PLOS ONE

RESEARCH ARTICLE

An AI-based intervention for improving undergraduate STEM learning

Mohammad Rashedul Hasan[1]*, **Bilal Khan**[2]

1 Electrical and Computer Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States of America, **2** Depts. of Community & Population Health, and Computer Science & Engineering, Lehigh University, Bethlehem, PA, United States of America

* hasan@unl.edu

Check for updates

OPEN ACCESS

Citation: Hasan MR, Khan B (2023) An AI-based intervention for improving undergraduate STEM learning. PLoS ONE 18(7): e0288844. https://doi.org/10.1371/journal.pone.0288844

Editor: Paulo F. Carvalho, Carnegie Mellon University, UNITED STATES

Received: March 9, 2022

Accepted: July 5, 2023

Published: July 19, 2023

Copyright: © 2023 Hasan, Khan. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The file named "2019-Fall-CSCE235-All.csv" is available from the openICPSR database. URL: https://www.openicpsr.org/openicpsr/project/184642/version/V1/view.

Funding: Our project was supported in part by a grant from the U.S. National Science Foundation (NSF DUE 2142558). The second author received support from the U.S. National Institutes of Health (NIH NIGMS P20GM130461 and NIH NIAAA R21AA029231), which facilitated the longitudinal analyses in this work. There was no additional external funding received for this study.

Abstract

We present results from a small-scale randomized controlled trial that evaluates the impact of just-in-time interventions on the academic outcomes of N = 65 undergraduate students in a STEM course. Intervention messaging content was based on machine learning forecasting models of data collected from 537 students in the same course over the preceding 3 years. Trial results show that the intervention produced a statistically significant increase in the proportion of students that achieved a passing grade. The outcomes point to the potential and promise of just-in-time interventions for STEM learning and the need for larger fully-powered randomized controlled trials.

Introduction

Even as the number of jobs requiring science, technology, engineering, and mathematics (STEM) knowledge and skills rapidly increase [1], retention rates in post-secondary STEM fields have fallen below 50% [2]. The graduation rate of STEM students is roughly 20% below their counterparts in non-STEM majors [3]. This is attributed to students' poor academic performance, particularly in the first few years of college [4]. Improving student performance in STEM programs is thus a critical national need. Many large-scale systemic reforms might help, including department-wide implementation of evidence-based practices, faculty development, and leadership, adoption of successful curricular approaches, stronger teacher preparation programs, student learning communities, professional development of faculty, etc. [2, 5, 6]. Yet, the rate at which systemic changes can take place is limited by each institution's financial resources and inertia [7], so it is important to develop alternative, cost-effective, incremental, and contextually appropriate interventions that might improve STEM performance. We report on one such attempt here.

Education researchers approach this need for intervention from many angles, e.g., designing active learning strategies to improve learning in the classroom [8], developing light-touch interventions to improve learning outside the classroom [9], and forming STEM learning communities to address the social aspects of learning [10, 11], etc. Other approaches engage psychological interventions such as online sessions conveying a growth mindset [12], performance-related warnings through course management system-based alerts [13], and emails containing grade forecast [7] to improve academic achievement. Such approaches implicitly

Competing interests: The authors have declared that no competing interests exist.

leverage non-cognitive psychological drivers (e.g., motivation) of academic performance [6], drawing support from Social Cognitive Theory [14].

Most recently, Artificial Intelligence (AI)-based strategies have been proposed for the delivery of psychological interventions [13, 15]. These approaches involve the creation of predictive models that use students' recent performance data (e.g., academic scores at the beginning of the semester) to forecast their final course outcomes. Such forecast messages are seen as interventions that serve to both inform students and motivate them to improve their academic performance [7, 13, 16]. The customizability and low implementation costs of AI-based solutions make them a potentially cost-effective, scalable approach for improving academic achievement, particularly in courses during the first two years of college, where STEM curriculum is fairly standardized, and performance is critical to long-term student retention [13, 16, 17].

This article presents an ML-based method for providing just-in-time interventions among undergraduate STEM students and evaluates its efficacy via a small randomized control trial.

Materials and methods

Study cohort

The study cohort consisted of 65 students who enrolled in the introductory course on discrete structures in the Fall of 2019 at the University of Nebraska-Lincoln (UNL). The course had prerequisites of introductory programming and precalculus-level mathematics. All 65 of the students were first-year students who were enrolled in the study by signing a written informed consent form. The study was approved by the UNL's Institutional Review Board (IRB #: 20180118001EX).

