0.208	0.915	1.092

4.2 Relationship of Moodle Engagement Analytics, Other Predictors and Final Grade

Correlation analysis found that there was a positive and low-level significant relationship between AvePreTest ($r_{TAction} = 0.270$; $r_{TSpent} = 0.283$; $r_E = 0.228$; $r_D = 0.156$; $p < 0.01$) or FG ($r_{TAction} = 0.248$; $r_{TSpent} = 0.194$; $r_E = 0.218$; $r_D = 0.098$; $p < 0.01$) and all analytics

Table 8 Correlation of Moodle engagement analytics with other predictors and final grade

Correlations	Total action	Time spent	Exam	Content	Discussion
AvePreTest	0.270**	0.283**	0.228**	0.053	0.156**
FG	0.248**	0.194**	0.218**	0.021	0.098**
TS	0.019	-0.026	0.033	-0.080*	0.051
MaA	0.118**	0.104**	0.093**	0.050	0.059
SRS	0.135**	0.153**	0.055	0.136**	0.074*

**Significant at the 0.01 level

*Significant at the 0.05 level

excluding C ($r_C=0.053$ and $r_C=0.021$; $p>0.05$). No significant relationship was found between TS and analytics other than C ($r_C = -0.080$; $p < 0.05$). Relationship of MaA with Total Action, Time Spent and E is significant and positive ($r_{TAction} = 0.118$; $r_{TSpent} = 0.104$; $r_E = 0.93$; $p < 0.01$). Significant relationship was found between SRS and analytics other than E ($r_{TAction} = 0.135$; $r_{TSpent} = 0.153$; $r_C = 0.136$; $r_{D=0.74}$; $p < 0.05$) (Table 8).

Table 8 showed small-level significant relationships observed between final grade, Moodle analytics, and other predictors. Moreover, although there is a significant relationship between some variables at 0.01 or 0.05 level, the relationship level is deficient. ($r_{FG-D} = 0.098$; $r_{TS-C} = -0.080$; $r_{MaA-E} = 0.093$; $p < 0.05$).

Table 9 Performance of classification

Method	AUC	CA	F1	Pre	Recall	LogLoss
Tree	0.668	0.644	0.644	0.645	0.644	0.651
Naive Bayes	0.678	0.633	0.633	0.632	0.633	0.701
Neural Network	0.652	0.625	0.624	0.624	0.625	0.813
Random Forest	0.653	0.620	0.619	0.618	0.620	0.928
SVM	0.611	0.596	0.596	0.596	0.596	0.669
CN2 rule inducer	0.593	0.584	0.584	0.583	0.584	0.831
kNN	0.567	0.515	0.490	0.562	0.515	3.127

Table 10 Classification rates of tree and Naive Bayes by actual status

Actual	Tree (Predicted)		Σ	Naive Bayes (Predicted)		Σ
	0	1		0	1	
0	60.3%	39.7%	536	58.2%	41.8%	536
1	32.2%	67.8%	673	32.7%	67.3%	673
Σ	540	669	1209	532	677	1209

