Predicting Post-Test Performance from Online Student Behavior: A High School MOOC Case Study

Sabina Tomkins[1], Arti Ramesh[2], Lise Getoor[1]
[1]University of California, Santa Cruz
[2]University of Maryland, College Park
satomkin@ucsc.edu, artir@cs.umd.edu, getoor@soe.ucsc.edu

ABSTRACT

With the success and proliferation of Massive Open Online Courses (MOOCs) for college curricula, there is demand for adapting this modern mode of education for high school courses. Online and open courses have the potential to fill a much needed gap in high school curricula, especially in fields such as computer science, where there is shortage of trained teachers nationwide. In this paper, we analyze student post-test performance to determine the success of a high school computer science MOOC. We empirically characterize student success by using students' performance on the Advanced Placement (AP) exam, which we treat as a post test. This post-test performance is more indicative of long-term learning than course performance, and allows us to model the extent to which students have internalized course material. Additionally, we analyze and compare the performance of a subset of students who received in-person coaching at their high school, to those students who took the course independently. This comparison provides better understanding of the role of a teacher in a student's learning. We build a predictive machine learning model, and use it to identify the key factors contributing to the success of online high school courses. Our analysis demonstrates that high schoolers can thrive in MOOCs.

Keywords

online education, high school MOOCs, student learning

1. INTRODUCTION

Massive Open Online Courses (MOOCs) have emerged as a powerful mode of instruction, enabling access around the world to high quality education. Particularly for college curricula, MOOCs have become a popular education platform, offering a variety of courses across many disciplines. Now open online education is being deployed to high schools worldwide, exposing students to vast amounts of content, and new methods of learning. Even as the popularity of high school MOOCs increases, their efficacy is debated [8]. One challenge is that the large amount of self direction MOOCs require may be lacking in the average high school student.

To understand the applicability of the MOOC model to high schoolers, we analyze student behavior in a year-long high school MOOC on Advanced Placement (AP) Computer Science. This course is distinguished from traditional college-level MOOCs in several ways. First it is a year-long course, while college MOOCs average 8-10 weeks in duration. This provides ample opportunity to mine student interactions for an extended period of time. Secondly, while traditional MOOCs have no student-instructor interaction, the high school MOOC that we consider incorporates instructor intervention in the form of coaching and online forum instructor responses. Evaluating the effectiveness of this hybrid model allows us to investigate the effect of human instruction on high school students, a group which may particularly benefit from supervision.

Finally, we introduce a post test as a comprehensive assessment occurring after the termination of the course. A valid post test should assess students' knowledge on critical course concepts, such that students' course mastery is reflected in their post-test score. We treat the Advanced Placement (AP) exam as a post test and consider students' performance on this test as being indicative of long term learning. Previous MOOC research evaluates students on course performance [4]. While course performance can be a good metric for evaluating student learning in the short term, post-test performance is a more informative metric for evaluating long-term mastery.

We propose and address the following research questions, aimed at evaluating the success of MOOCs at the high school level.

1. Can high school students learn from a MOOC, as evidenced here by their post-test (AP exam) performance?
2. How does coaching help students achieve better course performance and learning?
3. How can we predict student's post test performance from course performance, forum data, and learning environment?

Our contributions in this paper are as follows:

1. We perform an in-depth analysis of student participation and performance to evaluate the success of MOOCs at the high school level. To do so, we identify two course success measures: 1) course performance scores, and 2) post-test performance scores.

2. We evaluate the effect of two important elements of this high school MOOC: discussion forums and coaching, on student performance.
3. We use a machine learning model to predict student post test scores. First constructing features drawn from our analysis of student activities, then determining the relative predictive power of these features. We show that this process can be used to draw useful insights about student learning.

2. RELATED WORK

Research on online student engagement and learning, is extensive and still growing Kizilcec et al. [5], Anderson et al. [1], and Ramesh et al. [11] develop models for understanding student engagement in online courses. Tucker et al. [13] mine text data in forums and examine their effects on student performance and learning outcomes. Vigentini and Clayphan [14] analyze the effects of course design and teaching effect on students' pace through online courses. They conclude that both the course design and the mode of teaching influence the way in which students progress through and complete the course. Simon et al. [12] analyze the impact of peer instruction in student learning.

