Seiji Isotani · Eva Millán ·
Amy Ogan · Peter Hastings ·
Bruce McLaren · Rose Luckin (Eds.)

LNAI 11625

Artificial Intelligence in Education

20th International Conference, AIED 2019
Chicago, IL, USA, June 25–29, 2019
Proceedings, Part I

Part I

AIED 2019

Springer

Integrating Students' Behavioral Signals and Academic Profiles in Early Warning System

SungJin Nam(✉) and Perry Samson

University of Michigan, Ann Arbor, USA
{sjnam,samson}@umich.edu

Abstract. In this paper, we investigated how students' behavioral signals and incoming profiles can be integrated to describe and predict student success in a higher education's STEM course. The results include three major findings. First, we found behavioral signals like the number of correct responses to in-class questions, the number of confusing slides, and the number of viewed slides and videos are stable predictors of student success across different periods of a semester. Second, from the mixed-effect modeling results, we could identify significant gender gaps between mid-level incoming GPA student groups. We also showed some possible course advising scenarios based on the interaction between student behaviors and incoming profile factors. Third, using both behavioral signals and incoming profiles, our weekly forecast model on student success achieved a 72% prediction accuracy. We believe these findings can set the stage for subsequent early warning system studies that use different types of student data. Further investigations on the causal relationships for suggested results and developing other novel predictive features for student success would be beneficial for designing a better early warning system.

Keywords: Early warning system · Behavioral logs · Incoming profile · Grade prediction · Mixed effect model

1 Introduction

Several Early Warning Systems (EWSs) have been developed that harness the predictive power of Learning Management System (LMS) and Student Information System (SIS) data to identify at-risk students and allow for more timely pedagogical interventions [1,4,15,16]. For example, data on student online activity in a web-based LMS may provide an early indicator of students' academic performance [28]. Other studies also found that there is a strong relationship between LMS usage patterns and student exam scores [7,10].

Recent studies suggest that EWSs can benefit from using relevant signals from LMS data, such as access to course resources [25], the usage patterns of

S. Isotani et al. (Eds.): AIED 2019, LNAI 11625, pp. 345–357, 2019.
https://doi.org/10.1007/978-3-030-23204-7_29

a digital coaching application [6], and students' incoming prior academic performance [6,25]. Similarly, studies like [8,11] showed applications for identifying at-risk students for dropping out from higher education institutes, using student activity data collected from a software lecture environment [11] or incorporating the incoming data with semester-wise enrollment information [8]. Although many studies claimed that students' behavioral signals and incoming academic profile are important predictors of academic outcome, not many of them addressed how these different types of signals are related to each other.

In this paper, we explored how students' behavioral signals and incoming profiles can be integrated to describe and predict student success in a higher education STEM course. First, we identified the list of significant behavioral variables collected from a lecture software to describe students' academic success from a course. Second, we compared the different likelihood of succeed in a course between student groups. We also investigated how behavioral patterns are different in each student group. Third, we compared the predictive performance of statistical models with weekly accumulated training data, including each model's *day-one* performance and how the performance changes by adding students' weekly behavioral data. The findings of this paper can set the ground for integrating different types of behavioral signals and incoming academic profiles for building an EWS. Detailed investigation of these signals can also provide data-driven evidence for developing customized course-taking strategies.

2 Related Works

2.1 Predicting Student Performance

In many studies, behavioral signals are significant predictors of the student's academic outcome [2,22]. In a higher education context, studies used behavioral data to predict the student's performance in a course. Studies like [27,28,30] suggested how different behavioral signals observed in LMS, such as frequency or length of interactions, can be significant predictors of student success in higher education courses. Studies from Waddington et al. [25,26] specifically focused on using different types of course resources accesses to predict student success in entry-level STEM courses. They showed that course resources related to exam preparations and lecture materials are more significantly related to the course grade than other resource types, such as course information or assignments. Real-time stream data from courses can illustrate more details of student behaviors during learning. Studies like [5] showed behavioral signals, such as the number of interactions with video, exercise, or assignment, can be useful predictors of student engagement in MOOC courses. Other studies on MOOC suggested the use of implicit signals, such as click patterns or use of language in the forum group [23,29], and interaction with video lectures or timely completion of assignments [20,21] to predict different cognitive states of students.

