ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE

Paper ID #37868

Predicting learning outcome in a first-year engineering course: a human-centered learning analytics approach

Laura Melissa Cruz Castro

Laura M. Cruz Castro is a Ph.D. candidate in the Department of Engineering Education under the guidance of Dr. Kerrie Douglas. She holds a B.S. in Statistics from Universidad Nacional de Colombia, a M.S. in Business Intelligence from Universität de Barcelona, and a M.S in Electrical and Computer Engineering from Purdue University. Her research interests include educational data analytics, ethical considerations regarding the use of data in education, incorporation of the Data Science curriculum in computing education, and broadening participation of underrepresented populations in computing professions. She is currently a dean's teaching fellow for the College of Engineering at Purdue University.

Tiantian Li

Tiantian Li (Olivia) is a PhD student in Engineering Education at Purdue University. She is a Purdue graduate with a Bachelor of Science degree in Biological Engineering, with a concentration in Pharmaceutical Processing Engineering. She has completed Purdue's Certificate of Systems Engineering and Quantitative Research, Assessment, and Evaluation in Education Certificate. Her research interest is in the assessment of systems thinking skills and systems awareness. She is also interested in studying international scholars' experience within engineering.

Leyla Ciner

Kerrie A Douglas (Assistant Professor of Engineering Education)

Dr. Kerrie Douglas, Assistant Professor of Engineering Education at Purdue, studies how to improve the quality of classroom assessments and evaluation of online learning in a variety of engineering education contexts. She holds a PhD in Educational Psychology and a M.A. in Educational Studies, with focus on school counseling. She is a co-PI on the SCALE project, leading the evaluation and assessment efforts. She recently received an NSF award to study engineering instructor decisions and student support during COVID-19 and impact the pandemic is having on engineering students. She also recently won the prestigious CAREER award from the U.S. National Science Foundation to study increasing the fairness of engineering assessments. In total, she has been on the leadership of more than $24 million dollars in research awards. Her research on evaluation of online learning (supported by two NSF awards #1544259,1935683,) has resulted in more than 20 peer-reviewed conference and journal publications related to engineering learners in online courses. She was a FutureLearn Research Fellow from 2017-2019; a 2018 recipient of the FIE New Faculty Fellow Award and was the 2021 Program Chair for the Educational Research Methods Division of ASEE.

Christopher Greg Brinton

Predicting learning outcome in a first-year engineering course: a human-centered learning analytics approach

Abstract:

First-year engineering courses are relatively large with several sections; thus, it can be rather difficult for an individual instructor to recognize when a particular student begins to lose engagement. Learning management systems (LMS) (e.g., Canvas, Blackboard, Brightspace) can be valuable tools to provide a consistent curriculum across several sections of a course and generate data regarding students' engagement with course materials. However, a human-centered approach to transform the data needs to be utilized to extract valuable insights from LMS data.

The purpose of this Complete Research paper is to explore the following research questions: What type of LMS objects contain information to explain students' grades in a first-year engineering course? Is the inclusion of a human operator during the data transformation process significant to the analysis of learning outcomes? For this, data from LMS is used to predict the learning outcome of students in a FYE course. Two predictive models are compared. The first model corresponds to a usual predictive model, using the data from the LMS directly. The second model considers the specifics of the course, by transforming the data from aggregate user interaction to more granular categories related to the content of the class by a human operator. A logistic regression model is fitted using both datasets. The comparison between predictive measures such as precision, accuracy, and recall are then analyzed.

The findings from the transformed dataset indicate that students' engagement with the career exploration curriculum was the strongest predictor of students' final grades in the course. This is a fascinating finding because the amount of weight the career assignments contributed to the overall course grade was relatively low. Additionally, while both models produced adequate fit indices, the human-informed model performed significantly better and resulted in more interpretable results.

Introduction and Background

First-year engineering programs

As the 21st century approached, Bordogna and Ernst claimed that the engineering education paradigm shift needed to consider integration [1]. From a philosophical point of view, the authors argued that the field of engineering education was prepared to shift from a disconnected curriculum to an integrated curriculum. A more integrated curriculum would help students appreciate the importance of complexity and reflect the disciplinary integration of the engineering profession [1]. From their perspective, integration meant holistic education, an understanding that all knowledge acquired was with an engineering purpose. From this paradigm, first-year engineering (FYE) programs have been established as an essential part of engineering education and have significantly contributed to student success [2].

Before institutions established FYE programs, engineering students entered engineering and took classes focused heavily on science without direct connections to engineering[1]. Some FYE programs, also called Freshmen Engineering Programs, used approaches to integrate a set of courses from science, humanities, and engineering (e.g.,[3]). While others consisted of a set of engineering courses that exposed students to engineering early on without explicit integration with other science and humanities courses (e.g., [4]). Additionally, while some programs were designed to host students across all engineering disciplines, others were created for specific disciplines (e.g., [5], [6]).

Over the years, many FYE programs have been established. These programs are different across institutions, varying in length, the number of credits, and the disciplines that the students are prepared to enter [7]. Nonetheless, FYE programs have proven successful for engineering students. For instance, when comparing two different programs at different institutions, Richardson and Dantzlet [8] found that the retention and academic performance of students were improved by their participation in FYE programs. In addition, a study involving 28 years of data, caused Budny and colleagues to find a link between FYE completion and graduation success [9]. Marra and other researchers, after analyzing the data from a cross-sectional and longitudinal study, concluded that there was an improvement in intellectual development when comparing students who had completed an FYE course to their counterparts [10]. Additionally, retention, success, and motivation are some of the improved metrics that are attributed to the establishment of FYE programs at specific institutions [11]–[13].

