UNLV College of EDUCATION
Department of Educational Psychology & Higher Education

MALT·lab
Research on Metacognition and Motivation in Advanced Learning Technologies

Predicting STEM Achievement with Learning Management System Data: Prediction Modeling and a Test of an Early Warning System

Michelle M. Dominguez[1], Matthew L. Bernacki[1], P. Merlin Uesbeck[2]
doming90@unlv.nevada.edu matt.bernacki@unlv.edu uesbeck@unlv.nevada.edu

[1]Department of Educational Psychology & Higher Education, [2]Department of Computer Science

Overview

- Prediction modeling often operates using a fully data-driven approach.
- We employed 1) 4 weeks of logged LMS behaviors (Blackboard Learn) in a large scale lecture course to build a model predicting achievement and 2) used learning theory to guide principled representation of data in order to create higher-order features and improve accuracy of models predicting college achievement.
- Once a serviceable model was produced, we used it to inform an Early Warning System to alert students at risk of poor course performance.

Research Questions

1. Will early LMS log data be sufficient to predict student performance?
2. Can classifying types of student activities by the kind of content accessed & selecting a corresponding data representation (i.e., dichotomous vs. count) improve the accuracy of our prediction models?
3. When these prediction models are applied to new data sets, does it maintain the same level of accuracy in predicting student outcomes?

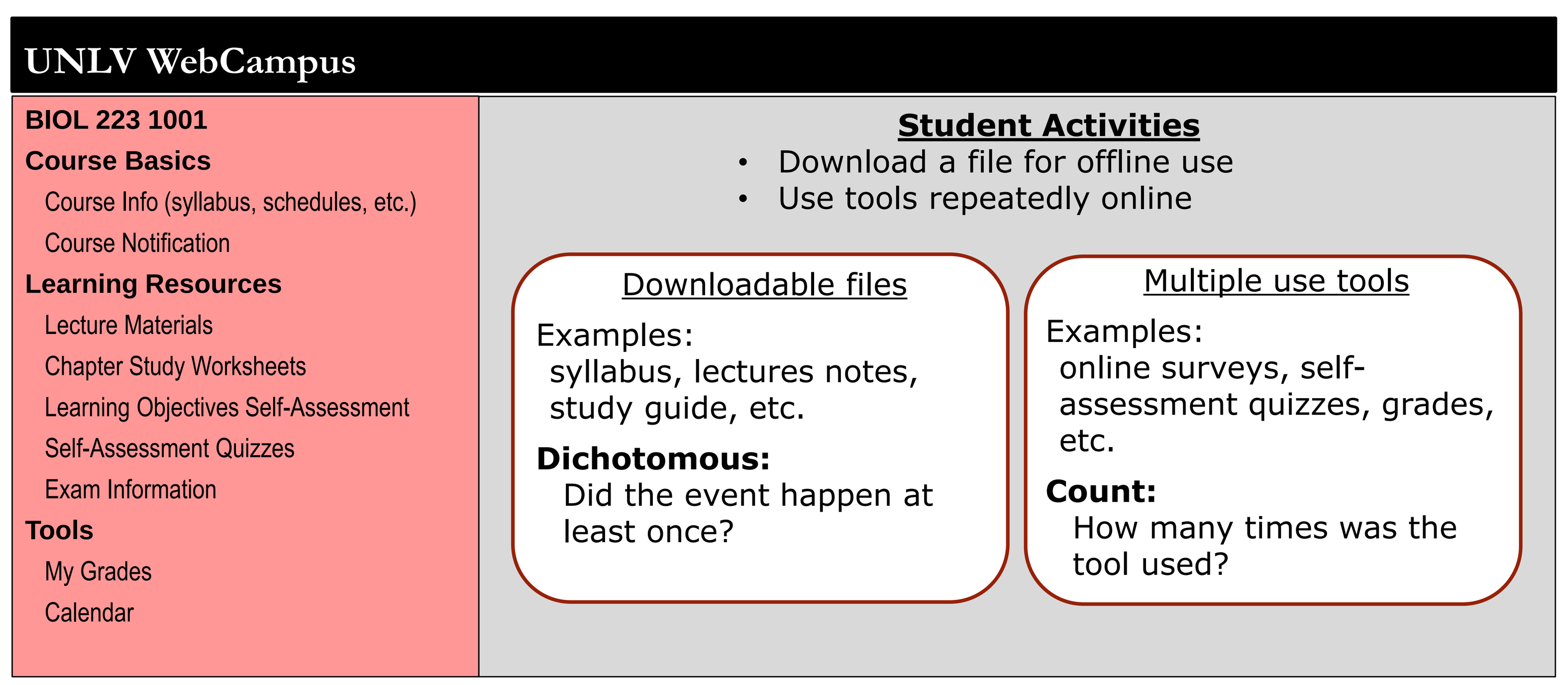

Figure 1. Blackboard Learn LMS Course Site (i.e., UNLV Webcampus)