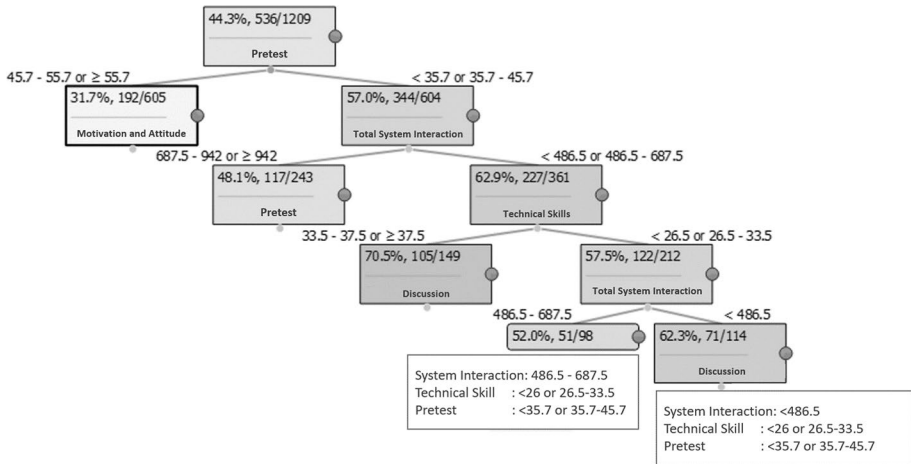


Fig. 5 Low prior knowledge, total system interaction, and technical skills

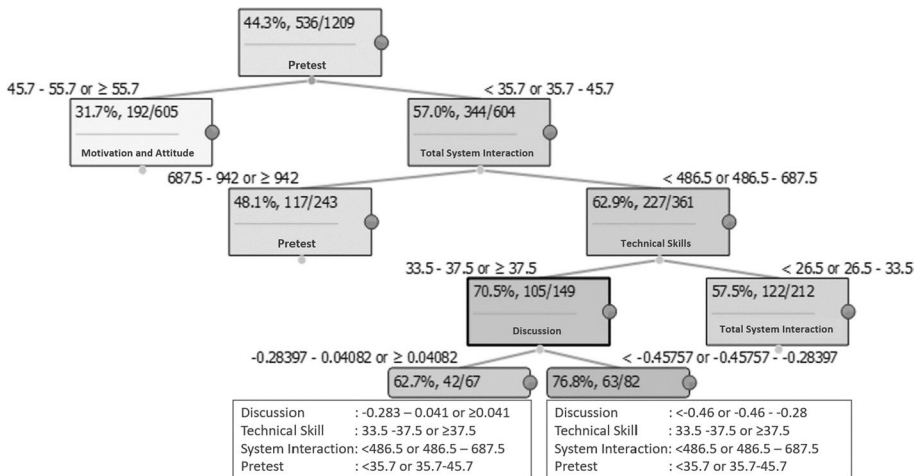


Fig. 6 "Low prior knowledge and total system interactions" and "high technical skills and high participation in the discussion"

4.3 Classification of Final Performance

The classification used to Moodle engagement analytics (total action, time spent, exam, content, discussion), learner characteristics (TS, MaA, and SRS), and prior knowledge (AvePreTest). The performance of classification was presented in Table 9.

Table 9 showed that the decision tree (CA=0.644, Pre=0.645) and Naive Bayes (CA=0.633, Pre=0.633) algorithm had the highest accuracy and precision rates. When the confusion matrix table was examined (Table 10), the decision tree correctly classified 67.8% of "learners with high performance (LwHP)" and 60% of "learners with low performance (LwLP)," Naive Bayes correctly classified 67.3% of LwHP and 58.2% of LwLP.

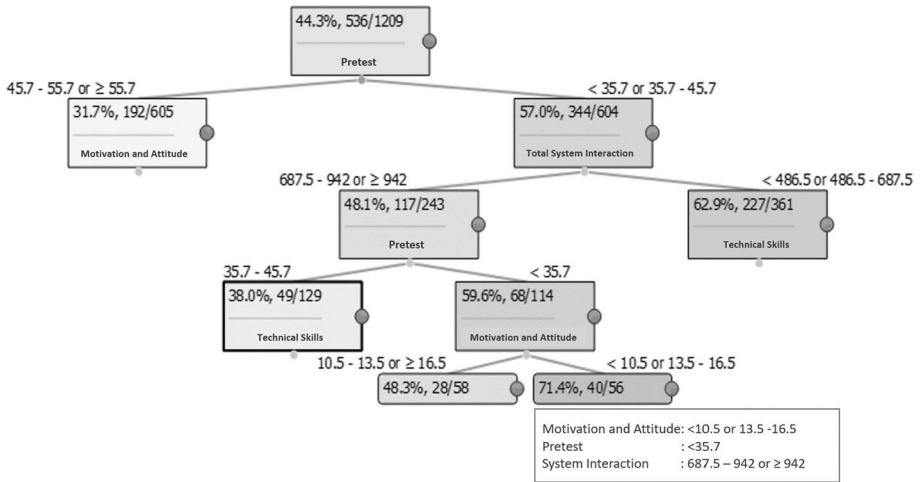


Fig. 7 "High total system interaction and low prior knowledge" and "motivation and attitude"

In Figs. 5, 6, and 7, rules created for decision trees according to cases of LwLP were presented. 57.5% of the students had low prior knowledge (pre-test average <45.7) and who had low total system interactions (<687.5), and who had not have a high level of technical skills (TS_{max} =40; TS <33.5) were LwLP. When the technical skills were greater than 33.5, the probability of LwLP increased to 70.5% (4th depth in Fig. 5). 62.3% of the students had low prior knowledge (pre-test average <45.7) and who had not a high level of technical skills (TS_{max} =40; TS <33.5) and who had a low performance of total system interactions (<486.5) were LwLP. When the total system interactions were greater than 486.5, the probability of LwLP decreased to 52% (5th depth in Fig. 5). Based on this finding, firstly, even if the students have very high technical skills, if they have a low level the prior knowledge and total system interactions, the probability of being LwLP is relatively high. Second, when the system interaction of the students with lower-level prior knowledge and technical skills increases slightly, their likelihood for LwHP may increase.

62.7% of students who had a low prior knowledge (Pretest <45.7) and who had low total system interactions (<687.5) and who had high technical skills (>33.5) and who had discussion interactions (≥ -0.28) were LwLP; whereas the proportion was higher (76.8%) for those who had lower discussion interactions (< -0.28) (5th depth in Fig. 6). Accordingly, if the learners who have low-level prior knowledge and system interactions and high technical skills participate in the discussion, the probability of LwHP may increase.

59.6% of students who had dramatically low prior knowledge (pre-test average <35.7) and who had high system interaction levels (>687.5) were LwLP. In this group, the probability of being LwLP (71.4%) of very low or high motivated students was higher than those with low or very high motivation and attitude (48.3%) (5th depth in Fig. 7). In this context, the high level of system interactions of learners with low-level prior knowledge reduces their probability of being LwLP. In terms of motivation and attitude, it is difficult to state that "motivation and attitude" increase final performance linearly.