All students were informed at the beginning of class that they would be receiving 3 messages from "an AI" at a regular interval and that these 3 messages would contain "a forecast of your future performance". More specifically, the AI would determine if their prospects were "Good", "Fair", "Prone-to-Risk", "At-Risk", but in some cases, the AI might declare that it was "Unable to make a prediction". Students were told that the messages they received would correspond to the final grades the AI had predicted for them, in accordance with Table 1.

The AI was instrumented by an ML app accessible through computer/cellphone browsers. Students were instructed to use their course management system (i.e., Canvas) credentials to log in to the app. When new forecasting was computed by the AI, the app notified the students by sending an automatic message to their email accounts. The message contained their forecasted performance. Students could also check their forecasting messages by logging in to the app. The app does not provide any other information apart from three periodic forecasting messages.

Randomized assignment

Just prior to week 6, the cohort was split into two groups. A randomly chosen one-half (33) of the students were assigned to the control group, and the remaining (32) were assigned to the

Table 1. Mapping of the AI's forecast to the AI's message.

AI's forecast of final grade	Message from AI
$90\% \leq$ grade $\leq 100\%$	"You are Good"
$80\% \leq$ grade $< 90\%$	"You are OK"
$70\% \leq$ grade $< 80\%$	"You are Prone-to-Risk"
$0\% \leq$ grade $< 70\%$	"You are At-Risk"

https://doi.org/10.1371/journal.pone.0288844.t001

intervention group. Students were not informed about the fact that a randomized assignment had taken place, or which group they were placed in.

All students (regardless of group) received 3 messages over the course of the semester, at the end of weeks $t = 6, 9$, and 12. The message at week 6 preceded the course midterm exam by 7 days; the message at week 12 preceded the final exam by 7 days; the message at week 9 was transmitted 1/3 of the way through the time interval between the midterm and final exams. These 3 timepoints were chosen in advance by contemplating natural pivots in the course delivery.

Intervention group

Within the intervention group, each student received a message at the end of week $t = 6, 9$, and 12. Each message indicated whether the AI had determined the student's prospects to be "Good", "Fair", "Prone-to-Risk", or "At-Risk". To generate the message for each student, data on the student's formal assessments (to date) were fed as input into an appropriate previously trained predictive ML model. The model's prediction was then translated into a message in accordance with Table 1. The AI-generated messages were delivered to each student via the app.

Control group

Within the control group, each student received a message at the end of week $t = 6, 9$, and 12. Each message stated that the AI had been "Unable to make a prediction" concerning the student's prospects in the course. These messages were delivered to each student via the app.

Predictive machine learning models

The predictive models described in this section were developed by the authors in previous work [15], and only a brief summary description is provided here. Three distinct predictive machine learning models were developed, $\mathcal{M}_6$, $\mathcal{M}_9$, and $\mathcal{M}_{12}$, to be used in determining intervention message content at weeks 6, 9, and 12, respectively.

The **training data set** consisted of academic assessments collected from 537 students who were enrolled in the same course between Fall 2015 and Spring 2018. The number of cases in the training data set was thus 537. The dimensions of the data set were 17, consisting of 17 numerical predictors (numerical scores on homework assignments, quizzes, and exams), along with each student's numerical final grade. The numerical grade was replaced with a categorical label of "A", "B", "C" or "Below C" using the binning scheme described in Table 2.

Before building the model, **feature selection** was carried out by retaining only those predictors whose Pearson correlation with the final grade exceeded 0.45. The resulting features selected are presented in Table 3.

Table 2. Binning of numerical final grade.

Numerical final grade	Label
$90\% \leq \text{grade} \leq 100\%$	"A"
$80\% \leq \text{grade} < 90\%$	"B"
$70\% \leq \text{grade} < 80\%$	"C"
$0\% \leq \text{grade} < 70\%$	"Below C"

https://doi.org/10.1371/journal.pone.0288844.t002

Table 3. Outputs of the feature selection process.

