How does Performance in an Online Primer Predict Achievement in a Future Computer Science Course?

Soniya Gadgil
Eberly Center for Teaching Excellence and Educational Innovation
soniyag@andrew.cmu.edu

Steven Moore[1], John Stamper[2]
Carnegie Mellon University, HCII
StevenJamesMoore@gmail.com[1]
jstamper@cs.cmu.edu[2]

ABSTRACT: We describe a feature engineering approach to predict future course performance based on students' interactions with an online math primer. To help incoming computer science freshman students gain competency on core discrete math concepts, we developed a primer course deployed in an interactive learning environment. The primer covered three foundational topics — logic, sets, and functions. Students completed this primer in the summer prior to their first semester as computer science undergraduates. We used random forest modeling and linear regression to understand which features predict performance in a subsequent face-to-face math course. Results indicated that students' performance on two of the three units (sets and functions) was positively associated with final grades, whereas total time spent in the course was negatively associated with final grades. We discuss implications for iterative course design as well as utility of educational data mining approaches for tracking preparation for future learning.

Keywords: Feature engineering · Computer Science Education · Educational Data Mining · Prediction Modeling · Interactive Learning Environment

1 INTRODUCTION

The proliferation of data on students interacting with online learning environments has opened up enormous possibilities for understanding student behavior within the last decade or two (Baker & Inventado, 2014). It has also enabled iterative improvements of these learning environments to promote student learning. However, a key challenge is to understand what aspects of students' behavior are most predictive of success in future learning situations.

In recent years, there have been calls to assess learning in terms of "robust" learning outcomes, going above and beyond traditional pretests and posttest which often measure only shallow encoding and retrieval (Koedinger, Corbett, & Perfetti, 2012). Robust learning refers to whether learning occurs in a way that transfers, prepares students for future learning, and is retained over time. While research on learning in online learning environments has been rapidly increasing, much less work has looked at how online courses prepares students for learning during future learning opportunities, including both online or in-person (Beaubouef, 2002). For example, if a student takes an online introductory course in mathematics, we can tell how the student performed within the system itself, but whether this interaction with online learning prepared the student for future math courses is often unclear (Reilly & Emmett, 2011).

Prior research has attempted to predict student performance on tests of transfer. Specifically, such work has found that avoiding help seeking and making fast responses after bugs were negatively associated with transfer (Baker, Gowda, & Corbett, 2011). Hershkovitz et al. showed that student performance on a transfer test can be predicted by calculating moment-to-moment probabilities of learning a particular skill. Other research has focused on using early course data to predict future success, and develop early warning systems to students identified as at risk for failure (Costa et al., 2017; Dominguez, Bernacki, & Uesbeck, 2016). While prior work on online learning and transfer sheds important light on what attributes of student behavior are critical to transfer, it has largely focused on performance within a single online course. No prior studies to our knowledge have looked at the impact of student performance across multiple online sequential courses or on a future face-to-face course. In this paper, we analyze learning analytics data from an online math primer course and develop a prediction model for performance on an in-person computer science follow-up course students complete.

2 TOOLS & METHODS

2.1 Open Learning Initiative

The Open Learning Initiative (OLI) is an open-ended learning environments that allows instructors to develop online courses consisting of interactive activities and diverse multimedia content. Detailed student interactions with the course materials, such as watching videos or answering questions are logged in the course's database. OLI courses, such as the one used in this study, are often intended to be used asynchronously without an instructor. Prior research has compared student learning from a stand-alone OLI course on introductory statistics to face-to-face equivalent instruction, and found that students showed increased learning gains in half the time as compared to students with the traditional face-to-face instruction (Lovett, Meyer, & Thille, 2008). While this system has been proven to be effective, no studies around it have measured the transfer of the content to future in-person courses. This is true for many online learning environments, while they are proven effective for learning, studies do not look at their transfer and retention when the knowledge is required for a follow-up in-person course, such as a traditional undergraduate one.