Behavioral signals contain rich information on how students interact with the learning materials. However, other factors, such as students' incoming academic profile [17] and temporal conditions [18], can add more contextual information

about students' learning. Many studies included non-behavioral factors, such as demographic information, in their models to control the effect of behavioral variables with student learning [6,25]. The non-behavioral factors also can be used to capture more complex patterns in student behaviors. Studies showed that interaction with peers [9] or intervention from the course [13] need to be designed differently by students' cultural background or the learning style that they are accustomed to [12]. To provide customized advice for each student's learning, identifying interactions between non-behavioral factors and behavioral variables would be important. In our study, we will investigate how different types of information, such as students' behavioral interactions with the lecture software and their incoming profile factors, can be used to describe and predict student success in a higher education course.

2.2 Early Warning Systems

EWS is a computerized system that focuses on identifying at-risk students early and providing data-driven evidence for developing strategies that can maximize students' academic success [1,15]. Many existing studies on early warning systems are designed to predict the student's retention in a course. Studies like [11,24] evaluated the quality of signals from earlier weeks and model's performance by making a prediction on the later weeks. More broadly, [8] modeled students' dropout from an institution in a semester level. The authors used various types of data. They used semester-level variables like GPA, credit hour, or enrolled year, along with incoming academic profiles like demographic factors and entrance exam information. Earlier identification of at-risk students can help institutions to improve students' course retention [11] and degree completion rates [8]. However, it can also benefit high achieving students to keep their success in a course. In our study, we will suggest an EWS that can be used in higher education institutions where the retention rate can be a lesser problem, and focus on the earlier prediction of students' *success* in a course.

2.3 Overview of the Current Study

Based on previous studies, we formulated the following research questions:

RQ1: Can we identify significant behavioral predictors of student success in a course across different weekly ranges? Answering the first research question will examine the significance and stability of behavioral predictors collected from a lecture software for describing and predicting student success in a course. Selected behavioral variables will also be used to answer later research questions.

RQ2: What is the benefit of including incoming profile factors in models to describe student success? The second research question will compare the model developed from RQ1 with a mixed-effect model, which can address variances between student groups for describing student success in a course.

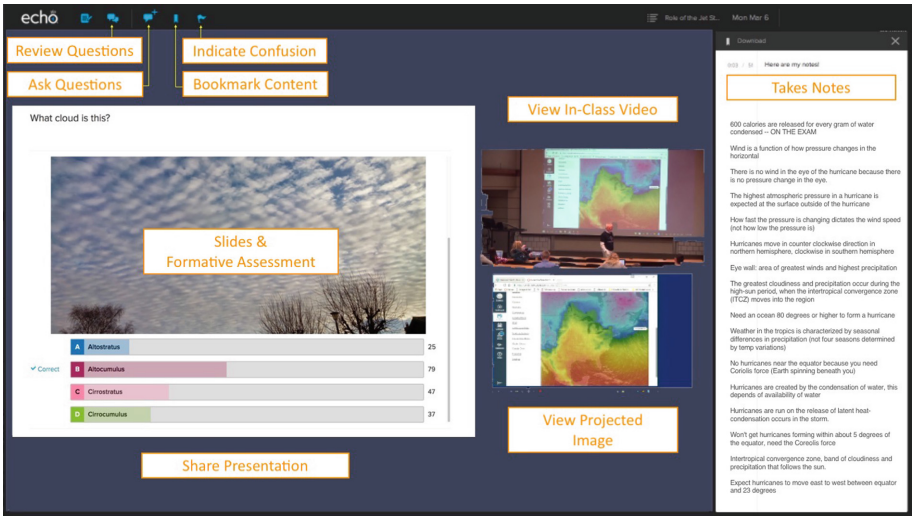

Fig. 1. A screenshot of Active Learning Platform. Students could attend or review the lecture by seeing the lecture video and slides simultaneously. Also, they could take notes or ask questions related to the lecture.