Methods

- Blackboard Learn logs were taken from 326 student during the Fall semester of a large lecture in Biology. Events were organized by name of content used, classification of item, and week.
- Prediction modeling approaches attempted included forward selection logistic regression, J-48 and J-Rip decision trees, and a Naïve Bayes model.
- In order to test the usability of the prediction model, behavioral data was taken from additional 298 students from the following Spring semester and another 349 students from the following Fall semester.

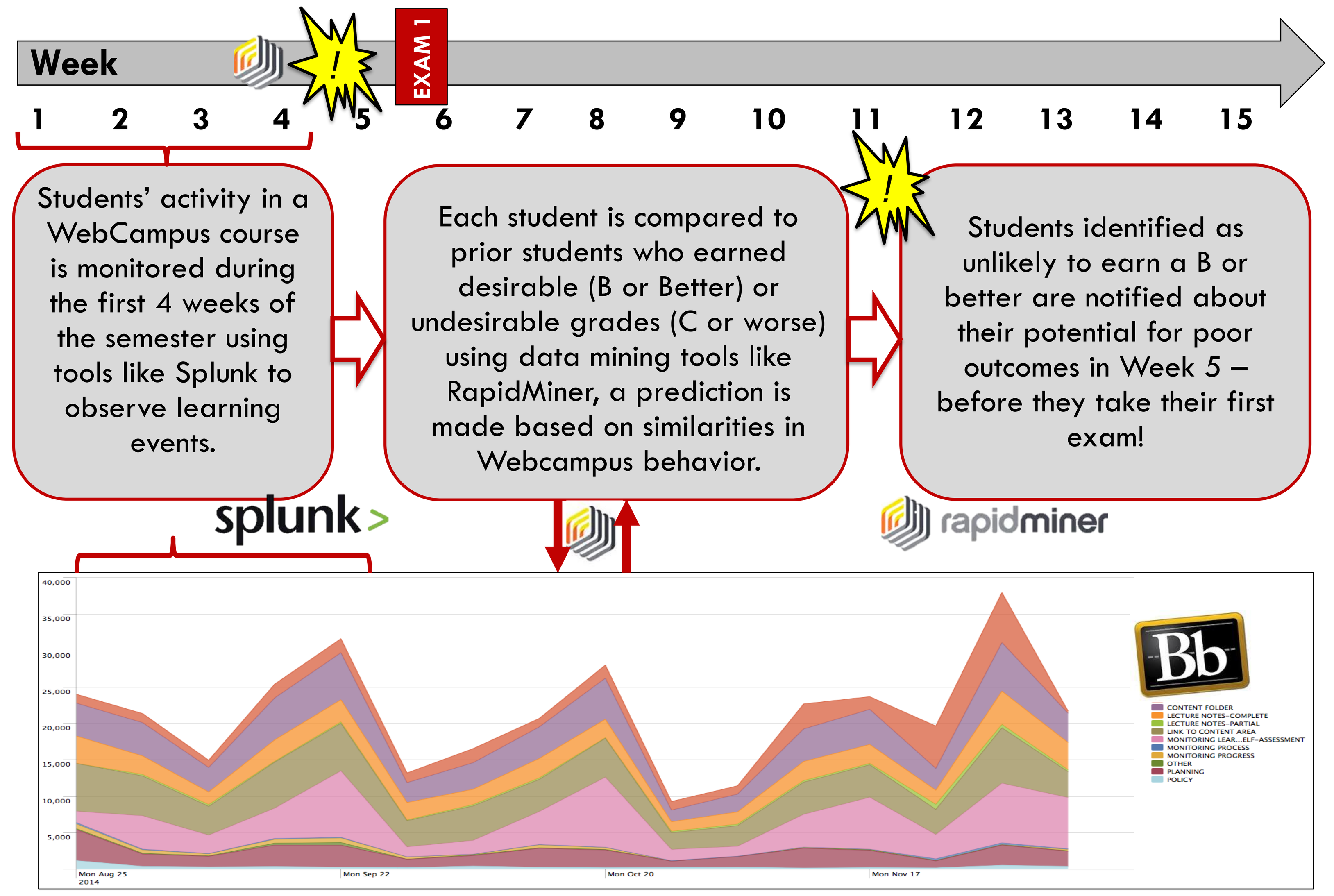

Figure 2. Student activities on Blackboard Learn over the course of the semester broken into different content types

Results

- The principled representation of the data produced the best combination of kappa (κ) = 0.212 and recall = 84.24%
- The prediction model based on this representation was used on two subsequent semesters: Spring and Fall
- The similar recall values suggest that the model is consistently accurate at identifying students likely to fail to obtain the grade needed to advance on in their major.