Since the establishment of the first FYE programs, they have been transformed and updated to accommodate the changing needs of students and the needs of the environment [14], as well as the need to fulfill external factors such as accreditation standards [15]. In the current times, it is common that large FYE programs receive thousands of students each semester, and the content and strategies used are the result of a long history of amendments. For instance, the increasing number of students in some programs led to multi-sectional courses [16], while some of the ABET renewed criteria drove programs to include ethics and professional skills in FYE [15]. Therefore, FYE programs are complex. For example, when preparing students for different disciplines, FYE programs have multiple stakeholders with a wide range of interests [17]. Therefore, the required content for these courses can face constant changes and can be over-extended [17]. Furthermore, FYE courses are often large. Whether the approach is to have a shared large lecture or multiple sections, large enrollment is a factor that affects FYE programs. For instance, large enrollment implies a need to dynamically organize instructional support teams, graders, teaching assistants, and instructors, among others [18].

Need for content evaluation of FYE

The complexity of FYE programs related to multiple stakeholders and large enrollment affects the ability to evaluate the content and teaching strategies used. For example, attending to various stakeholders can reflect on constant changes related to the content, which may be difficult for the deployment of interventions to evaluate changes to the courses. In addition, the prominent enrollment aspect jeopardizes the ability to form student-instructor relationships. Large enrollment hinders the communication between instructors and students [19], limiting the ability of the instructors to gauge the understanding or situation of their students. Consequently, limited communication restrains FYE program instructors' and instructional support teams' knowledge of how changes in teaching strategies used and content selected affect students' experiences.

All the evaluations made for FYE bring a common understanding of the good that this type of program brought to engineering education, along with the need for a careful evaluation of the content and strategies used in FYE programs. FYE programs receive thousands of students each semester, with varying levels of academic preparation and prior knowledge about engineering and higher education [7]. These two aspects, enrollment size and diversity of students' backgrounds of FYE programs, require greater responsibility from course designers, instructional teams, and stakeholders. FYE students are in a crucial stage of their academic career, with their first-year experiences affecting their motivation to continue and their success [7]. Due to the importance of FYE programs' role in student development and the programs' complexity, there is a need to assist instructors and instructional support teams with useful information to make both course- and program-level decisions.

Although an evaluation of FYE programs at the content level is necessary, finding information for such evaluation can be challenging. For instance, some content changes that are incorporated into FYE programs are done for a short period of time, further complicating the timely evaluation of the content. Nevertheless, with the increasing use of data from Learning Management Systems (LMS), components that are critical for students' success in their FYE experience can be understood.

LMS data and data analytics as possibilities for the evaluation of FYE content

In recent years, the use of LMS has increased in higher education, with LMS becoming a standard platform of communication between instructors and students [20]. The use of LMS data creates a possibility to regularly evaluate the content and strategies on FYE programs. LMS data can provide granular information about the use of materials, student access to specific content, and interaction with discussion forums, among other things [21]. Additionally, the use of LMS in large-enrollment classrooms is beneficial since these systems allow for a standard structure across multi-sectional classes and multiple instructional users [22]. Therefore, the current use and possibility for data retrieval from LMS can contribute to the regular evaluation of FYE programs.

Previous research with LMS data has focused on predictive analytics, emphasizing the performance of the machine learning models [23]. However, given the complexity of FYE programs, we hypothesize that using a human-centered approach will generate valuable insights into the evaluation of FYE content and strategies. In learning analytics, a human-centered approach refers to approaching learning analytics solutions while constantly responding to the *why* and *how* humans will use the solution [24]. The main benefit of using a human-centered approach is that consideration for how the results can be interpreted is part of the design of the model, whereas non-human centered machine learning models often have results that are difficult to interpret [24]. One of the most popular approaches to analyzing LMS data in learning analytics has addressed the use of data collected in the LMS to predict learning outcomes (typically in the form of learner's satisfaction or grade) [25]. The importance of such analysis has been justified by the possibility of predictions evolving into a vital tool to find at-risk students and identify student behaviors that can lead to success [25]. However, although learning outcome predictions offer much insight into educational research, they are not without limitations.

Despite the potential of learning outcome prediction, at least two limitations have accompanied the significance of such studies. First, the predictions made seldomly address the fact that they were system- and context-dependent, meaning that the variables collected were unique to a specific course and a particular platform [26]. This limitation influenced the generalizability of the prediction models, which was scarcely discussed in publications. Second, in the majority of the cases, the techniques used in these predictions did not return many insights on education itself and only focused on the methodologies used for predictions, making it difficult to understand the students' behaviors that lead to success [26]–[28]. These limitations contradicted the original purpose of LMS, which were designed primarily to help educators supervise the learning process of individuals and organizations [29]. As a result, such predictions offered limited assistance to instructors in using the outcomes to make data-informed and diagnostic decisions in their courses.