RQ3: Can we identify distinctive behavioral patterns between student groups in terms of success in a course? The third research question will aim to expand the model from RQ2 by adding random slope structures. The results will identify if there are significantly different behavioral patterns exist between student groups.

RQ4: Can we predict student success at a weekly level by using the findings from RQ1-3? The last research question will investigate the predictive power of models suggested in RQ1-3, especially in a weekly-level early warning scenario.

3 Methods

3.1 Data Source

Behavioral Signals. For this study, we collected students' behavioral signals from lecture software. Data were collected across seven semesters, from the year 2015 to 2018. Students were enrolled in an entry-level STEM course, *CLIM 999*. *CLIM 999* was a survey course, and part of 'science distribution' requirement at the university. The course covered the physics of extreme weather events and potential relationships with a changing climate.

CLIM 999 was a blended learning course [30] using *Echo 360*'s Active Learning Platform (ALP) software[1]. By using ALP, students could access the physical classroom session through a web browser. It included video recordings and

[1] One of the authors, Dr. Samson, is a consultant to Echo360 Inc. and uses the Echo360 Active Learning Platform in his class.

Table 1. Distribution of students' incoming GPA and gender profiles. `None` represents the students without incoming GPA information (e.g., the first semester Freshmen or transfer students).

Incoming GPA	Female	Male	All
None	50 (3.9%)	83 (6.5%)	133 (10.4%)
(0.0,2.5]	21 (1.6%)	59 (4.6%)	80 (6.2%)
(2.5,3.0]	75 (5.8%)	157 (12.3%)	232 (18.1%)
(3.0,3.5]	180 (14.1%)	279 (21.7%)	459 (35.8%)
(3.5,4.0]	153 (11.9%)	226 (17.6%)	379 (29.5%)
All	479 (37.3%)	804 (62.7%)	1283 (100.0%)

slides of the lecture. It also provided a virtual learning environment, where students could answer in-class activity questions, take notes, ask questions to other students and instructors, and mark where they felt confused from the lecture (Fig. 1). ALP recorded students' interactions with the system. In this study, we used frequencies of different behavioral signals as predictors (Table 2). Following the previous studies [25,26], we normalized recorded frequencies from ALP by using a percentile rank method for each semester. We expected this would normalize the outlier data-points, and make behavioral signals collected from multiple semesters easily comparable.

Academic Profiles. For the study, we also collected students' incoming profile factors to represent their different academic profiles. To do this, we used a combination of incoming GPA and gender information, retrieved from the university's data-warehouse. To help with better convergence of mixed-effect modeling results, we filtered out data from 63 students since they did not have gender information. As results, we used data from 1283 students for the analysis. These students' profiles were represented by the combination of 5-level incoming GPA labels and 2-level gender labels (Table 1).

Student Success. We labeled the student's performance as *successful* if she achieved 80% or better on the average of three exams from the course. As results, we labeled 53.4% (685) of students as *successful*[2].

The threshold for *successful* label was decided based on the nature of our dataset and previous studies. Our data were collected from students who enrolled in the R1 university of the U.S.[3]. Students in these institutions may aim higher than just passing the course. For example, in our dataset, only 4.53% of students

[2] The distribution of grades were: 178 students with A (90 or higher, 13.87%), 507 with B (80–90, 39.52%), 373 with C (70–80, 29.07%), 167 with D (60–70, 13.02%), and 58 with F (60 or lower, 4.52%).