Rules created for decision trees according to cases of LwHP were presented in Figs. 8, 9, and 10. The students who had high prior knowledge (pre-test average >45.7)

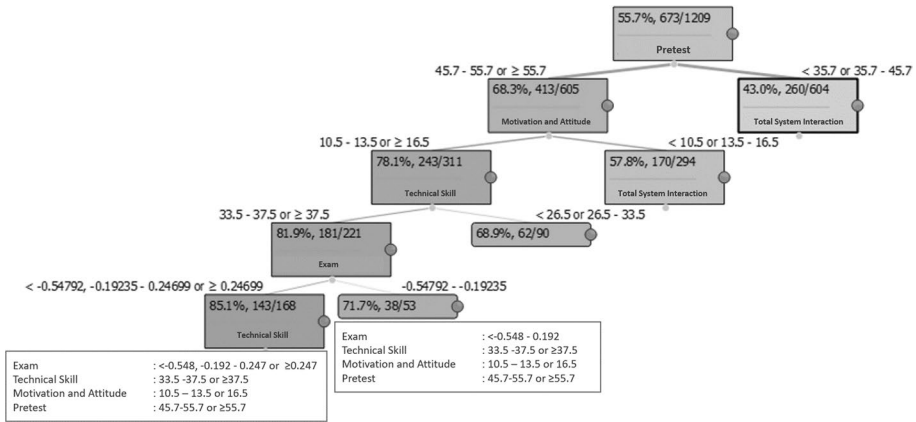


Fig. 8 High prior knowledge, "low or very high motivation," and high technical skill

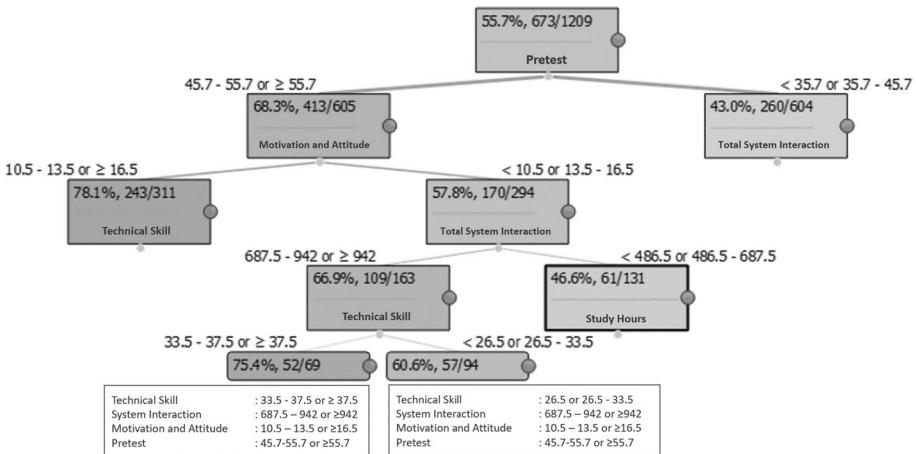


Fig. 9 High prior knowledge, "low or very high motivation," high total system interaction and technical skill

and who had low or very high motivation and attitude, and who had high technical skills mainly were successful (81.9%). In this group, the probability of being LwHP was very high (71.7%), even if their interactions with the exam were low (exam: between -0.547 and -0.192) (Fig. 8). There were some inconsistencies in terms of both exam interactions, "motivation and attitude," as in the case of LwLP. For example, those with low and very high "motivation and attitude" (78.1%) were more likely to be successful than those with very low and high "motivation and attitude" (57.8%). Typically, those with high and very

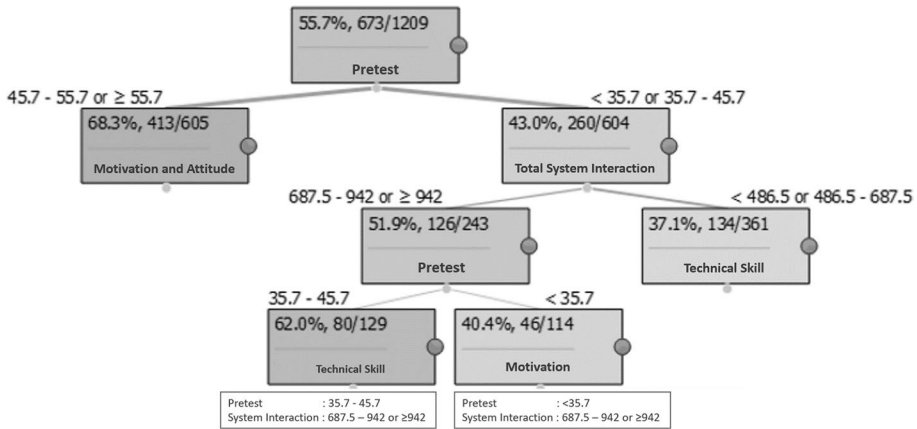


Fig. 10 High prior knowledge and system interaction

high "motivation and attitude" (78.1%) would be expected to be LwHP. In this context, motivation or exam interactions, categorically divided into four levels, may not linearly increase the probability of final performance.