Particularly relevant to our findings is the impact of gaming the system on long-term learning. Baker et al. [2] investigate the effect of students gaming an intelligent tutor system on post-test performance. In the high school MOOC setting, we observe a similar behavior in some students achieving high course performance, but low post-test performance. We identify plausible ways in which these students can be gaming the system to achieve high course performance and present analysis that is potentially useful for MOOC designers to prevent this behavior.

There is limited work on analyzing student behavior in high school MOOCs. Kurhila and Vihavainen [6] analyze Finnish high school students' behavior in a computer science MOOC to understand whether MOOCs can be used to supplement traditional classroom education. Najafi et al. [9] perform a study on 29 participating students by splitting them into two groups: one group participating only in the MOOC, and another group is a blended-MOOC that has some instructor interactions in addition to the MOOC. The report that students in the blended group showed more persistence in the course, but there was no statistically significant difference between the groups' performance in a post-test. In our work, we focus on empirically analyzing different elements of a high school MOOC that contribute to student learning in an online setting. We use post-test scores to capture student learning in the course and examine the interaction of different modes of course participation with post-test performance. Our analysis reveals course design insights which are helpful to MOOC educators.

3. DATA

This data is from a two-semester high school Computer Science MOOC, offered by a for-profit education company. The course prepares students for College Board's Advanced Placement Computer Science A exam and is equivalent to a semester long college introductory course on computer science. In this work, we consider data from the 2014-2015 school year for which 5692 students were enrolled. The course is structured by terms, units, and lessons. Lessons provide instruction on a single topic, and consist of video lectures and activities. The lessons progress in difficulty beginning with printing output in Java, and ending with designing algorithms. Each lesson is accompanied with activities. These activities are not graded, instead students receive credit for attempting them. Students take *assessments* in three forms: assignments, quizzes, and exams, each released every two weeks.

At the end of the year students take an Advanced Placement (AP) exam. Students can use their AP exam performance exam as a substitution for a single introductory college course. The AP exam score ranges from 1 to 5. In all, we have data for 1613 students who take the AP exam. This number is a lower limit on the total number of students who may have taken the course and the AP. The course provides a forum service for students, which is staffed with paid course instructors. Approximately, 30% of all students who created course accounts also created forum accounts, 1728 students in all.

This course is unique in that it provides a coach service which high schools can purchase. This option requires that the school appoint a coach, who is responsible for overseeing the students at their school. The coach is provided with additional offline resources, and has access to a forum exclusive to coaches and course instructors. The average classroom size is approximately 9 students with a standard deviation of approximately 12 students. The largest classroom size coached by a single coach is 72, while some coaches supervise a single student. Of all students who have enrolled in the course, approximately, 23% (1290) are coached and 77% (4402) are independent. From here on we refer to the students enrolled with a coach as *coached students*.

We summarize the class statistics in Figure 1 below. The majority of coached students sign up for the student forum, and many persist with the course to take the final AP exam at the end of the year.

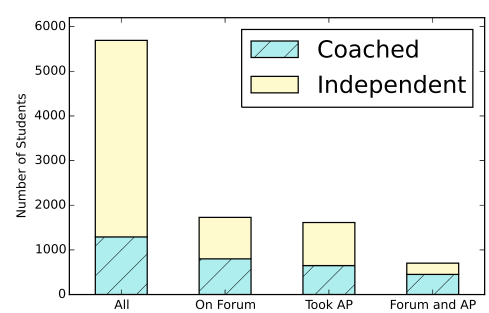


Figure 1: Student participation varies between coached and independent students.

4. EMPIRICALLY CHARACTERIZING SUCCESS OF A HIGH-SCHOOL MOOC

In this section, we use post-test performance and course performance to question the success of MOOCs for high school students. With an empirical analysis, we provide insights on how to adapt high school MOOCs to benefit different groups of students. To investigate this question, we focus on the subset of students for whom we have post-test data. To evaluate student success in the course, we identify three measures of course participation in MOOCs that are relevant to the high school population: *overall score*, *course completion*, and *post-test score*.