[3] We follow the definition of R1 university in here: https://en.wikipedia.org/wiki/List_of_research_universities_in_the_United_States.

Table 2. List of behavioral variables (`Beh.`) and student profile factors (`Aca.`). Numbers in parentheses represent the number of times each behavioral variable was selected from the step-wise process.

Type	Name	Description
Beh.	`ActivitiesParticipation`(9)	Activities they submitted an answer to
	`ActivitiesCorrect`(10)	Students get score for correct answer on quiz
	`Attendance`(0)	Students entered the SW interface during classtime
	`NotesCount`(0)	Notes word count
	`NotesInteraction`(0)	Number of times user interacted/edited notes
	`QnA`(3)	Each time they created a question or response
	`SlideDeckView`(10)	Viewed over 5% of a presentation
	`SlideConfused`(10)	Marked confusing in slide
	`VideoView`(10)	Viewed over 5% of a video
	`VideoConfused`(0)	Marked a scene confusing in video
Aca.	`GPA`	Students' average GPA from previous semesters
	`Gender`	Binary label (female or male) for students' gender.

achieved the 60 or less average score (F). Previous works also specified that academic advisers in these institutions consider getting B (i.e., 80% of total grade) or better is an important goal for their students, to pursue entering graduate schools or getting a job easily in the STEM field [17,25,26]. We realize that the results of our study may not be generalized for preventing students from failing the course or degree [8,11]. However, we believe our findings can provide unique insights on designing an EWS for high achieving students to keep succeeding in their institution, and sets the stage for subsequent work that includes the analysis on other types of higher education courses.

3.2 Building Models

Selecting Behavioral Predictors. To answer the first research question (RQ1), we used a cross-validation method to identify more meaningful behavioral predictors. In each cross-validation fold, a generalized linear regression model (GLM) model was initially fitted using all 10 behavioral predictors from the data excluding the held-outs. Then the backward and forward step-wise process

was applied to the model, using Akaike Information Criterion (AIC) with R's [19] `step` function. After the cross-validation process is finished, we counted the number of times that each variable was selected in individual folds (Table 2).

Descriptive and Predictive Models. Using selected behavioral predictors as fixed-effect predictors, we firstly built descriptive models using the whole dataset without cross-validation. The GLM descriptive model expected to show how the entire dataset can be described solely using behavioral signals (RQ1). We also built generalized linear mixed-effect regression models (GLMER) to address variances across incoming profile factors, incoming GPA and gender (RQ2), and different behavioral patterns between these student groups (RQ3). We used `glmer` function with `nAGQ=0` argument from `lme4` [3] to fit GLMER models. Like the `GLM` model, GLMER descriptive models were also fitted with the whole dataset.

We also developed prediction models that can provide a weekly forecast of student success in the course (RQ4). The structure of predictive models followed GLM and GLMER models that used for the previous research questions (RQ1-3). The performance of predictive models was evaluated with held-out sets from 10-fold cross-validation. Average scores of evaluation metrics, such as prediction accuracy, area under the curve scores of receiver operation characteristics (ROCAUC) and precision-recall curve (PRAUC), were used to compare the prediction performance of each model. The evaluation included how early the model can effectively predict student success in the course. For example, testing models' performance with `Week 0` subset was considered as a *day-one* scenario, which allowed models to predict student success without any behavioral signals. Testing with `Week 1-5` subset used behavioral signals accumulated until the fifth week of the semester. Similarly, `Week 1-15` subset included all students' behavioral data recorded across the entire semester.

4 Results

4.1 Selecting Behavioral Variables

In Table 2, we noted how many times each fixed-effect variable were selected from the variable selection process. It shows that behavioral variables like `ActivitiesCorrect`, `SlideDeckView`, `SlideConfused`, and `VideoView` were significant predictors across all cross-validation folds. We call these variables as `Unanimous` set. Another variable set is `MoreOnce` set, which contains additional variables of `ActivitiesParticipation` and `QnA` that were selected more than once from the variable selection process.