If the pre-test was high, "motivation and attitude" were very low or high, and there was high total interaction, 57.8% of students were LwHP. However, 75.4% of the students in this group were LwHP when their technical skills were high. 60.6% of the students were successful when technical skills were low (Fig. 9). In this context, technical skills may have a positive effect (high prior knowledge and very low or high motivation and attitude) and a negative impact on final performance in some cases (low prior knowledge and system interaction).

51.9% of the students with low pre-test scores and high system interaction were LwHP. 40.4% of the students in this group were LwHP, where pre-test scores are deficient (<35.7). Besides, 62% of the students in this group were LwHP when the prior knowledge increased from <35.7 to 35.5–45.7. Therefore, the higher the level of prior knowledge, the higher the probability of being LwHP for students with a high total system interaction.

5 Discussions

Each of the students brings their knowledge and experience to the learning process. Throughout the learning process, they have different levels of prior knowledge, motivation, self-regulation, and other ways of interacting with the content, instructors, and peer learners. The current study revealed the following key findings based on quantitative measures to understanding learners.

The result of regression showed that prior knowledge has a significant positive impact on final performance. Yau and Ifenthaler's (2020) study, a systematic review of the previous literature, stated that while learning outcomes are predicted especially for educational institutions (e.g., universities), study history is still the most fundamental indicator. The study history includes data such as previous academic achievement, assessing learning progress with various assessment methods, or prior knowledge tests. For example, Shulruf et al. (2018) found that prior academic achievement had the most significant predictive value, with medium to substantial effect sizes (0.44–1.22) in five undergraduate medical

schools in Australia and New Zealand for predicting the binary outcomes (completing or not completing course; passing or failing examination). As expected, prior knowledge is the variable that has the most substantial effect on the final performance in the current study.

Technical skills (e-readiness dimension) include using technologies such as computers, the Internet, social networks for information search or communication. Liu (2019) confirmed that, for the orientation of students taking online courses, their technical skills should be considered part of the instructional design, social competence, working strategy, and communication dimensions. Therefore, learners with high levels of technical skills can more easily adapt to the online environment. They are more likely to have benefitted from the opportunities offered by the system. Indirectly, final performances may be expected to be positively affected. The current study found that although technical skills significantly affect the final grade, the effect size is minimal. Therefore, this finding concludes that there may not be a linear relationship between technical skills and final grades.

In MOOCs, since the effects of motivation on participation (e.g., De Barba et al., 2016) or self-regulation on goal attainment (Kizilcec et al., 2017) and autonomous control (Papamitsiou & Economides, 2019) are known, it is essential to design interventions to increase motivation (e.g., Aguilar et al., 2021; Herodotou et al., 2020) or to support self-regulated learning (e.g., Aguilar et al., 2021; Jivet et al., 2020). In the current study, it is difficult to state that "motivation and attitude" (an e-readiness dimension) or self-regulation skills increase or decrease success linearly. In multiple regression analyses, "motivation and attitude" and self-regulation skills also had no role in predicting final performance. Moreover, in classification analysis, both variables were found unlikely to play a role in classifying learners with high performance (≥ 70). These findings may be explained in course-related characteristics (e.g., position relative to other courses, course structure) based on the classification of learning analytics indicators paired with three data profiles (Yau & Ifenthaler, 2020).

The context of the ICT course is not a situation where the learner chooses a course that appeals to individual interest, supporting professional development, as in MOOC settings. The course (ICT101) is planned only for asynchronous delivery within the first year of undergraduate university programs. Learners had face-to-face learning experiences in other courses. Therefore, no matter how motivated the students may be, they may have faced particular challenges adapting to the lesson delivery style or medium or did not care much for the lesson/course. Accordingly, they may have experienced lower levels of motivation as the course weeks progressed.

When a sufficient variety of learning resources is given, and rich learning experiences are offered, students can prefer different media and learning paths. Some students may choose handouts in the current study, while others prefer videos or interactive activities in the ICT course. Therefore, contrary to the current study's findings, learners with higher self-regulation skills could be expected to be more successful in a course generally offered at this level of diversity. However, this research does not show that self-regulation skills are not crucial for the final performance. Viberg et al. (2020) draw attention to the insufficient number of studies (20%) showing evidence of improvement in learning outcomes for LA (including interventions to support SRL). However, researchers state that the evidence in these studies has been slight so far, and there is a generalization problem. Therefore, supporting self-regulation skills may not guarantee high learning outcomes in the current study context (e.g., ICT course structure, assessment structure). For example, linearly designed content divided into weeks in the course (ICT101) may contribute positively to

learners with low self-regulation skills. For this reason, the effect of self-regulation skills may not have been reflected in the final performance.

Moreover, a rich learning experience in both synchronous and asynchronous activities should be provided for a practical learning experience in the design of online courses by replicating classroom biases pedagogies (e.g., only content transfer oriented). These learning experiences can occur in various forms, such as acquiring, researching, applying, producing, discussing, and collaborating (Laurillard et al., 2013). In addition, the interaction of learners with assessment activities is also a learning experience (Holmes et al., 2019). The activity that the student will do for each learning experience is also different. In this study, the learning experiences in the ICT course were designed with assimilation activities (watching interactive videos, reading notes) and assessment activities (pre-test and post-test for each topic). The learner is expected to demonstrate the required minimum 70% success for each subject by repeating the assimilation activities and using their experience in the post-test (can be repeated an unlimited number of times). This design does not offer flexibility in the learner's use of self-regulation skills. Everything is evident in this design. Students with high or low self-regulation skills are also forced into the same experience. However, the role of motivation or self-regulation skills can be felt more when the learner is offered more autonomy for a lesson designed with activities such as producing, discussing, and collaborating more.