Importance of context and a Human-Centered Learning Analytics approach to evaluate FYE content

In addition to these prediction studies using LMS data, some other studies have offered a good understanding of the importance of the context in the ability to use predictive models in FYE programs. For example, Marbouti and colleagues found that almost all predictive studies were not easily reused [30]. Their proposal then was to use students’ current academic performance to build predictive models. For this purpose, the researchers used semester performance as a predictor. The focus of their research was to design an accurate model with information available to the instructors at the beginning of the semester. They found that their model was able to predict 82.6% of the students at risk only with the information from two weeks of the semester. They also discussed the importance of reliable grading for the use of this approach and the contextual nature of such models.

Marbouti and colleagues [25], [30], [31] bring to the front the issue of data quality and data usability. Although identified as one of the most critical factors in the success of learning analytics approaches, data quality is still one of the aspects that fail to garner enough attention from researchers [32]. As a result, in this study, we emphasize data transformation, which is critical to the use of LMS data in predictive studies. Data transformation is viewed as one vital aspect of data quality, which refers to transforming raw data into variables on an intermediate level to fulfill different objectives (e.g., standardize, summarize). Under complex conditions, such as predicting learning outcomes, data transformation requires human intervention. For instance, Human Factors Engineering researchers argue that even though one could think that humans are not involved in some machine learning and data science processes, better solutions could be achieved in a hybrid space in which human operators are considered as processing elements and interpreters [33].

Among the various predictive models built with LMS data, most have been used to mainly identify at-risk students; however, we argue that another purpose of predictive models could be to evaluate the content of the course and inform instructional teams of such results. For instance, we assert that, from a learner-centered perspective, the ultimate goal of using these predictions is not only to find at-risk students but also to help instructors improve their courses. Therefore, we advocate for an understanding of the prediction variables so that instructional teams are aware of the content and strategies that are more valuable to the students and advocate for the emphasis on such objects.

The purpose of this paper is to evaluate the content of an FYE course from a human-centered approach using LMS data. For this purpose, we explore the following questions:

(1) What type of LMS content objects contain information that can explain students' grades in a first-year engineering course?
(2) Is the inclusion of a human operator during the data transformation process significant to analyzing learning outcomes?

Framework: Human-centered learning analytics

According to Shum and colleagues, human-centeredness in learning analytics (HCLA) means taking into account the range of users that will engage, interact, and use the data, as well as the circumstances of such user-tool interactions [24]. In their summary, Shum and colleagues specify that HCLA needs to account for human factors, considering at least *Why* and *How* analytics tools will be used. As we progress with the use of data to evaluate the content and strategies of FYE programs, it is critical to use a human-centered approach. For instance, we perform the analysis from the perspective that the benefits of data analysis on LMS data should be for instructors, programs, and students. In addition, we will ensure that our results are of use to this specific program, accounting for the context and human knowledge of the course.

For this purpose, we will compare two different approaches. In the first approach, we will use a traditional prediction approach. In the second approach, we will use a human-centered approach. The second approach is human-centered due to two conditions. First, the objective responds to the stakeholders directly; we are performing a predictive analysis on FYE program data to obtain valuable insights that help us evaluate the content of an FYE course. In addition, the stakeholders will be able to use these insights directly in the classroom by promoting higher student access of the content tagged as relevant and identifying students early on that are not complying with the access of these materials. Second the complexity of the context is accounted for by transforming the data with the help of an expert grader of the class to reflect the main components of the classroom.

Methods:

Context and participants

The course this study used was an FYE course in Spring 2019 with approximately 187 students divided into two main sections at a midwestern university. Section 1 had a total of 100 students, while Section 2 had a total of 87 students. The course content was concerned with data analytics, professional habits, engineering modeling and design, communication, and teaming. In addition, this course served students who had not chosen their engineering disciplinary areas yet. The grade breakdown of the course is generally 40% projects, 37% exams and quizzes, 10% assessments related to the team, 9% class preparation and participation, and 4% career exploration assignments. The university's first-year engineering cohort in Spring 2019 consisted of 26% female students and 74% male students. Among these students, 47% identified as White, 26% international, 11% Asian, 5% of two or more races, 5% Hispanic/Latino, 4% Black or African American, and 1% Other.

The course's final grade corresponded to a letter grade (A, B, C, D, or F), with a plus or minus indicating the student's achievement level. In 2019, these courses were hybrid (online and in-person), in which smaller portions of the course with approximately 25 students were able to go to the classroom in person once a week. Thus, during any given class period, some of the students could attend the lecture with the instructor in the classroom while others were using a streaming device. Because of the hybrid nature of the course and the number of students registered, students and instructors alike were accessing the LMS continuously.

In this study, we studied the difference between statistical analysis performed using LMS data with and without involving human operators with knowledge of the undergraduate first-year engineering course. By enlisting the contextual knowledge of the human operator, we were able to categorize students' clickstream behaviors with this knowledge.

Data Categorization

The LMS information contained 389 and 387 distinct objects for Section 1 and Section 2, respectively. These objects correspond to content that is by default in the platform before the course starts (e.g., welcome to the course, general use of the platform) and objects uploaded or created by instructional members (e.g., quizzes, lectures). Three researchers participated in this study, reviewed the objects, and discussed the possible categories until agreement was reached with eight broader categories and 38 granular categories to which the distinct objects could belong. One of the researchers had experience as a former student and teaching assistant in the class. After reaching an agreement on the categories, one of the researchers classified the objects into both broader and granular categories, leaving some of the objects for further discussion with the team. The remaining objects were classified through discussion until an agreement was reached. The definitions of the broader classification categories are found in Table 1.