From the ANOVA analysis using residual deviance scores with the GLM descriptive model setting, we found that the model trained with `MoreOnce` variable set (1570.8) show similar performance from the model using all 10 behavioral variables ($1568.8, p = 0.738$). This model also showed significantly better performance than the model using `Unanimous` variable set ($1577.5, p < 0.05$). Thus,

we used `MoreOnce` variable set for further analyses, keeping the model structure simpler but minimizing the performance sacrifice.

To answer the first research question (RQ1), we built descriptive GLM and GLMER models, using `MoreOnce` variable set, and see how coefficients change across different week ranges. Figure 2 shows that coefficients for variables in `MoreOnce` set were most significant in different week and model conditions, except `QnA` and `ActivitiesParticipation` variables.

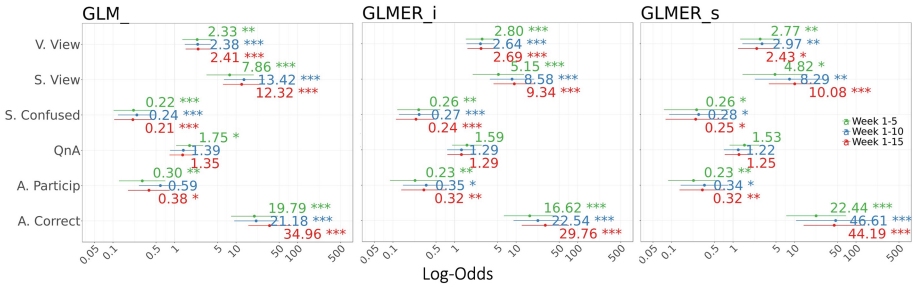

Fig. 2. Coefficients for behavioral variables were relatively consistent across different models and weekly accumulation ranges. Error bars indicate the 95% confidence interval of coefficient estimates. ($^{*}: p < 0.05$, $^{**}: p < 0.01$, $^{***}: p < 0.001$)

4.2 Integrating Student Profile Information

To answer RQ2 and RQ3, we built GLMER models that include additional random intercepts from incoming profile factors (`GLMER_i`) and random slopes from behavioral variables (`GLMER_s`). From each GLMER model, we expected to see if the model captures different starting points (RQ2) and some behavioral pattern differences (RQ3) between different student groups.

Figure 3 shows variances between different student groups (random intercepts) and behavioral patterns (random slopes) from the `GLMER_s` model[4]. First, in the most left panel for random intercepts, we could find significant gender differences among mid-level incoming GPA groups. While there were positive relationships between students' success in incoming GPA levels, the estimated random intercept for female students in the `(3.0,3.5]` incoming GPA group was almost as similar as the intercept of lower incoming GPA male students (`(2.5,3.0]`). This pattern was also observed between female students in the `(2.5,3.0]` incoming GPA group and male students in the `(0.0,2.5]` group. These gender gaps were not observed in the highest (`(3.5,4.0]`) or the lowest (`(0.0,2.5]`) incoming GPA groups, or the first semester students without incoming GPA information (`none`).

Second, other panels in Fig. 3 provide some potential customized advising scenarios. For example, we could see some significant random slope coefficients. For students with the highest incoming GPA group (`(3.5,4.0]`), providing correct

[4] The results for random intercepts were similar between `GLMER_i` and `GLMER_s` models.

answers to in-class activity questions (`ActivitiesCorrect`) and more reviewing of the slides (`SlideDeckView`) were less important for predicting their success in the course. Additionally, in the highest incoming GPA group of female students, marking confusion (`SlideDeckConfused`) was more significantly related to their success. These results show that different student groups may need different study strategies to succeed in a course.