Although a significant low-level relationship, the correlation result showed that it is worth exploring further the role of system interaction on the final performance. Ordinarily, it can be expected to be higher relationships between the interactions of content, exams, or discussion and final performance. However, there was a low-level positive significant relationship between students' final grades and the interactions of exams and discussions in online lessons. In the literature, both parallel (e.g., Schumacher & Ifenthaler, 2021; Strang, 2016) and opposing results (e.g., Saqr et al., 2017) can be found compared to the current study results. For example, Schumacher and Ifenthaler (2021) investigated whether trace data can inform learning performance. The researchers found that "only participants' number of views of the handout was a significant predictor of their learning performance in the transfer test and stated, "trace data did not, as expected, provide explanation for learning performance" (Schumacher & Ifenthaler, 2021, pp. 10–11). However, another study found that the Moodle engagement analytics showed significant positive correlations with student performance, especially for parameters that reflected motivation and self-regulation (Saqr et al., 2017). In parallel with the current study, this difference may be considered in terms of the impact of learning design (Er et al., 2019; Holmes et al., 2019) on both performance and learner behavior in the online learning environment.

While traditional instructional design focuses on content transfer, learning design (or the new interpretation of instructional design based on constructivist theories) focuses on activities in the learning process (Holmes et al., 2019; van Merriënboer & Kirschner, 2017). Learning design is broadly referred to as the design of sequences of learning activities (such as reading texts, analyzing data, practicing exercises, producing videos, participating in discussion forums, or collaborating in group projects) in line with the activity's aims, outcomes, teaching methods, assessment, learning approach, duration, and necessary resources (Holmes et al., 2019). Although the impact of a particular learning design on course success has not been observed, its effect on learner behavior has been proven (Holmes et al., 2019). Therefore, the fact that the course design applied in the current study is content-transfer-oriented and linear may have caused the system interactions of learners to be seen as similar.

Regression and correlation analysis to explain the final performance based on the research findings of RQ1 and RQ2 showed the importance of e-readiness (in terms of technical skills) and prior knowledge. It showed that "motivation and attitude" and self-regulation skills were not significant in the context of the ICT course. However, these analyses indicate a linear relationship between the variables that are not considered essential and the final performance. In this context, testing the classification models that predict the final performance over *probability* may emphasize the importance of some variables in terms of the final performance.

In the classification model generated by using learners' prior knowledge, Moodle LMS engagement analytics, and learner's characteristic variables, 67.8% of learners were correctly classified according to best probability. Strang (2017) showed that the effect of Moodle engagement analytics on assessment grades was high. For example, the General Linear Model with four online interaction predictors captured 77.5% of grade variance within an undergraduate business course. On the contrary, Conijn et al. (2017) revealed that the purposes of early intervention or when in-between assessment grades were taken into account, the LMS data proved to be of little value. In the current study, the results were obtained parallel to Conijn et al. (2017). For example, the effect of Moodle Engagement Analytics (partially excluding discussion interactions) was not observed to classify the performance.

The partial effect of discussion interaction was observed in low-level prior knowledge and system interactions and highly technical skills. For example, when the learners participated more in the discussion, the possibility of low performance (< 70) decreased (from 76.8 to 62.7%). On the other hand, when technical skills are combined with high prior knowledge, most learners have high performance (81.9%). In this context, students with high technical skills may not need much effort to be successful by comparing themselves with the class in the general discussions. If learners with low-level prior knowledge have a high level of total system interactions, their probability of high-performance increases from 40.4 to 60.2%. When these findings are evaluated together, the computer literate students are expected to score higher on their pre-tests and pass the course with less effort when compared to new beginners to the topic. So, prior knowledge, which is consistent with previous research (Duffy & Azevedo, 2015; Moos & Azevedo, 2008), can be stated as increasing the probability of success. Nevertheless, system interactions alone may not be sufficient for high performance.

6 Conclusions, Limitations, and Future Research

This study, in an ICT course and Turkey context, examined indicators that affect final performance within an online learning environment by investigating learners' characteristics (e.g., technical skills, "motivation and attitude," self-regulation), assessment scores (e.g., pre-test scores), and behaviors in using an LMS (Moodle engagement analytics). Consequently, final performance seems to be related to but not "motivation and attitude" and self-regulation skills. The effect of prior knowledge, system interactions, and technical skills on the final performance should be viewed as a typical situation. The contribution of this research is that it shows that leading reference indicators (e.g., self-regulation and motivation) in most learning analytics research may not impact final performance. However, it should not be understood from the results of this research that self-regulation and motivation are unimportant in online learning. It would be helpful to look at results in the context of the characteristics of the course and the differences in learning design.