Table 1.
Broader Classification Categories

Category	Definition	# Objects Section 1	# Objects Section 2
Career exploration	Refers to videos, plans of studies, and external links that were designed to help students explore the different engineering disciplines.	69	69
Assessment	Refers to any traditional grading object, such as quizzes, homework, and activities for extra credit for which the students receive a grade. This category was left out from the modeling, as it was directly related to the learning outcome.	48	41
Miscellaneous	Refers to links or content objects related to management (e.g., regarding requests) or general information (e.g., university policies, emergency evacuation).	99	83
Student Support	Refers to materials published for students to get advice on specific topics that can help them succeed in the course. This category includes materials related to effective learning practices, among others.	2	4
Materials	Refers to links, documents, or content objects that include information related to tools and topics not covered directly through the course, but that might be necessary to succeed (e.g., basic excel operations materials, data collection strategies, datasets, among others).	37	52

Lecture content	Refers to videos, PowerPoint slides, and other documents that cover topics directly related to the course's objectives and that are typically posted in sequential order. Lecture material does not include content related to career exploration or design projects. It was not possible to code this category at a more granular level.	43	45
Teaming	Refers to content on group assessment and assignments, as well as content on creating good team groups, such as strategies for teamwork.	24	24
Project	Refers to any object type related to projects. These include but are not limited to videos, PowerPoint slides, and other documents available for students to review.	67	72

Data analysis:

Before modeling, we summarized the data available to better understand the variables that we were measuring to evaluate the content of this first-year engineering course. Afterward, the analysis of the data was performed using a series of logistic models. The response variable corresponded to a binary classification of the learning outcome. Students whose final grade was A, A+, or A- were coded as one, and those whose final grade was lower than A- were coded as zero. The decision for the binary classification was made because an expert grader argued that, based on their experience in this course, there was a large difference between students that earned an A and students who had less than this letter grade. Further, there were few observations that had a final grade lower than A-.

A logistic linear model was used as the prediction tool due to two main reasons. First, a logistic type model serves well the objective of predicting students that can be at risk, while offering an approach in which it is still possible to understand which factor affects the model. Second, it was likely to have an unbalanced dataset and logistic regression is one of the most common techniques used in such cases [34]. In addition, logistic type models have appeared previously in LMS prediction of at-risk students' literature [25].

Two types of models were trained. The first type corresponded to models using summaries of the total accesses, such as the number of total accesses and the number of unique objects accessed. Three models were under this category, Model 1.1, Model 1.2, and Model 1.3. Model 1.1 corresponds to a model in which the only predictor for the achievement of an A was the number of clicks (NC) on the LMS. Model 1.2 corresponds to the model in which the only predictor was the number of unique objects visited (UNO). Finally, Model 1.3 is a model in which the explanatory variables correspond to NC and UNO. These first models reflect on the usual approach of summarizing the data with the information that comes from the platform without any human transformation. The second type corresponds to models that included the accesses discriminated by the broader categories (See Table I). For model 2.1, the NC to each broader category was included as an explanatory variable. In model 2.2, the UNO from each category was used as an explanatory variable. Model 2.3 corresponds to all variables from NC or UNO. From Model 2.3, a stepwise regression procedure [35] was used to select the variables that explained the largest variance in the grade. Model 2.4 was obtained using both backward and forward selection. Backward and forward selection are model selection strategies used frequently to select the more meaningful variables in models and therefore get the most parsimonious model [35]. In backward selection, all variables are considered at first, and the variables without relevance in the model are removed until all variables are meaningful to the model. Forward selection starts with the variable that contributes the most to the variance explanation of the model and add variables until no significant difference is obtained from including further variable [35]. After obtaining the final model, we ran verifications of linear assumptions [36]. For instance, the Variance Inflation Factor (VIF), was used to detect possible multicollinearity [37], as some of the variables were hypothesized as possibly correlated.

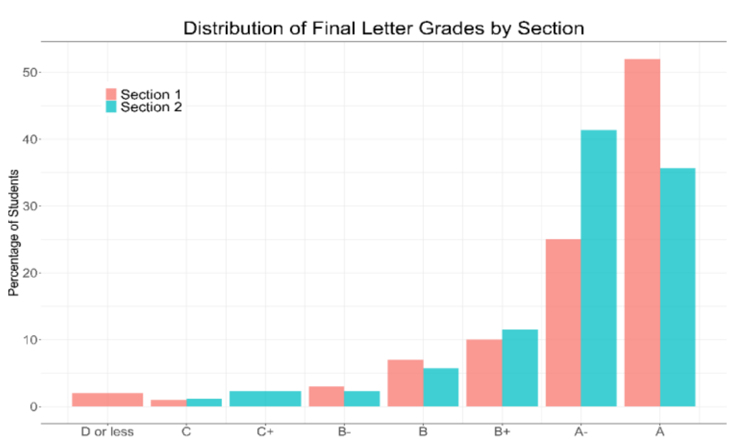
The evaluation of the performance of the different models was conducted by using Section 1 as the training dataset and Section 2 as the testing data set. Different measures were used to compare the models. The accuracy measures the probability for the model to be correctly classifying a student, regardless of whether the students is at risk or not. Precision measures the effectiveness of the model when predicting students that are not at risk. The recall corresponds to the ability of the model to detect students that are at risk. Finally, the F1 score corresponds to the harmonic mean of the precision and recall, summarizing the performance of the model when classifying both students at risk and students not at risk. For each model, the four measures were obtained to understand the usability and generalizability of the model.