2.2 Data Description

Our predictor data came from the Discrete Math Primer (DMP) OLI course, which was completed by incoming freshmen at Carnegie Mellon University during the summer of 2016. This course serves as a prerequisite for core computer science courses, providing students with a foundation for key concepts in the field, such as the notion of data structures. The course is divided into three units — Logic, Sets, and Functions, with which students interacted in a sequential manner. The final grades from the follow-up in-person course, Mathematical Foundations for Computer Science (MFCS), were used as our predictive variable. The final grade was calculated as a percent out of 100. This course was taken by the same students the following semester during Fall 2016 and was taught in a traditional in-person lecture and recitation format. From the syllabus of the follow-up course, proofs is one of the five listed key topics, which makes use of the Logic unit. Functions and Sets is another key topic of the five listed covered in the follow-up course, which takes a deeper dive on the concepts than what is covered by the online DMP course. Performance in this online course is appropriate for predicting the performance on the follow-up as it directly builds upon the topics covered in the DMP course and is thusly a prerequisite of the MFCP course.

Our dataset consists of 34,999 transactions from 139 students. The transactions consist of student actions in the OLI course, such as selecting an answer in a multiple-choice question, requesting a hint, and submitting an answer. These data entries detail UI events, question correctness, time on task, performance on checkpoints, and hints where relevant. In total, the data spans 198.5 hours of student activity in the course. The course consists of twenty three pages, not including the three quizzes, and is comparable in length to a textbook page. Each page consists of instructional text that is interspersed with low-stakes questions that give detailed feedback intended to foster learning. The Logic unit consists of forty-three questions, Sets has twenty-three, and Functions consists of fifty-one for a total of 106 questions we had student data from in the course. Table 1 shows the variables that we used for our analysis.

Table 1: A description of each variable used in the dataset

Variable	Description
ID	A hashed string corresponding to the student
Duration	The time, in seconds, a student interacted with an element, such as a question
Student Response Type	Denotes the student's action, whether it be a hint request, question attempt, page view, or saving their question answer
Level (Module)	States which of the three units the transaction came from
Step Name	The unique name for the part(s) of a problem, each step contains an opportunity for a correct or incorrect response
Outcome	If applicable, whether the student got the problem correct or incorrect
Attempt at Step	Denotes the amount a student has attempted a given question step
Skill	The label for the skill associated with the particular problem step

2.3 Feature Engineering

We performed feature engineering to construct seven key predictors. Prior research has shown that students who perform at or below a failing grade level in an online course tend to have fewer interactions (Davies & Graff, 2005). Each entry in our dataset represents a student transaction, so we were able to count the numbers of transactions each individual student made through the course. Once the data was filtered on a per-student transaction basis, the total duration each transaction took could be summed to generate a student's total duration in seconds within the course.

As previously described, the course is divided into three units — Logic, Sets, and Functions. Each of these concludes with a summative quiz covering the core material covered in the unit. Each quiz consisted of eight questions, and students were only allowed a single attempt per quiz question. The grade for each quiz was calculated by summing the number of correct questions out of eight possible points. This yielded three of our seven analyzed features, which were the final quiz grades for each unit.

For each student transaction that details the submission of a question, the OLI platform denotes if it is the student's first attempt at the question. Subsequently if they attempted the problem again, such as changing their answer and submitting, the following entry for the attempt would be marked with a two in the corresponding column. Using this attempt count in conjunction with the outcome, correct or incorrect of the problem attempt, we are able to determine the accuracy of a student's overall attempts as a percentage. Knowing the student's number of attempts at a question and its outcome also allows us to calculate their accuracy on the last attempt, our final feature. In total, this gives us the following seven features:

1. Number of transactions
2. Duration in course
3. Logic quiz grades
4. Sets quiz grade
5. Functions quiz grade
6. Accuracy of overall attempts
7. Accuracy on last attempt

2.4 Random Forest & Linear Regression

We used random forest model, implemented in the R programming language, to predict final grade performance in the follow-up in-person MFCS course. Random forest modeling is a classification and regression algorithm that estimates the amount of increase in mean squared error for each variable, when it is replaced by a set of random values. This provided us with a weighting of how important each of our seven defined features is in the prediction of the final grade. Following this, we used linear regression to predict the nature of the relationship of the predictor variables from our model and to estimate what percentage of variation in final grades was explained by each predictor variable.