Model	Data Considered	Features selected
$\mathcal{M}_6$	Weeks 1–6	Quiz 1–3 & Homework 1, 2
$\mathcal{M}_9$	Weeks 1–9	Quiz 1–5 & Homework 1, 2, 3 & Midterm 1
$\mathcal{M}_{12}$	Weeks 1–12	Quiz 1–7 & Homework 1, 2, 3, 4 & Midterm 1

https://doi.org/10.1371/journal.pone.0288844.t003

Rather than training a single classifier to predict all 4 class labels ("A", "B", "C", and "Below C") in one shot, we followed a hybrid approach [18]. Building each model $\mathcal{M}_t$ (t = 6, 9, 12) required the training of two classifiers:

$\mathcal{C}_t^1$ is a 3-label classifier that predicts either "A", "B" or "C and Below". It is trained on a transformed data set where the students who were labeled "C" or "Below C" have been relabelled as "C and Below". The distribution of the labels for this classifier is: A = 252, B = 156, and C and Below = 129.

$\mathcal{C}_t^2$ is a binary classifier that predicts either "Below C" or "not Below C". It is trained on a transformed data set where the students who were labeled "A", "B", or "C" have been relabelled as "not Below C". The distribution of the labels for this classifier is: not Below C = 396 and below C = 141.

The predictions of the two classifiers $\mathcal{C}_t^1$ and $\mathcal{C}_t^2$ are combined to create the predictive model $\mathcal{M}_t$ as follows:

- If $\mathcal{C}_t^1$ predicts "A" or "B" then this output is taken to be the output of $\mathcal{M}_t$.
- If $\mathcal{C}_t^1$ predicts "C and Below" then $\mathcal{C}_t^2$ is consulted:
  - If $\mathcal{C}_t^2$ predicts "Below C" then that is taken to be the output of $\mathcal{M}_t$.
  - If $\mathcal{C}_t^2$ predicts "not Below C" then the output of $\mathcal{M}_t$ is "C".

This 2-stage design was chosen to address challenges associated with a lack of data and features, especially in early predictions (e.g., the $\mathcal{M}_6$ model). Performance measures for the $\mathcal{M}_6$, $\mathcal{M}_9$, and $\mathcal{M}_{12}$ models were described by the authors in their prior work [15], but are summarized in Table 5 in S1 Appendix.

Measures of impact on student outcomes

We are interested in assessing the impact of the intervention on student performance outcomes. Specifically, we would like to know whether the intervention significantly increased the proportion of students who passed the course (i.e., achieved a grade > 64). To answer this question, we considered the 2x2 contingency table containing counts, where the groups (columns) were taken as Intervention and Control, while the outcomes (rows) were taken to be Failed ($<= 64$) or Passed (> 64). By design, the total number of individuals in each of the intervention versus control groups was fixed, so the columns of the contingency table were conditioned, while the row sums were not. Our assignment of individuals to intervention versus control was at random; this, together with our assumption that social network effects were minimal, allowed us to conclude that responses aggregated in the table were indeed independent. Given these data characteristics, together with the small sample size, we chose to apply Barnard's test [19]– an unconditional exact test for two independent binomials. Let p_I denote the binomial probability of "Failing" for members of the intervention group and p_C the

analogous probability for members of the control group. Our null hypothesis is then H_0: $p_I \geq p_C$ and alternative hypothesis H_1: $p_I < p_C$. The hypothesis was tested at the 5% significance level via a one-sided Barnard's test engaging the Wald statistic with unpooled variance, by using the implementation in the `scipy` library [20] (see "Supporting information" section for code). The findings from this analysis are presented in the Results section below.

Results

Prior to the first intervention at week 6, we did not observe a significant difference in the distribution of various performance types among the control and treatment groups.

Impact on student outcomes

Table 4 shows the `2x2` contingency table derived from the data of our randomized control trial.

Barnard's test yields 1.92 as the value of the Wald statistic, with a p-value of 0.0352. It follows that under the null hypothesis (that the intervention does not lower the chance of a student failing), the probability of obtaining trial results at least as extreme as the observed data is approximately 3.5%. Since this p-value is less than our chosen significance level of 5%, we have sufficient evidence to reject the null hypothesis in favor of the alternative.

Additionally, Relative Risk (RR) was used to assess the substantive significance within the Randomized Controlled Trial (RCT). The RR was found to be 0.34. Given that the RR value was <1, the intervention was confirmed to reduce the number of students below the threshold relative to the control group.