Results

We had a total of 187 students participating in this study. Of these students, 77% had an A as their final grade in their course. The percentage of students getting an A was not statistically different across sections (see Figure 1). From Table 2 it is possible to see that the average number of clicks (NC) for Section 1 and Section 2 was close to the number of unique objects (389 and 387, respectively). However, for Section 1, the NC is higher than the number of objects available, while for Section 2, it is less than the number of objects available. In addition, students from Section 1 visited an average total of 220 unique objects (NUO), while students from Section 2 visited an average of 211 unique objects. Both sections show very dispersed distributions, with a standard deviation that reached almost a third of the mean value for the NC variable and a quarter of the mean value for NUO.

Figure 1

Distribution of Grades per section



Analyzing the information by category, we found that, on average, students clicked more objects than the number of objects available when it corresponded to categories such as assessment, student support, materials, teaming, and project. Also, students' number of clicks was less than the number of objects available for the categories career exploration, miscellaneous, and lecture content. When analyzing from the NUO, student support, assessment, and materials were the categories with the highest number of objects accessed. In contrast, career exploration and miscellaneous had less than 50% of the objects accessed.

Table 2

Summary statistics of objects clicked per student

	Section 1		**Section 2**	
	NC ***Mean (SD)***	**NUO** ***Mean (SD)***	**NC** ***Mean (SD)***	**NUO** ***Mean (SD)***
All objects available in the LMS	392.48 (109.72)	220.80 (50.53)	374.20 (118.59)	211.87 (52.20)
Career exploration	0.64 (0.30)	0.44 (0.22)	0.62 (0.30)	0.43 (0.23)
Assessment	1.33 (0.35)	0.74 (0.14)	1.32 (0.37)	0.75 (0.16)
Miscellaneous	0.67 (0.21)	0.42 (0.13)	0.63 (0.20)	0.40 (0.12)
Student Support	1.22 (0.65)	0.94 (0.39)	1.18 (0.57)	0.85 (0.24)
Materials	1.58 (0.55)	0.82 (0.15)	1.23 (0.45)	0.68 (0.15)
Lecture Content	0.97 (0.34)	0.59 (0.15)	0.96 (0.42)	0.52 (0.13)
Teaming	1.20 (0.53)	0.62 (0.20)	1.08 (0.54)	0.60 (0.24)
Project	1.29 (0.57)	0.61 (0.16)	1.24 (0.61)	0.60 (0.18)

*Number of unique objects (NUO) refers to the count of unique objects accessed by the student in the semester

**Number of clicks (NC) refers to the number of clicked objects during the semester

From Table 3, we can observe that the LMS accessing behavior does contain information that is relevant to predicting a student's learning outcomes. Regarding Model 1.1 and Model 1.3 it is possible to observe that the accuracy of both models is the same. However, Model 1.1 is more precise when predicting students' achievement an A in the course, while the F1 score favors Model 1.3 Both models seem to include relevant information regarding the prediction of student outcome.

Another relevant result from Table 3 corresponds to good accuracy achieved by even simple models. For instance, a logistic regression paired with summarized accessing behavior (Model 1.3), achieved 83.9% accuracy. Although the accuracy in Model 1.3 intuitively could be caused by correlation between NUO and NC, an analysis of the VIF overruled this possibility. The maximum VIF for Model 1.3 was 3.48

For Model 2.1, the VIF was also calculated, with a maximum of 3.31 obtained when evaluating the NUO on lecture content. Although none of the variables were significant in this model, in the linear relationship, NUO of students' support, miscellaneous, projects, and teaming showed a negative relationship with respect to students gaining an A in the class. While assessment, career exploration, lecture content, and materials are categories that showed a positive relationship.

For Model 2.2, VIF had a maximum value of 3.95 corresponding to NC on lecture content. This model resulted in two significant variables, NC career exploration and NC miscellaneous objects. When using just the number of clicks, most of the variables resulted in a positive relationship except for NC materials and NC miscellaneous.

Table 3
Model comparison

Model ***	**Model**	**Precision**	**Accuracy**	**Recall**	**F1**
Model 1.1	Number of unique objects (NUO*)	83.1%	81.6%	95.5%	88.9%
Model 1.2	Number of clicks (NC**)	81.5%	81.6%	98.5%	89.2%
Model 1.3	NUO+NC	83.5%	83.9%	98.5%	90.4%
Model 2.1	NUO by category	86.8%	80.5%	88.1%	87.4%
Model 2.2	NC by category	83.3%	82.8%	97.0%	89.7%
Model 2.3	NUO by category + NC by category	84.9%	81.6%	92.5%	88.6%
Model 2.4	NC career exploration + NUO materials	87.5%	85.1%	94.0%	90.6%
Model 2.5	NC career exploration		85.1%		

*Number of unique objects (NUO) refers to the count of unique objects accessed by the student in the semester
**Number of clicks (NC) refers to the number of clicked objects during the semester
*** The first type of model, models using summaries of the total accesses, such as the number of total accesses and the number of unique objects accessed are identified with a starting numbering of one. The second type of model, models using categorized summaries, are identified with a starting numbering of two.