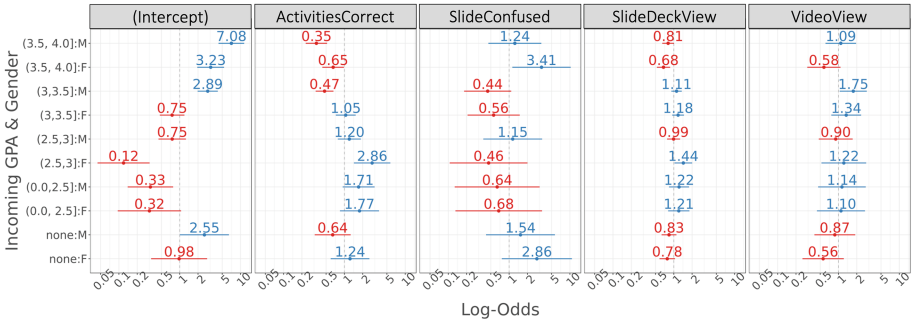

Fig. 3. Random effect results from the `GLMER_s` model, using Week 1–15 accumulated data. Significant gender gaps in random intercepts were observed in `(3.0,3.5]` and `(2.5,3.0]` incoming GPA groups. Also, some random slope coefficients for `ActivitiesCorrect`, `SlideDeckView`, and `SlideDeckConfused` showed significantly different patterns by incoming profiles.

4.3 Predictive Modeling: Weekly Performance

Lastly, we examined the prediction performance of GLM and GLMER models (RQ4)[5]. Figure 4 shows that all models reach better performances when training data is accumulated through more weeks. All models quickly achieved meaningful ROCAUC scores of grater than 0.5 (maximum scores were 0.724 (`GLM`), 0.791 (`GLMER_i`), and 0.789 (`GLMER_s`)). PRAUC results were similar to ROCAUC results. Across the semester weeks, both GLMER models using random effect structures achieved the maximum accuracy near 72%. Both models performed consistently better than the `GLM` model, which used behavioral predictors only (66.3% maximum accuracy; 8.7% relative worse than GLMER models).

To measure the *day-one* performance in `Week 0`, models predicted the majority label (`GLM` model) or relied on the incoming profile information only (GLMER models). For the `GLM` model, the average accuracy was 53.33%, which is equivalent to the likelihood of students achieving an average of 80 or better in exams. However, GLMER models could predict students' success in the course with 70% accuracy using incoming profiles alone (Fig. 4). The `GLM` model showed maximum accuracy in week 11 and 12. For both GLMER models, the maximum accuracy scores were achieved in week 15. However, adding behavioral predictors to GLMER models did not provide meaningful gains to their predictive performances.

[5] Detailed prediction results can be found at http://bit.ly/nam-EWSpreds.

We also examined a prediction performance for models with a single important behavioral variable (`ActivitiesCorrect`) from Sect. 4.2. A GLM model only achieved 59.1% accuracy and 0.67 ROCAUC scores. Both GLMER models showed around 70% accuracy and 0.77 ROCAUC scores, which are not so different from the *day-one* performance. All these best performing scores were observed in week 12 or 13.

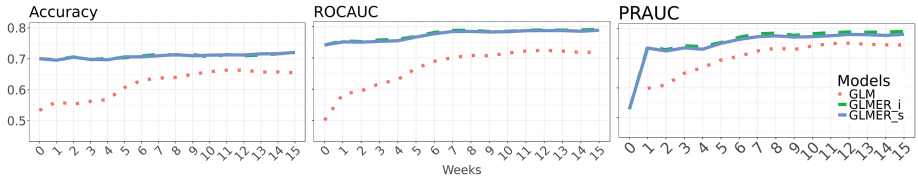

Fig. 4. Weekly prediction performance of GLM and GLMER models. All models' performed increasingly better with more training data. While both GLMER models performed better than the `GLM` model, the differences between the two were insignificant.