Learning analytics for online learning is primarily based on motivation and self-regulated learning (Wong et al., 2019). Self-regulated learning that the student controls, monitors, and influences their thinking and learning process requires knowledge and skills (Kocdar et al., 2018). This approach is often used by learning analytics researchers in intervention design, provided that the responsibility for learning is not left entirely to the student. Nevertheless, recent research found little evidence for LA that improved learning outcomes (Schumacher & Ifenthaler, 2021; Viberg et al., 2020). For example, Viberg et al. (2020) found evidence that the implementation of LA to support students' SRL improves their learning outcomes in only 20% of the systematically reviewed papers ($n=11$). The researchers suggested the LA research should focus more on measuring different SRL parts rather than on its support to improve learner performance—one of the LA's key goals—for online learning. Schumacher and Ifenthaler (2021) found that the prompts based on self-regulation (e.g., cognitive, motivational prompts) might have only limitedly impacted declarative knowledge and knowledge transfer. The researchers stated, "prompts might not have been efficient, as they were not related to students' characteristics or behavior, resulting in inappropriate support." (Schumacher & Ifenthaler, 2021, pp. 11). The current study supports recent researches (Schumacher & Ifenthaler, 2021; Viberg et al., 2020) and, unlike these studies, discuss interpreting learning analytics with learning design (Macfadyen et al., 2020; Mangaroska & Giannakos, 2019; Lockyer & Dawson, 2011). The problem points to the situation in which the indicator should be given importance rather than which indicator is necessary. Instructors/researchers, therefore, should focus more on learning design in terms of LA usage, intervention, or which indicators they should use.

This research was a small-scale, techno-centric, and exploratory study designed to deepen understanding of the online learning process in ICT course and Turkey. When considering insights from this study, actions can be planned to increase system interactions for learners with low levels of prior knowledge in the current course design (linear structure, only asynchronous, video-based content transfer, and pretest–posttest assessment). For example, these students can be provided with extra materials to fill their knowledge and skills gaps. Those with low levels of system interaction can be alerted weekly via email. Moreover, dynamic visualization tools, where learners can compare themselves with system interactions of high performance, can be integrated into the system. However, it would be helpful to repeat that it cannot be generalized in different course designs (e.g., discussion-intense course designs or productive activity intense course designs; Holmes et al., 2019). By researching different course designs that offer more flexible opportunities for learners, the effects of other dimensions of e-readiness (e.g., motivation) or self-regulation on final performance and final performance may be investigated again. Moreover, student opinions may determine variables that can affect performance, and the resulting variables may be added to regression and classification analysis.

This study has limitations in a different aspect other than generalizability. For example, in the analyses made for classifying the final performance, various methods (such as using indicators as categorical or continuous variables or the number of categories of indicators) have been tried to achieve the best estimation results. It is known in the literature that different techniques show different classification performances (Cui et al., 2020; Hung et al., 2019). Therefore, this study is limited to the data transformation steps and classification algorithms it uses. When different methods and techniques are applied, the findings related to variables (e.g., motivation and attitude) whose significance cannot be consistently shown in this study may differ.

Acknowledgements This research was supported and funded by the Scientific and Technological Research Council of Turkey (TUBİTAK).

Data Availability Statement The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethical Approval The authors of the current study have approved the manuscript for submission. All procedures performed in the current study involving human participants were in accordance with the ethical standards of the institutional research committee. With regard to ethical considerations, a consent form was prepared and explained to the participants both while collecting log data and applying the scales. Written informed consent was also obtained from the all participants prior to the research in LMS.

References

- Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, *162*, 104085. <https://doi.org/10.1016/j.compedu.2020.104085>
- Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, *20*(2), 13–29. <https://doi.org/10.24059/olj.v20i2.790>
- Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and meta-cognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, *45*(4), 210–223. <https://doi.org/10.1080/00461520.2010.515934>
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5–6), 318–331. <https://doi.org/10.1504/IJTEL.2012.051815>
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, *18*(6), 683–695. <https://doi.org/10.1080/13562517.2013.827653>
- Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. <https://doi.org/10.1109/TLT.2016.2616312>
- Council of Higher Education. (2020). *Procedures and principles regarding distance education in higher education institutions* [Yükseköğretim Kurumlarında Uzaktan Öğretime İlişkin Usul Ve Esaslar]. Retrieved from <https://t.ly/xX7H>
- Cui, Y., Chen, F., & Shiri, A. (2020). Scale up predictive models for early detection of at-risk students: A feasibility study. *Information and Learning Sciences*, *121*(3/4), 97–116. <https://doi.org/10.1108/ILS-05-2019-0041>
- De Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning*, *32*(3), 218–231. <https://doi.org/10.1111/jcal.12130>
- Du, X., Yang, J., Shelton, B. E., Hung, J. L., & Zhang, M. (2019). A systematic meta-review and analysis of learning analytics research. *Behaviour & Information Technology*. <https://doi.org/10.1080/0144929X.2019.1669712>
- Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348. <https://doi.org/10.1016/j.chb.2015.05.041>
- Er, E., Gómez-Sánchez, E., Dimitriadis, Y., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Álvarez-Álvarez, S. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments*, *27*(5–6), 685–698.
- Gašević, D., Dawson, S., Rogers, T., & Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting learning success. *The Internet and Higher Education*, *28*, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>