For Model 2.3, the maximum VIF obtained was 12 corresponding to the NC of career exploration and the NUO of career exploration. This model failed the multicollinearity assumption. Nonetheless, the NUO of materials, and the NUO of students support result was significant with NUO of student support having a negative effect and NC materials having a positive effect on the final grade.

Model 2.4 is the result of using both backward and forward model. From both model selection strategies, the same model resulted in the best model. This is model 2.4. Model 2.4 has only two variables: NC career exploration and NUO materials; both significantly explain the final grade on the course for the students. Both significant variables have a positive relationship with the final grade. These results point to the importance of access to career exploration and other materials.

When comparing the different models, the models that contained human categorized data (e.g., model 2.4) achieved better precision, accuracy, and recall than those that used only summaries of the total access. Students' number of clicks in objects categorized as career exploration achieved an 83.9% accuracy on predicting whether a student would get an A or a lower grade in the course. This finding is relevant, as career exploration materials account for only 4% of the grade, and access to the career exploration materials occurred mainly at the beginning of the semester.

Discussion

The purpose of this study was to evaluate the content of a course of an FYE engineering program. Numerous evaluations of the FYE program have been made at the program level [9], [10], and the benefits of such programs have been proven. In this study, we offer insights into the content that is relevant for students' success. For instance, we found that although the career exploration assignments weighed very little in overall course grade, students' access to the career exploration materials was most predictive of student overall grade. The importance of career exploration might be underestimated by instructors without the results of this study since it might be considered unrelated to the subject matter of the course and takes up only 4% of the total grade. These results indicate the significance of career exploration content objects can be interpreted in two ways: in the light of internal and external motivation. In terms of internal motivation, more motivated students might look for more information about career exploration than those who only access objects in the course with greater weight in the course grade. On the other hand, in the light of external motivation, it is possible that students who explore these objects find more motivation to pursue a path in engineering, and because of that, they obtain a greater grade in the course. Either way, the categorized results give the instructional team insights on the importance of these objects and help outline some strategies to help students succeed. For example, giving students opportunities to reinforce the access to these objects, improving these materials, or early monitoring of students' access to these materials.

In addition, while there have been many prediction studies conducted in the context of large courses, this research compared two analytic approaches to determine which yielded the better prediction and returned the most insights into the classroom. The research presented differs from other predictive models at different levels. For instance, the use of the prediction is not only to identify students at risk (e.g., [25]) but also to provide insights into the relevant content. In addition, the use of a human-centered approach to transform the data is crucial to address the complexity related to content. We showed the differences in the prediction metrics and the interpretation of the models. While a model without categories would give us only a prediction, a categorized model can open room to understand what content has more relevance. For instance, the results highlighted the importance of the number of clicks that a student had on objects categorized as career exploration.

The supervision of learning processes for individuals and organizations is an essential goal for higher education institutions [38]. For instance, while institutions create room for a more diverse background of students and increase the number of students on their student board, monitoring learning processes for different groups of students becomes a necessity. This is of particular importance for FYE programs due to their complexity. In addition, the richness of LMS data can lead educators to find important metrics about student engagement, find students at risk of dropping the class and generate insights for instructors to make decisions and improve materials [39], [40].Due to its complexity, LMS data analysis might require human intervention to ensure data quality, particularly for categorizing the data [32].

The involvement of a human operator who has internal knowledge of the general structure, content, format, and design of the course yielded unexpected results regarding the significance of access to certain types of content objects and their influence on student success. As shown in our model comparisons, the findings in this study revealed what would most likely remain hidden without the involvement of human operators in the data transformation and analysis process. These findings supported previous literature on the importance of HFE [33] and emphasized the context-dependent nature of predictions in education [27]. Research findings that take the aforementioned factors into account have the potential to lead researchers to discover new patterns and relationships regarding student behaviors via LMS data and explore the underlying explanations for these behaviors, provide directions for future research, and offer instructors a new perspective to examine their courses for improvement to facilitate student learning.

Although it was necessary to employ a human operator for this study, the overarching argument of the study was to show the importance of LMS data and the need for a human centered approach when considering the evaluation of content in FYE programs. Therefore, a human operator might not be required if there is a coding schema that links the course content with the LMS objects. That is, to not only predict at-risk students but also evaluate the content of FYE courses, it is necessary to develop a naming scheme that gives enough information to use LMS data for FYE content evaluation. A general naming with the broader categories used in this article is possible; however, better naming conventions can be used if granular content categories are defined and followed when using an LMS.

Limitations

Some of the limitations of our work are related to the contextual dependence of the data. For instance, the data obtained, and the categories used are relevant for the course; however, they cannot be used for other FYE programs. In addition, even though the models used in this study are interpretable, more complex modeling should be considered in future studies to place more emphasis on the trade-off between interpretability and robustness of the methods. Lack of time data due to the storage of information from the LMS is another limitation. The LMS used did not provide access to the time in which the students clicked the objects.

Conclusion

FYE programs are a crucial part of engineering education . They are also complex environments. Evaluation of the content and strategies used in FYE is essential but, at times, also challenging to execute. LMS data represents an alternative for evaluating an FYE program's content. However, currently, LMS data are often only used for at-risk prediction. Despite yielding meaningful insight, such predictions often disregard the need to also provide insights into course decision making. In this case study, we present a human-operator approach that predicts students at risk while also returning insights into the classroom. We conclude that career exploration items contained the highest amount of explained variance when predicting students' final grade. The single use of students' access to career exploration content was able to predict student's final grades with an 84% precision in the class. The results generated with a human-categorized model give the instructional team insights on the importance of these objects and point to some strategies to help students succeed. For example, giving students opportunities to reinforce the access to these objects, improving these materials, or early monitoring the students' access to these materials.