5 Conclusion and Discussion

In this paper, we identified significant behavioral predictors of student success from a lecture software. We also found gender gaps in the mid-level incoming GPA groups, and differences in behavioral patterns between student groups. By combining behavioral signals and incoming profile factors, our early warning prediction model achieved up to a 72% accuracy for predicting student success in an entry-level STEM course. We believe these findings can provide more contexts to student behaviors in higher education settings. Practically, it can give data-driven evidence for developing more personalized course-taking strategies, which would be helpful for academic advisers, instructors, and students.

Based on our findings, we have a few discussion points for future studies. First, like other studies in entry-level STEM courses [14], we could observe the gender gap between students. As specified in Sect. 4.2, it was surprising that the gender gap existed even among students that share similar GPAs from the previous semesters. However, in this study, we were not able to provide further explanations of what are the relevant causes of this gap. Including additional demographic or enrollment information to the analysis, or comparing student behaviors between multiple STEM courses may provide deeper contexts for our current findings. Second, further investigation on why different behavioral patterns were observed between student groups would be interesting. For the results in Sect. 4.2, we only suspect that higher incoming GPA students might have explored more on incorrect answers with in-class questions, or spent lesser time on lecture slides because they tend to have their own study notes. Using qualitative methodologies, such as interviewing students or observing them during class, may provide more detailed insights on these behavioral differences between students. Lastly, developing more effective predictive features would improve

the models' prediction performances. As we saw in Sect. 4.3, adding behavioral signals to GLMER models provided only marginal increases to the prediction performances from `Week 0`. Adding multiple behavioral variables to the model was more effective than adding a single important behavioral variable. However, since the improvement was not very significant, we may need to explore other exogenous features that can provide additional information to the model that are not addressed with current incoming profile factors. Developing non-frequency based variables, such as a semantic representation of student notes or predicted student engagement states during the lecture, can be also helpful options for developing more sophisticated and accurate predictive models.

References

1. Aguilar, S., Lonn, S., Teasley, S.D.: Perceptions and use of an early warning system during a higher education transition program. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, pp. 113–117. ACM (2014)
2. Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z.: Off-task behavior in the cognitive tutor classroom: when students game the system. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 383–390. ACM (2004)
3. Bates, D., Máchler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. J. Stat. Softwa. **67**(1), 1–48 (2015). https://doi.org/10.18637/jss.v067.i01
4. Beck, H.P., Davidson, W.D.: Establishing an early warning system: predicting low grades in college students from survey of academic orientations scores. Res. High. Educ. **42**(6), 709–723 (2001)
5. Bote-Lorenzo, M.L., Gómez-Sánchez, E.: Predicting the decrease of engagement indicators in a MOOC. In: Proceedings of the Seventh International Learning Analytics and Knowledge Conference, pp. 143–147. ACM (2017)
6. Brown, M.G., DeMonbrun, R.M., Teasley, S.D.: Don't call it a comeback: academic recovery and the timing of educational technology adoption. In: Proceedings of the Seventh International Learning Analytics and Knowledge Conference, pp. 489–493. ACM (2017)
7. Campbell, J.P., DeBlois, P.B., Oblinger, D.G.: Academic analytics: a new tool for a new era. Educase Rev. **42**(4), 40 (2007)
8. 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. In: Proceedings of the Eighth International Learning Analytics and Knowledge Conference, pp. 270–279. ACM (2018)
9. Davis, D., Jivet, I., Kizilcec, R.F., Chen, G., Hauff, C., Houben, G.J.: Follow the successful crowd: raising mooc completion rates through social comparison at scale. In: Proceedings of the Seventh International Learning Analytics and Knowledge Conference, pp. 454–463. ACM (2017)
10. Goldstein, P.J., Katz, R.N.: Academic analytics: the uses of management information and technology in higher education, vol. 8. Educause (2005)
11. Hlosta, M., Zdrahal, Z., Zendulka, J.: Ouroboros: early identification of at-risk students without models based on legacy data. In: Proceedings of the Seventh International Learning Analytics and Knowledge Conference, pp. 6–15. ACM (2017)