3 RESULTS

The results of the random forest modeling indicated the following variables contributed to the increase in mean square error: total number of transactions, quiz grades for the Sets unit, quiz frade for the Functions unit, number of correct and incorrect attempts, and the duration of time spent in the course, see Figure 1.

A simple linear regression analysis was conducted to predict final grade based on the variables found to be associated with an increase in the mean square error. A significant regression equation was found, R2 = .43, F(7,120) = 12.55, $p < .001$. Results indicated that the accuracy scores on the Sets ($t = 2.25$, $p = .02$) and Functions quizzes ($t = 2.10$, $p = .037$) had a significant positive association with final grade. The total number of transactions was negatively correlated with final grade ($t = -3.07$, $p = .002$). The regression performed on last attempt correct and incorrect was found to not be significant. It is interesting to note that while only the scores on the Functions and Sets quizzes were positively associated with the final grade on the subsequent course, it was not because students were already performing at ceiling levels on the Logic module. Mean scores for the Logic and Sets quizzes were 78% and 77% respectively, whereas mean for the Functions quiz was significantly lower at 55%.

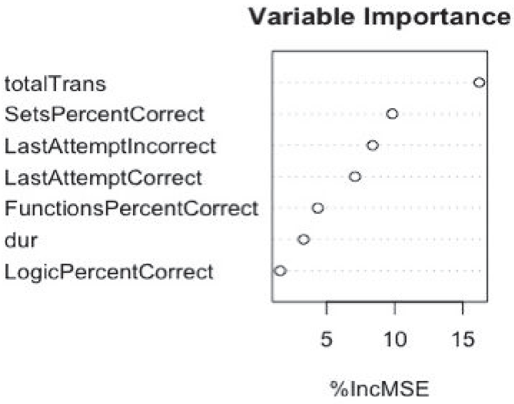

Figure 1: Variable importance plot using random forest modeling

4 DISCUSSION

In this paper, we describe some preliminary results on how students' performance in an online course can be used to predict their learning and performance in a future course. We found that students' performance on two of the three modules in the online OLI course significantly predicted final grades on the subsequent in-person course. The predictive power of the Sets and Functions units, but not the Logic unit, may be explained by the sequence they occur in for both the online DMP course and then in the in-person MFCP course. In the follow-up course, the Proofs section uses subject matter from the Logic unit, and occurs early on in the course. It may be the case that this is a minor section and not a heavily contributing portion of their final grade, since it is the very first part. However, Sets and Functions occurs in the middle of the follow-up course and is taught together. Since these two units are taught together in the follow-up course, it is likely that a student who did not perform well on these two units from the online DMP course will lack the required prior knowledge for this topic and vice-versa. Additionally as it falls in the middle of the course, it might be the case that midterms, an often large portion of a student's grade, occurs during this unit and contains a sizeable portion of material from Sets and Functions.

We found the total number of transactions was found to be negatively associated with final grades in the subsequent course. This is in contrast with prior work that showed that fewer interactions with the online learning system were associated with less learning (Rovai & Barnum, 2007). The system the course is implemented in, OLI, is intended for students to practice on low-stakes activities, not necessarily getting the questions right on the first attempt. However, if students read the accompanying instructional materials on the page, they should be able to answer the questions on the first try. This result of more transactions correlating to a lower grade could be attributed to guess-and-check behavior, where students omit reading the materials and attempt the questions until they achieve the correct answer. Attempting the problems in this system and many others is not technically discouraged, since they contain rich feedback that serves as an instructional moment. Unfortunately many students do not always read the feedback and believe they understand the content once the correct answer is achieved, even if it is by guessing.