Research to practice

This paper informs practice in FYE classrooms in the following manner. First, analyzing information on student-LMS interaction is critical; this information is often readily available and can be valuable for uncovering insights into the content and strategies used in FYE classrooms. Nonetheless, for the data from LMS to be useful, it is necessary to use a naming convention and do so with a person who is knowledgeable about the content and strategies used in the classroom. In addition, while there exist previous prediction studies in FYE classrooms, there is a need to go beyond prediction studies and think from a human-centered approach about *How* and *Why* users will use the information from these studies. Finally, results of this case study show career exploration materials to have an effect on the student's final grade. This would suggest that there can be space to investigate if this is the case for other settings, and if so, give more relevance to these types of materials in FYE courses.

Acknowledgments

This material is based upon work supported by the Charles Koch Foundation.

Bibliography

[1] J. Bordogna, E. Fromm, and E. W. Ernst, "Engineering Education: Innovation Through Integration," *Journal of Engineering Education*, vol. 82, no. 1, pp. 3–8, 1993, doi: 10.1002/j.2168-9830.1993.tb00065.x.

[2] E. Jamelske, "Measuring the impact of a university first-year experience program on student GPA and retention," *Higher Education*, vol. 57, no. 3, 2009, doi: 10.1007/s10734-008-9161-1.

[3] N. A. Pendergrass *et al.*, "Improving first-year engineering education," in *Journal of Engineering Education*, 2001, vol. 90, no. 1, pp. 33–41. doi: 10.1002/j.2168-9830.2001.tb00564.x.

[4] S. Odeh, S. McKenna, and H. Abu-Mulaweh, "A unified first-year engineering design-based learning course," *International Journal of Mechanical Engineering Education*, vol. 45, no. 1, 2017, doi: 10.1177/0306419016674508.

[5] S. A. Ambrose and C. H. Amon, "Systematic design of a first-year mechanical engineering course at Carnegie Mellon University," *Journal of Engineering Education*, vol. 86, no. 2, 1997, doi: 10.1002/j.2168-9830.1997.tb00281.x.

[6] N. S. Samsuri, K. Mohd-Yusof, and A. Abdul Aziz, "Preparing first year engineering students to become engineers: The impact of an 'introduction to engineering' course," *Journal of Technical Education and Training*, vol. 9, no. 1, 2017.

[7] R. Feldman and M. Zimbler, “Engendering College Student Success: Improving the First Year and Beyond,” *The McGraw-Hill Research Foundation Policy Paper, University of Massachusetts Amherst, September 2011*, 2011.

[8] J. Richardson and J. Dantzler, “Effect of a freshman engineering program on retention and academic performance,” in *Proceedings - Frontiers in Education Conference*, 2002, vol. 3. doi: 10.1109/fie.2002.1158656.

[9] D. Budny, G. Bjedov, and W. LeBold, “Assessment of the impact of the freshman engineering courses,” in *Proceedings - Frontiers in Education Conference*, 1997, vol. 2. doi: 10.1002/j.2168-9830.1998.tb00372.x.

[10] R. M. Marra, B. Palmer, and T. A. Litzinger, “The effects of a first-year engineering design course on student intellectual development as measured by the perry scheme,” in *Journal of Engineering Education*, 2000, vol. 89, no. 1. doi: 10.1002/j.2168-9830.2000.tb00492.x.

[11] P. L. Cox, E. Dunne Schmitt, P. E. Bobrowski, and G. Graham, “Enhancing the First-Year Experience for Business Students: Student Retention and Academic Success,” *Journal of Behavioral and Applied Management*, 2005, doi: 10.21818/001c.14572.

[12] K. C. Ellis, “Ready for college: Assessing the influence of student engagement on student academic motivation in a first-year experience program,” 2013.

[13] B. D. Jones, M. C. Paretti, S. F. Hein, and T. W. Knott, “An analysis of motivation constructs with first-year engineering students: Relationships among expectancies, values, achievement, and career plans,” *Journal of Engineering Education*, vol. 99, no. 4, 2010, doi: 10.1002/j.2168-9830.2010.tb01066.x.

[14] J. E. Froyd and G. J. Rogers, “Evolution and evaluation of an integrated, first-year curriculum,” in *Proceedings - Frontiers in Education Conference*, 1997, vol. 2. doi: 10.1109/fie.1997.636048.

[15] A. Katz, K. Reid, D. Riley, and N. van Tyne, “Overcoming Challenges to Enhance a First Year Engineering Ethics Curriculum,” *Advances in Engineering Education*, 2020.

[16] H. A. Diefes-Dux and L. M. Cruz Castro, “Reflection types and students’ viewing of feedback in a first-year engineering course using standards-based grading,” *Journal of Engineering Education*, vol. 111, no. 2, pp. 283–307, Apr. 2022, doi: 10.1002/jee.20452.

[17] M. T. Gordon, S. Chun, X. S. Zhao, M. J. C. Nalbandian, M. K. Han, and M. Oyanader, “Design Course for First-Year Students in Multiple Engineering Disciplines,” 2019.