- Gülbahar, Y. (2012). Study of developing scales for assessment of the levels of readiness and satisfaction of participants in e-learning environments. *Ankara University, Journal of Faculty of Educational Sciences*, 45(2), 119–137. https://doi.org/10.1501/Egifak_0000001256
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Advanced diagnostics for multiple regression: A supplement to multivariate data analysis*. Prentice Hall.
- Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q. (2018). Predicting academic performance by considering student heterogeneity. *Knowledge-Based Systems*, 161, 134–146. <https://doi.org/10.1016/j.knosys.2018.07.042>
- Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., & Rienties, B. (2020). How can predictive learning analytics and motivational interventions increase student retention and enhance administrative support in distance education? *Journal of Learning Analytics*, 7(2), 72–83. <https://doi.org/10.18608/jla.2020.72.4>
- Holmes, W., Nguyen, Q., Zhang, J., Mavrikis, M., & Rienties, B. (2019). Learning analytics for learning design in online distance learning. *Distance Education*, 40(3), 309–329. <https://doi.org/10.1080/01587919.2019.1637716>
- Horzum, M. B., Kaymak, Z. D., & Gungoren, O. C. (2015). Structural equation modeling towards online learning readiness, academic motivations, and perceived learning. *Educational Sciences: Theory and Practice*, 15(3), 759–770. <https://doi.org/10.12738/estp.2015.3.2410>
- Hung, J. L., Shelton, B. E., Yang, J., & Du, X. (2019). Improving predictive modeling for at-risk student identification: A multistage approach. *IEEE Transactions on Learning Technologies*, 12(2), 148–157. <https://doi.org/10.1109/TLT.2019.2911072>
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Iglesias-Pradas, S., Ruiz-de-Azcárate, C., & Agudo-Peregrina, A. F. (2015). Assessing the suitability of student interactions from Moodle data logs as predictors of cross-curricular competencies. *Computers in Human Behavior*, 47, 81–89. <https://doi.org/10.1016/j.chb.2014.09.065>
- Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachslers, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *The Internet and Higher Education*, 47, 100758.
- Joosten, T., & Cusatis, R. (2020). Online learning readiness. *American Journal of Distance Education*. <https://doi.org/10.1080/08923647.2020.1726167>
- Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, 104, 18–33. <https://doi.org/10.1016/j.compedu.2016.10.001>
- Kocdar, S., Karadeniz, A., Bozkurt, A., & Buyuk, K. (2018). Measuring self-regulation in self-paced open and distance learning environments. *International Review of Research in Open and Distributed Learning*, 19(1), 25–43. <https://doi.org/10.19173/irrodl.v19i1.3255>
- Laurillard, D., Charlton, P., Craft, B., Dimakopoulos, D., Ljubojevic, D., Magoulas, G., Masterman, E., Pujadas, R., Whitley, E. A., & Whittlestone, K. (2013). A constructionist learning environment for teachers to model learning designs. *Journal of Computer Assisted Learning*, 29(1), 15–30. <https://doi.org/10.1111/j.1365-2729.2011.00458.x>
- Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—A literature review. In A. Peña-Ayala (Ed.), *Learning analytics: Fundamentals, applications, and trends* (pp. 1–23). Springer. https://doi.org/10.1007/978-3-319-52977-6_1
- Liu, J. C. (2019). Evaluating online learning orientation design with a readiness scale. *Online Learning*, 23(4), 42–61. <https://doi.org/10.24059/olj.v23i4.2078>
- Lockyer, L., & Dawson, S. (2011). Learning designs and learning analytics. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 153–156). <https://doi.org/10.1145/2090116.2090140>
- Lu, C.-H., & Yu, C.-H. (2019). Online data stream analytics for dynamic environments using self-regularized learning framework. *IEEE Systems Journal*, 13(4), 3697–3707. <https://doi.org/10.1109/JSYST.2019.2894697>
- Macayan, J. V. (2017). Implementing outcome-based education (OBE) framework: Implications for assessment of students' performance. *Educational Measurement and Evaluation Review*, 8(1), 1–10.
- Macfadyen, L. P., Lockyer, L., & Rienties, B. (2020). Learning design and learning analytics: Snapshot 2020. *Journal of Learning Analytics*, 7(3), 6–12. <https://doi.org/10.18608/jla.2020.73.2>
- Mangaroska, K., & Giannakos, M. (2019). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, 12(4), 516–534. <https://doi.org/10.1109/TLT.2018.2868673>