Acceptability of the intervention

We conducted a general user survey on the preliminary version of the AI-based app at the end of the Fall 2019 semester. The survey was voluntary and there was no incentive provided for completing it (e.g., no financial incentives or extra credit points). A total of 40 (of the 65) students completed the survey, of which 8 were female and 32 were male. Of the 40 respondents, 24 students belong to the intervention group, while 16 were in the control group. The survey gathered data on students' experience and demographics through a series of multiple-choice questions, the results of which are presented below:

Question Q1 read *Did you receive the message "Unable to make a prediction" during the semester?* The question was used to identify membership in the intervention arm since only the students in the control arm received this message.

Question Q2 read *How many times have you used the performance forecasting app during the semester?* This question was used to identify "active users" within the intervention arm. Survey data showed that 14 students (58% of the 24 who received the intervention) actively used the app ("Few times a month" = 7, "Every 2/3 weeks" = 5, and "Almost every week after the prediction was made" = 2). The other 10 students "never" used the app. Further analysis showed variability in usage by gender: Of the 24 students in the intervention arm, 6 were

Table 4. Contingency table.

	Intervention	Control	Totals
Failed	3	9	12
Passed	29	24	53
Totals	32	33	65

https://doi.org/10.1371/journal.pone.0288844.t004

female and 18 were male. However, among the females, only 17% (1 of 6) were active users (5 female students "never" used the app, and the remaining 1 female student used the app "Few times a month"), while among the males 78% (14 of 18) were active users.

Question Q3 read How useful were the predictions? There were 4 possible answers: "Very useful", "Somewhat useful", "Not at all useful", and "I have never looked at the predictions". Data from this question revealed that 12 of the 14 active users (86%) reported finding the app's predictions to be useful ("Very useful" = 5, "Somewhat useful" = 7, and "Not at all useful" = 2).

Question Q4 read Did you put more effort into your studies after seeing the predictions? There were 4 possible answers: "Yes", "Not always", "No", and "I have never looked at the predictions". The question was used to determine that 12 of the 14 active users (86%) reported they believed they put more effort in after seeing the predictions ("Yes" = 6, "Not always" = 6, and "No" = 2).

Discussion

Though we observed an overall positive impact of the AI-based interventions on students' academic outcomes, we identified some limitations in the design of the interventions, analysis of the study results, and acceptability survey of the interventions.

Unintended intervention to the control group. Our app attempted to mask the randomized assignment into the control/intervention arms by sending the control group students messages that read "Unable to make a prediction" whenever members of the intervention group received a forecast-based prompt. These "deceptive" messages might have influenced the outcome of some students in the control group and impacted their course-related behaviors (compared to if they had received no messages at all). Unfortunately, in the present study, we did not collect any information about the psychological impact of these deceptive messages on students in the control group. We acknowledge that there are potential ethical concerns surrounding deception in our experiments, given the unknown impacts. Our future studies will attempt to circumvent the need for deceptive messages altogether or will, at minimum, seek to measure the impact of deceptive messages to ascertain that no harm is incurred.

Limited to intent-to-treat analysis. Our app did not capture metadata on user engagement with the app, so we were not able to evaluate treatment compliance except via self-report at the end-of-study survey. Unfortunately, in hindsight, the granularity of this usage/acceptability survey was too coarse to enable refined analyses. For example, in question 2, we intended to identify "active users" in the intervention arm based on whether they reported using the app "At least a few times in a month". However, since the app sent only three messages over the semester, this question may have yielded an undercount of active users. We plan to alleviate these limitations in our future study by sending intervention messages more frequently and collecting metadata about participant viewing behaviors to support more nuanced as-treated analysis.

Limitations to generalizability. Our system is based on the premise that a model trained on previous offerings of a course can subsequently be used to deliver interventions to students in future offerings of the same course. This requires a clear correspondence between assessment items (i.e., homework, quizzes, exams) across semesters. In our RCT, the same instructor taught the same curriculum across all semesters from which the model was trained, and the intervention deployed. More broadly, if each assessment's topics remain the same, but the questions change, we believe it may be possible to limit the impact of cross-semester heterogeneity through data preprocessing and normalization. If, however, the topics are fundamentally altered over time, it will be challenging to train and apply a model. An ML-based intervention system that relies on academic assessment data alone may not be effective if the assessment

topics vary significantly over time. It would be interesting to investigate if the inclusion of non-academic data (e.g., course-related behavior, science identity, socioeconomic background, social connectedness, etc.) makes the interventions more robust to variability in the assessment instruments over time. In the future, we would like to explore such questions by acquiring non-academic features and using them alongside assessment data in the next version of our ML system.