12. Hoffmann, B., Ritchie, D.: Using multimedia to overcome the problems with problem based learning. Instr. Sci. **25**(2), 97–115 (1997)
13. Kizilcec, R.F., Cohen, G.L.: Eight-minute self-regulation intervention raises educational attainment at scale in individualist but not collectivist cultures. Proc. Natl. Acad. Sci. U.S.A. **114**(17), 4348–4353 (2017)
14. Koester, B.P., Grom, G., McKay, T.A.: Patterns of gendered performance difference in introductory stem courses. arXiv preprint arXiv:1608.07565 (2016)
15. Krumm, A.E., Waddington, R.J., Teasley, S.D., Lonn, S.: A learning management system-based early warning system for academic advising in undergraduate engineering. In: Larusson, J.A., White, B. (eds.) Learn. Anal., pp. 103–119. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-3305-7_6
16. Macfadyen, L.P., Dawson, S.: Mining LMS data to develop an "early warning system" for educators: a proof of concept. Comput. Educ. **54**(2), 588–599 (2010)
17. Nam, S., Lonn, S., Brown, T., Davis, C.S., Koch, D.: Customized course advising: Investigating engineering student success with incoming profiles and patterns of concurrent course enrollment. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, pp. 16–25. ACM (2014)
18. Park, J., Yu, R., Rodriguez, F., Baker, R., Smyth, P., Warschauer, M.: Understanding student procrastination via mixture models. In: Proceedings of the International Conference on Educational Data Mining (2018)
19. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2018). https://www.R-project.org/
20. Ramesh, A., Goldwasser, D., Huang, B., Daume III, H., Getoor, L.: Learning latent engagement patterns of students in online courses. In: Twenty-Eighth AAAI Conference on Artificial Intelligence (2014)
21. Sinha, T., Li, N., Jermann, P., Dillenbourg, P.: Capturing "attrition intensifying" structural traits from didactic interaction sequences of MOOC learners. In: Proceedings of the EMNLP 2014 Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses, pp. 42–49 (2014)
22. Skinner, E.A., Kindermann, T.A., Furrer, C.J.: A motivational perspective on engagement and disaffection: conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. Educ. Psychol. Meas. **69**(3), 493–525 (2009)
23. Tomkins, S., Ramesh, A., Getoor, L.: Predicting post-test performance from online student behavior: a high school mooc case study. In: Proceedings of the International Conference on Educational Data Mining (2016)
24. Veeramachaneni, K., O'Reilly, U.M., Taylor, C.: Towards feature engineering at scale for data from massive open online courses. arXiv preprint arXiv:1407.5238 (2014)
25. Waddington, R.J., Nam, S., Lonn, S., Teasley, S.D.: Improving early warning systems with categorized course resource usage. J. Learn. Anal. **3**(3), 263–290 (2016)
26. Waddington, R.J., Nam, S.: Practice exams make perfect: incorporating course resource use into an early warning system. In: Proceedings of the Fourth International Learning Analytics And Knowledge Conference, pp. 188–192. ACM (2014)
27. Wang, A.Y., Newlin, M.H.: Characteristics of students who enroll and succeed in psychology web-based classes. J. Educ. Psychol. **92**(1), 137 (2000)
28. Wang, A.Y., Newlin, M.H.: Predictors of performance in the virtual classroom: identifying and helping at-risk cyber-students. J. (Technol. Horiz. Educ.) **29**(10), 21 (2002)

29. Yang, D., Kraut, R., Rose, C.P.: Exploring the effect of student confusion in massive open online courses. J. Educ. Data Min. **8**(1), 52–83 (2016)
30. Zacharis, N.Z.: A multivariate approach to predicting student outcomes in web-enabled blended learning courses. Internet High. Educ. **27**, 44–53 (2015)