Table 1. Prediction models using different versions of data and using best model on subsequent semesters

Data representation	κ	Accuracy (%)	Precision (%)	Recall (%)	True: Predicted			
					1:1	1:0	0:1	0:0
count	.16	61	61	82	48	94	34	150
dichotomous	.17	60	63	72	63	79	52	132
both	.22	63	65	73	69	73	49	135
principled	.21←	63	63	84←	51	91	29	155
Future Semesters								
Spring	.07	53	52	85	33	117	22	126
Fall	.15	58	57	81	56	112	34	147

Note. The baseline for test data versions (count, dichotomous, both & principled) is 56%. The baseline for the Spring use data is 51% and the baseline for Fall use data is 52%.

- The prediction model with the best combination of kappa and recall was the principled representation of the data (treating downloadable files as dichotomous and multiple use tools as count).

Table 2. Prediction model with principled count & dichotomous model attributes and weights

Attribute	Weight
Intercept	-.42
Overall Activity (count)	
Week 1 Lecture Materials link	.01
Week 2 Unit 1 Lecture folder	.10
Any Access of Unit 1 Self-Assess folder	.07
Lectures (dichotomous)	
Week 1 Ch 4 Tissues	1.81
Week 1 Ch 11 Muscular System	-.49
Week 2 Ch 4 Tissues	.59
Any Access of Ch 2 Macromolecules	-.54
Planning (dichotomous)	
Week 2 Ch 1 Review of Intro to A&P	-.78
Monitoring Progress on Learning Goals (count)	
Week 3 Learning Goals Checklist Ch 4	-9.78
Monitoring Learning (count)	
Week 4 Ch 3 Self-Assessment Quiz	.01
Tools (data as count)	
Week 4 View of Course Calendar	-1.52

Ongoing Research

- The prediction model (Table 2) was programmed back into the Splunk data model so that a prediction could be made after four weeks for each student in the current semester of the biology course (i.e., likelihood of obtaining a B or better, and being able to move on to the subsequent course).
- Students were randomly assigned to Follow or Alert groups; an early warning message was sent from the instructor through the LMS correspondence tool to all in Alert Groups. Each message included a salutation, indication of the upcoming exam, and a redirect of the student to helpful resources available on the LMS for students to use. Four Alert conditions were created to test effects of **personalizing** the salutation (with first name) and of **providing feedback** (that their behaviors are similar to those of prior poor performers).

Research Question 1. How does the alert message affect behavior?

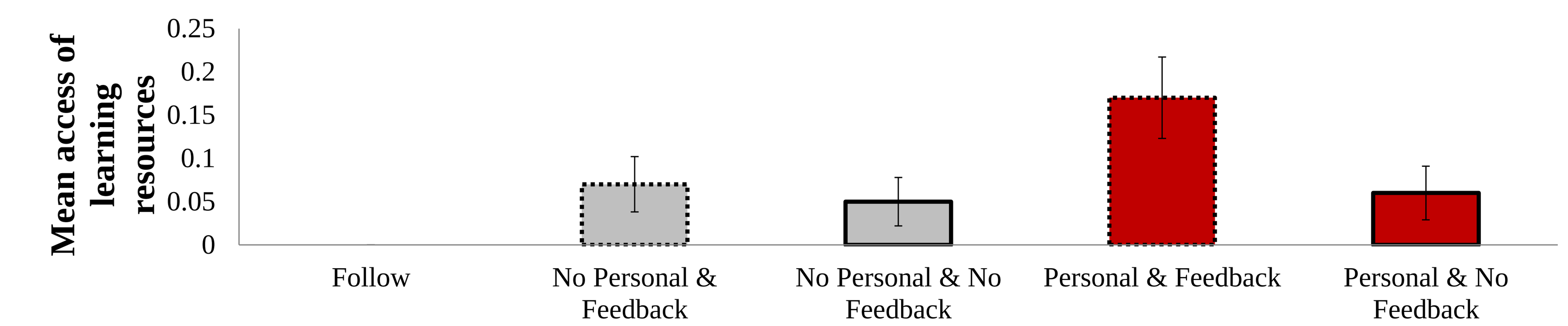

Figure 3. Access of learning resources among the groups that were predicted to not obtain a B or better

- Those who received alerts accessed learning resources at a greater rate than the follow group (i.e., received an LMS system notification "new content posted"), $F(4, 312)=4.00$ [$MS_E=.06$], $p=.004$, $\eta_p^2=.049$
- Access was more frequent when messages were Personalized with Feedback as compared to Follow Group; $p < .01$ level ($p=.001$, $d=.64$)
- Access by Personal with Feedback Group was marginally significantly greater than No Personal & No Feedback group; $p < .10$ level ($p=.057$, $d=.39$)

Research Question 2. How does the alert message affect performance?