Limitations due to scale. Our research findings are limited by statistical power implications of the small cohort size (65 students), short duration (one semester), predictor granularity (17 timepoints), and intervention frequency (3 timepoints). A future RCT that is larger on any/all of these axes will allow us to test richer hypotheses, such as whether it is possible to cluster students based on their distal characteristics and proximal trajectories, towards the design of tailored interventions.

Conclusions

We conducted a small-scale randomized controlled trial to evaluate the effectiveness of AI-based interventions in improving undergraduate STEM education. The study shows the potential for leveraging AI to build this type of intervention system:

- We found that the messaging intervention increased the proportion of students who passed (i.e., achieved a final grade > 64) and that this increase was statistically significant ($p = 0.0352$).
- Exit surveys from the RCT found that a significant fraction ($\sim 58\%$) of students had used the app frequently, and of those, a significant fraction ($\sim 86\%$) felt that messages were helpful and that they worked harder after receiving them. Male students were ~ 4 times more likely than females to use the app (72% versus 17%).

Future work

In the short term, we plan to address the limitations presented in the Discussion section by collecting data from larger, more carefully designed RCT. In addition to this, we plan to perform robustness testing. Longer term, we will pursue the promises of AI to tailor interventions by reflecting the natural heterogeneity among students. We will approach this by computing a typology of students based on their responsiveness to messaging interventions drawn from a range of design parameters. While other researchers have built classification schemes based on students' assessment scores and cognitive factors [21–23], few have considered students' relative responsiveness to interventions as the basis of a typology. Given such a typology, an AI system that can rapidly infer each student's type might be able to generate intervention messaging with tone and content optimized to yield favorable outcomes. This line of inquiry would be a natural extension of prior efforts in machine learning that have sought to automate the efficient determination and updating of student classifications [18, 24–26], and allow for tailored interventions throughout the semester.

Supporting information

S1 File.
(PDF)

S1 Data.
(CSV)

S1 Appendix.
(PDF)

Author Contributions

Conceptualization: Mohammad Rashedul Hasan, Bilal Khan.

Data curation: Mohammad Rashedul Hasan.

Formal analysis: Mohammad Rashedul Hasan, Bilal Khan.

Investigation: Mohammad Rashedul Hasan.

Methodology: Mohammad Rashedul Hasan, Bilal Khan.

Project administration: Mohammad Rashedul Hasan.

Resources: Mohammad Rashedul Hasan.

Software: Mohammad Rashedul Hasan.

Supervision: Mohammad Rashedul Hasan, Bilal Khan.

Validation: Mohammad Rashedul Hasan, Bilal Khan.

Visualization: Mohammad Rashedul Hasan.

Writing – original draft: Mohammad Rashedul Hasan.

Writing – review & editing: Bilal Khan.

References

1. Bureau of Labor Statistics. https://www.bls.gov/emp/tables/stem-employment.htm.
2. Sithole A, Chiyaka ET, McCarthy P, Mupinga DM, Bucklein BK, Kibirige J. Student Attraction, Persistence and Retention in STEM Programs: Successes and Continuing Challenges. Higher Education Studies. 2017; 7(1):46–59. https://doi.org/10.5539/hes.v7n1p46
3. Leary M, Morewood A, Bryner R. A controlled intervention to improve freshman retention in a STEM-based physiology major. Advances in Physiology Education. 2020; 44(3):334–343. https://doi.org/10.1152/advan.00038.2020 PMID: 32568008
4. Chen X. STEM attrition among high-performing college students: Scope and potential causes. Journal of Technology and Science Education. 2015; 5(1):1–19. https://doi.org/10.3926/jotse.136
5. Fry CL. Achieving Systemic Change: A Sourcebook for Advancing and Funding Undergraduate STEM Education, The Coalition for Reform of Undergraduate STEM Education. The Association of American Universities. 2013; p. 1–30.
6. Cromley JG, Perez T, Kaplan A. Undergraduate STEM Achievement and Retention: Cognitive, Motivational, and Institutional Factors and Solutions. Institute of Education Sciences (ED); National Science Foundation (NSF). 2016; 3(1):4–11.
7. Nostrand DFV, Pollenz RS. Evaluating Psychosocial Mechanisms Underlying STEM Persistence in Undergraduates: Evidence of Impact from a Six-Day Pre-College Engagement STEM Academy Program. CBE Life Sci Educ. 2016; 16(2).
8. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, et al. Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences. 2014; 111(23):8410–8415. https://doi.org/10.1073/pnas.1319030111 PMID: 24821756
9. Rodriguez F, Rivas MJ, Matsumura LH, Warschauer M, Sato BK. How do students study in STEM courses? Findings from a light-touch intervention and its relevance for underrepresented students. PLOS ONE. 2018; 13(7):1–20. https://doi.org/10.1371/journal.pone.0200767 PMID: 30063744
10. Solanki S, McPartlan P, Xu D, Sato BK. Success with EASE: Who benefits from a STEM learning community? PLOS ONE. 2019; 14(3):1–20. https://doi.org/10.1371/journal.pone.0213827 PMID: 30901339
11. Cohen GL, Garcia J, Apfel N, Master A. Reducing the Racial Achievement Gap: A Social-Psychological Intervention. Science. 2006; 313(5791):1307–1310. https://doi.org/10.1126/science.1128317 PMID: 16946074