- Miranda, S., & Vegliante, R. (2019). Learning analytics to support learners and teachers: The navigation among contents as a model to adopt. *Journal of e-Learning and Knowledge Society*, 15(3), 101–116. <https://doi.org/10.20368/1971-8829/1135065>
- Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology*, 33(2), 270–298. <https://doi.org/10.1016/j.cedpsych.2007.03.001>
- Moreno-Marcos, P. M., Munoz-Merino, P. J., Maldonado-Mahauad, J., Perez-Sanagustin, M., Alario-Hoyos, C., & Kloos, C. D. (2020). Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs. *Computers & Education*, 145, 103728. <https://doi.org/10.1016/j.compedu.2019.103728>
- Mwalumbwe, I., & Mtebe, J. S. (2017). Using learning analytics to predict students' performance in Moodle learning management system: A case of Mbeya University of Science and Technology. *The Electronic Journal of Information Systems in Developing Countries*, 79(1), 1–13. <https://doi.org/10.1002/j.1681-4835.2017.tb00577.x>
- Namoun, A., & Alshanjiti, A. (2021). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, 11(1), 237. <https://doi.org/10.3390/app11010237>
- Olive, D. M., Huynh, D. Q., Reynolds, M., Dougiamas, M., & Wiese, D. (2019). A quest for a one-size-fits-all neural network: Early prediction of students at risk in online courses. *IEEE Transactions on Learning Technologies*, 12(2), 171–183. <https://doi.org/10.1109/TLT.2019.2911068>
- Omedes, J. (2018). *Learning analytics 2018—An updated perspective*. IAD Learning. <https://www.iadlearning.com/learning-analytics-2018/>
- Papamitsiou, Z., & Economides, A. A. (2014). Temporal learning analytics for adaptive assessment. *Journal of Learning Analytics*, 1(3), 165–168. <https://doi.org/10.18608/jla.2014.13.13>
- Papamitsiou, Z., & Economides, A. A. (2019). Exploring autonomous learning capacity from a self-regulated learning perspective using learning analytics. *British Journal of Educational Technology*, 50(6), 3138–3155. <https://doi.org/10.1111/bjet.12747>
- Pardo, A., Han, F., & Ellis, R. A. (2016). Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. *IEEE Transactions on Learning Technologies*, 10(1), 82–92. <https://doi.org/10.1109/TLT.2016.2639508>
- Premalatha, K. (2019). Course and program outcomes assessment methods in outcome-based education: A review. *Journal of Education*, 199(3), 111–127. <https://doi.org/10.1177/0022057419854351>
- Ramirez-Arellano, A., Bory-Reyes, J., & Hernández-Simón, L. M. (2019). Emotions, motivation, cognitive–metacognitive strategies, and behavior as predictors of learning performance in blended learning. *Journal of Educational Computing Research*, 57(2), 491–512. <https://doi.org/10.1177/0735633117753935>
- Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher*, 39(7), 757–767. <https://doi.org/10.1080/0142159X.2017.1309376>
- Schumacher, C., & Ifenthaler, D. (2021). Investigating prompts for supporting students' self-regulation—A remaining challenge for learning analytics approaches? *The Internet and Higher Education*, 49, 100791. <https://doi.org/10.1016/j.iheduc.2020.100791>
- Shulruf, B., Bagg, W., Begun, M., Hay, M., Lichtwark, I., Turnock, A., Warnecke, E., Wilkinson, T. J., & Poole, P. J. (2018). The efficacy of medical student selection tools in Australia and New Zealand. *Medical Journal of Australia*, 208(5), 214–218. <https://doi.org/10.5694/mja17.00400>
- Strang, K. D. (2016). Do the critical success factors from learning analytics predict student outcomes? *Journal of Educational Technology Systems*, 44(3), 273–299. <https://doi.org/10.1177/0047239515615850>
- Strang, K. D. (2017). Beyond engagement analytics: Which online mixed-data factors predict student learning outcomes? *Education & Information Technology*, 22, 917–937. <https://doi.org/10.1007/s10639-016-9464-2>
- Sun, J. C. Y., Lin, C. T., & Chou, C. (2018). Applying learning analytics to explore the effects of motivation on online students' reading behavioral patterns. *International Review of Research in Open and Distributed Learning*. <https://doi.org/10.19173/irrodl.v19i2.2853>
- van Merriënboer, J. J. G., & Kirschner, P. A. (2017). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Routledge. <https://doi.org/10.4324/9780203096864>
- Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 524–533). <https://doi.org/10.1145/3375462.3375483>
- Wong, J., Baars, M., de Koning, B. B., van der Zee, T., Davis, D., Khalil, M., Houben, G., & Paas, F. (2019). Educational theories and learning analytics: From data to knowledge. In D. Ifenthaler, D.-K.

- Mah, & J.Y.-K. Yau (Eds.), *Utilizing learning analytics to support study success* (pp. 3–25). Springer. https://doi.org/10.1007/978-3-319-64792-0_1
- Yau, J. Y. K., & Ifenthaler, D. (2020). Reflections on different learning analytics indicators for supporting study success. *International Journal of Learning Analytics and Artificial Intelligence for Education: IJAI*, 2(2), 4–23. <https://doi.org/10.3991/ijai.v2i2.15639>
- Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior*, 70, 251–260. <https://doi.org/10.1016/j.chb.2016.12.085>
- You, J. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education*, 29, 23–30. <https://doi.org/10.1016/j.iheduc.2015.11.003>
- Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *Internet and Higher Education*, 27, 44–53. <https://doi.org/10.1016/j.iheduc.2015.05.002>

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