Next, the evidence that the online discrete math primer helped students' performance in the subsequent course is only correlational. There are many factors that can come into play between the completion of the online course and conclusion of the follow-up one. However, these results demonstrate how online learning environments may make use of data they are already collecting, quiz scores and formative assessment answers, in a way that feed into a greater predictive system. Predictive modeling is a growing research area with many resulting systems suggesting interventions for at-risk students, based on the input data (Roblyer & Davis, 2008; Essa & Ayad, 2012). Such systems or similar methods could be integrated into the OLI platform, make use of this data, and provide interventions to the students that might fall into the at-risk category.

In sum, predicting future performance using student interaction data in an online course is a promising area of research, and should continue to be explored in the educational data mining literature. The insights gained will help improve student learning not only as measured by pre and post tests within the course, but will ensure that robust learning that prepares students for future learning opportunities is supported.

5 FUTURE WORK

As predictive modeling research continues and integrates with more systems, we hope to find trends across platforms that indicate a set of features that are continuously correlational. Future work in this area could also focus more on not only proving the effectiveness of the system for immediate learning, but for robust learning that transfers to later contexts were it is then prior knowledge. Looking at the transfer of this material from an online course context to an in-person one, like in this study, can help to indicate what makes online learning effective or not. With so many instructional materials and services online that claim to be effective, gauging the long term retention of what they teach is key to them truly being successful for learning. Additionally, future work in CS education should also consider courses in the curriculum that do not strictly rely on programming, such as this studies DMP course. Mathematical foundations are essential in certain aspects of programming and computational thinking, yet many transfer studies focus solely on programming contexts.

One limitation of the present work is that we did not have a measure of students' incoming mastery of the content of the DMP course. We are currently replicating the study with a new cohort of students, who took a short pretest at the beginning of the course, and the quizzes for each module included three items from the pretest to serve as a posttest. Analyses of pre and posttests will give a clearer window into what students learned from the online course, instead of simply measuring their performance on a test. We suggest future work in this area do the same, providing students with a concrete pretest and posttest to effectively evaluate their learning from the online materials. To further obtain a stronger causal evidence for its efficacy, a randomized controlled experiment, where one group of students completes the OLI course, whereas another completes a comparable activity of similar duration would be recommended.

REFERENCES

Baker, R., Gowda, S., & Corbett, A. (2011). Towards predicting future transfer of learning. In Artificial intelligence in education (pp. 23-30). Springer Berlin/Heidelberg.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In Learning analytics (pp. 61-75). Springer, New York, NY.

Beaubouef, T. (2002). Why computer science students need math. ACM SIGCSE Bulletin, 34(4), 57-59.

Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. Computers in Human Behavior, 73, 247-256.

Davies, J., & Graff, M. (2005). Performance in e-learning: online participation and student grades. British Journal of Educational Technology, 36(4), 657-663.

Dominguez, M., Bernacki, M. L., & Uesbeck, P. M. (2016). Using Learning Management System Data to Predict STEM Achievement: Implications for early warning systems. Paper presented at the Educational Data Mining Conference, Raleigh, NC.

Essa, A., & Ayad, H. (2012, April). Student success system: risk analytics and data visualization using ensembles of predictive models. In Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 158-161). ACM.

Hershkovitz, A., Baker, R., Gowda, S. M., & Corbett, A. T. (2013, July). Predicting future learning better using quantitative analysis of moment-by- moment learning. In Educational Data Mining 2013

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. Cognitive Science, 36(5), 757-798.

Lovett, M., Meyer, O., & Thille, C. (2008). JIME-The open learning initiative: Measuring the effectiveness of the OLI statistics course in accelerating student learning. Journal of Interactive Media in Education, 2008(1), 13-26

Reilly, Christine F., & Emmett Tomai. An examination of mathematics preparation for and progress through three introductory computer science courses.&quot; Frontiers in Education Conference (FIE), 2014 IEEE. IEEE, 2014.

Roblyer, M. D., & Davis, L. (2008). Predicting success for virtual school students: Putting research-based models into practice. Online Journal of Distance Learning Administration, 11(4).

Rovai, A. P., & Barnum, K. T. (2007). On-line course effectiveness: An analysis of student interactions and perceptions of learning. International Journal of E-Learning & Distance Education, 18(1), 57-73.