[18] C. A. Twigg, “Improving Learning and Reducing Costs for Online Learning,” in *Encyclopedia of Distance Learning, Second Edition*, 2011. doi: 10.4018/978-1-60566-198-8.ch163.

[19] G. Marbach-Ad and P. G. Sokolove, “The use of E-mail and in-class writing to facilitate student-instructor interaction in large-enrollment traditional and active learning classes,” *Journal of Science Education and Technology*, vol. 11, no. 2, 2002, doi: 10.1023/A:1014609328479.

[20] F. Martin, D. Polly, S. Coles, and C. Wang, "Examining Higher Education Faculty Use of Current Digital Technologies: Importance, Competence, and Motivation," *International Journal of Teaching and Learning in Higher Education*, vol. 32, no. 1, 2020.

[21] S. J. Miah, M. Miah, and J. Shen, "Editorial note: Learning management systems and big data technologies for higher education," *Education and Information Technologies*, vol. 25, no. 2. 2020. doi: 10.1007/s10639-020-10129-z.

[22] A. E. Elhabashy, "Pedagogy in large classrooms best practices," 2019.

[23] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," *Computers and Education*, vol. 103, pp. 1–15, Dec. 2016, doi: 10.1016/j.compedu.2016.09.005.

[24] S. B. Shum, R. Ferguson, and R. Martinez-maldonado, "Human-Centred Learning Analytics," *Journal of Learning Analytics*, vol. 6, no. 2, pp. 1–9, 2019.

[25] F. Marbouti, "A Standards-Based Grading Model to Predict Students' Success in a First-Year Engineering Course.," Purdue University, 2016.

[26] Y. Kuo, A. E. Walker, B. R. Belland, and K. E. E. Schroder, "A predictive study of student satisfaction in online education programs | Kuo | The International Review of Research in Open and Distributed Learning," *The International Review of Research in Open and Distance Learning*, vol. 14, no. 1, pp. 15–39, 2013, [Online]. Available: http://www.irrodl.org/index.php/irrodl/article/view/1338/2416

[27] M. Cocea and S. Weibelzahl, "Cross-system validation of engagement prediction from log files," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4753 LNCS, pp. 14–25, 2007, doi: 10.1007/978-3-540-75195-3_2.

[28] Y. C. Yeh, O. M. Kwok, H. Y. Chien, N. W. Sweany, E. Baek, and W. A. McIntosh, "How college students' achievement goal orientations predict their expected online learning outcome: The mediation roles of self-regulated learning strategies and supportive online learning behaviors," *Online Learning Journal*, vol. 23, no. 4, pp. 23–41, 2019, doi: 10.24059/olj.v23i4.2076.

[29] W. R. Watson and S. L. Watson, "An argument for clarity: What are learning management systems, what are they not, and what should they become?," *TechTrends*, vol. 51, no. 2, pp. 28–34, 2007, doi: 10.1007/s11528-007-0023-y.

[30] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," *Computers and Education*, vol. 103, pp. 1–15, 2016, doi: 10.1016/j.compedu.2016.09.005.

[31] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," *Computers and Education*, vol. 103, pp. 1–15, 2016, doi: 10.1016/j.compedu.2016.09.005.

[32] A. Jain *et al.*, "Overview and Importance of Data Quality for Machine Learning Tasks," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Aug. 2020, pp. 3561–3562. doi: 10.1145/3394486.3406477.

[33] D. Hannon *et al.*, "A Human Factors Engineering Education Perspective on Data Science, Machine Learning and Automation," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 63, no. 1, pp. 488–492, 2019, doi: 10.1177/1071181319631248.

[34] H. Zhang, Z. Li, H. Shahriar, L. Tao, P. Bhattacharya, and Y. Qian, "Improving prediction accuracy for logistic regression on imbalanced datasets," in *Proceedings - International Computer Software and Applications Conference*, Jul. 2019, vol. 1, pp. 918–919. doi: 10.1109/COMPSAC.2019.00140.

[35] G. A. Seber and A. J. Lee, "Prediction and Model Selection," in *Linear Regression Analysis*, Second., Hoboken, NJ: John Wiley & Sons, 2003, pp. 391–456.

[36] G. A. Seber and A. J. Lee, "Departures from Underlying Assumptions," in *Linear Regression Analysis, Second Edition*, Second., Hoboken, NJ: Wiley, 2003, p. 227263.

[37] G. R. Franke, "Multicollinearity," in *Wiley International Encyclopedia of Marketing*, Chichester, UK: John Wiley & Sons, Ltd, 2010. doi: 10.1002/9781444316568.wiem02066.

[38] W. R. Watson and S. L. Watson, "An argument for clarity: What are learning management systems, what are they not, and what should they become?," *TechTrends*, vol. 51, no. 2, pp. 28–34, 2007, doi: 10.1007/s11528-007-0023-y.

[39] Y. Kuo, A. E. Walker, B. R. Belland, and K. E. E. Schroder, "A predictive study of student satisfaction in online education programs | Kuo | The International Review of Research in Open and Distributed Learning," *The International Review of Research in Open and Distance Learning*, vol. 14, no. 1, pp. 15–39, 2013, [Online]. Available: http://www.irrodl.org/index.php/irrodl/article/view/1338/2416

[40] M. Cocea and S. Weibelzahl, "Cross-system validation of engagement prediction from log files," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4753 LNCS, pp. 14–25, 2007, doi: 10.1007/978-3-540-75195-3_2.