- Students who received Alerts **did not** differ in performance from those who did not received an alert message on exams
- $F(3,296)=.846$, $p=.47$; Wilk's $\Lambda=.992$ (Figure 4)

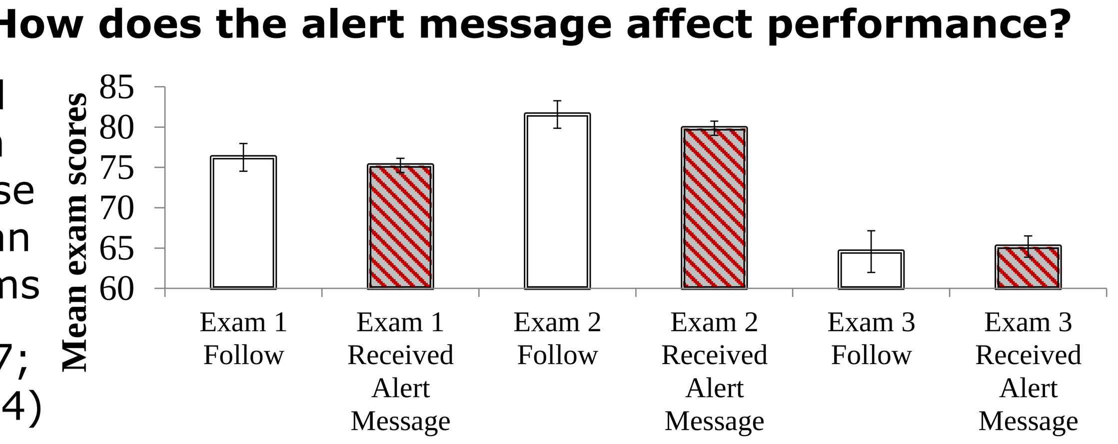

Figure 4. Effect of alert message on Exams 1, 2, & 3

Research Question 3. How do message *features* affect performance?

- Alerts seem to be effective in prompting resource access, but prompting use of that resource did not change performance. However, Features of the message do seem to impact performance. Within the Alert groups, those that received a personalized message performed better than those that did not on Exams 2 & 3.
- Exam 2: $F(1, 298)=4.08$ [$MS_E=178.183$], $p=.044$, $d=.25$ (Figure 5a)
- Exam 3: $F(1, 298)=4.40$ [$MS_E=407.194$], $p=.037$, $d=.12$ (Figure 5b)

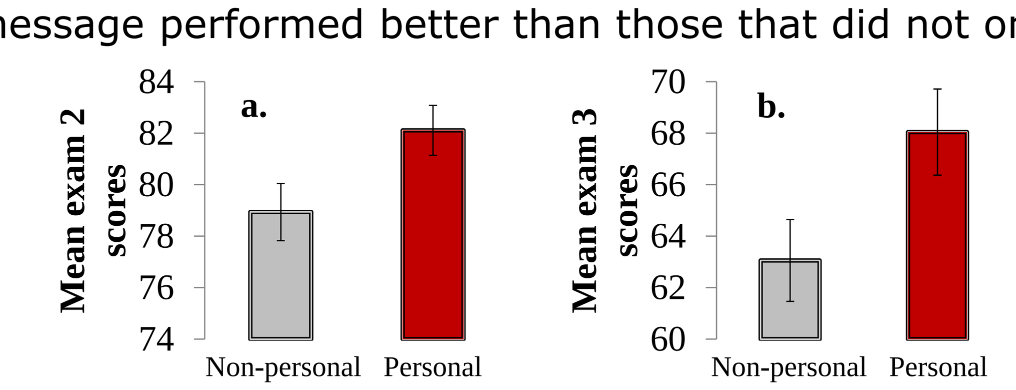

Figure 5. Effect of personalization on Exams 2 & 3

- Additional analyses are needed to explore differences among alerted groups: 1) personalization, feedback affect type or pattern of study behavior, or 2) potential mediation where alerts affect motivation, explain performance

Funded by National Science Foundation Award # DRL-1420491
In kind support from UNLV Office of Information Technology