12. Paunesku D, Walton GM, Romero C, Smith EN, Yeager DS, Dweck CS. Mind-Set Interventions Are a Scalable Treatment for Academic Underachievement. Psychological Science. 2015; 26(6):784–793. https://doi.org/10.1177/0956797615571017 PMID: 25862544

13. Arnold KE, Pistilli MD. Course Signals at Purdue: Using Learning Analytics to Increase Student Success. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. 2012; p. 267–270.

14. Bandura A. Social cognitive theory of mass communication. Media Psychology. 2001; 3:265–299. https://doi.org/10.1207/S1532785XMEP0303_03

15. Hasan MR, Aly M. Get More From Less: A Hybrid Machine Learning Framework for Improving Early Predictions in STEM Education. In: The 6th Annual Conf. on Computational Science and Computational Intelligence, CSCI 2019 (CSCI'19); 2019.

16. Page LC, Gehlbach H. How an Artificially Intelligent Virtual Assistant Helps Students Navigate the Road to College. AERA Open. 2017; 3(4):2332858417749220. https://doi.org/10.1177/2332858417749220

17. Chen Y, Johri A, Rangwala H. Running out of STEM: a comparative study across STEM majors of college students at-risk of dropping out early. LAK'18: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. 2018; p. 270–279.

18. Aly M, Hasan MR. Improving STEM Performance by Leveraging Machine Learning Models. In: the Proceedings of the International Conference International Conference of Frontiers in Education (FECS'19); 2019. p. 205–2011. Available from: https://csce.ucmss.com/cr/books/2019/LFS/CSREA2019/FEC7082.pdf.

19. Barnard G. A New Test for 2 × 2 Tables. Nature. 1945; 156(177).

20. Virtanen P, Gommers R, Oliphant TEea. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods; 17:261. https://doi.org/10.1038/s41592-019-0686-2 PMID: 32015543

21. Zhang G, Anderson TJ, Ohland MW, Thorndyke BR. Identifying Factors Influencing Engineering Student Graduation: A Longitudinal and Cross-Institutional Study. Journal of Engineering Education. 2004; 93(4):313–320. https://doi.org/10.1002/j.2168-9830.2004.tb00820.x

22. Jones BD, Paretti MC, Hein SF, Knott TW. An Analysis of Motivation Constructs with First-Year Engineering Students: Relationships Among Expectancies, Values, Achievement, and Career Plans. Journal of Engineering Education. 2010; 99(4):319–336. https://doi.org/10.1002/j.2168-9830.2010.tb01066.x

23. Van Soom C, Donche V. Profiling First-Year Students in STEM Programs Based on Autonomous Motivation and Academic Self-Concept and Relationship with Academic Achievement. PLOS ONE. 2014; 9(11):1–13. https://doi.org/10.1371/journal.pone.0112489 PMID: 25390942

24. Marbouti F, Diefes-Dux HA, Madhavan K. Models for Early Prediction of At-risk Students in a Course Using Standards-based Grading. Comput Educ. 2016; 103(C):1–15. https://doi.org/10.1016/j.compedu.2016.09.005

25. Essa A, Ayad H. Student Success System: Risk Analytics and Data Visualization Using Ensembles of Predictive Models. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. 2012; p. 158–161.

26. Macfadyen LP, Dawson S. Mining LMS data to develop an early warning system for educators: A proof of concept. Computers and Education. 2010; 54(2):588–599. https://doi.org/10.1016/j.compedu.2009.09.008