Overall Score The overall score captures the combined score across course assignments, quizzes, exams, and activities, each of which contributes to the final score with some weight. We maintain the same weights as those assigned by the course, exams are weighted most heavily, activities the least.

Overall Score = .3*(Assignment Score + Quiz Score) + .6*Exam Score + .1*Activity Score.

Course Completion The second success measure we use is course completion. Course completion measures the total number of course activities and assessments completed by the student.

$$\textit{Course Completion} = \frac{\text{Total Activities and Assessments Attempted}}{\text{Total Number of Activities and Assessments}}$$

Post-Test Score This score captures student scores in the post test that is conducted 2 weeks after the end of the course. The score ranges from 1 to 5. This score captures the advance placement (AP) score, hence we also refer to it as the AP score.

To evaluate the effectiveness of the high school MOOC on student performance, we first examine the relationship between course completion and course performance. We hypothesize that as students complete a higher percentage of the course, they should do better in the course assessments leading to higher course performance scores and post-test scores. Examining the correlation of course completion to post-test performance, we find that they are positively correlated. This suggests that the course indeed helps students in achieving good performance in the assessments. However, we find that of the students that achieve an overall score of 90 or greater, only 70% pass the post test. Similarly, of the students who complete 90% of the course, only 63% pass the post test. These initial observations indicate the need to perform a more detailed study in order to understand the different student populations in the course.

Next, we examine the relationship between overall score and post-test score, captured in Figure 2. From this plot, we see a positive linear relationship between course performance and post-test score. Notably, we observe that the average post-test score of the students who achieve a 90% or higher in the course is above a 4.0, and well above a passing score.

Students regularly complete three kinds of assessments: assignments, quizzes, and exams. Assignments are programming exercises, testing students' coding abilities. Programming assignments are submitted online through an interface capable of compiling programs and displaying error messages. Quizzes are multiple choice assessments on course material, with an emphasis on recently covered topics. Exams have a similar format to quizzes but are slightly longer. Both quizzes and exams are timed and students cannot change

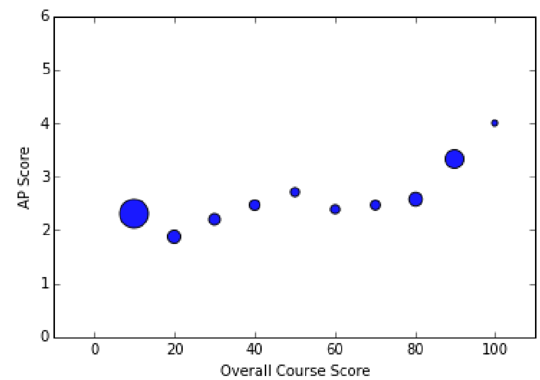


Figure 2: The dot sizes are proportional to the number of students achieving the overall score.

their answers once they submit them. In all, there are 15 assignments, 8 quizzes and 6 exams in the course. We will refer to them as $A_{1:15}$, $Q_{1:8}$, and $E_{1:6}$, in the discussion below.

In Figure 4, we present results of student performance across assessments. Figures 4(a), 4(b), and 4(c) present average student assignment, quiz, and exam scores for students who passed/failed the post test, respectively. We find that students who pass the post test do better on assessments. We also observe that the scores across all assessments show a decreasing trend as the course progresses. This signals that the assessments get harder for both groups of students as the course progresses. Another important observation is the increase in scores for both groups at assignment 8, quiz 5, and exam 4; these assessments are at the start of the second term in the course, indicating that students may have higher motivation at the start of a term.

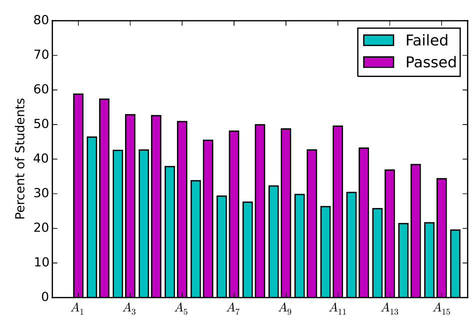
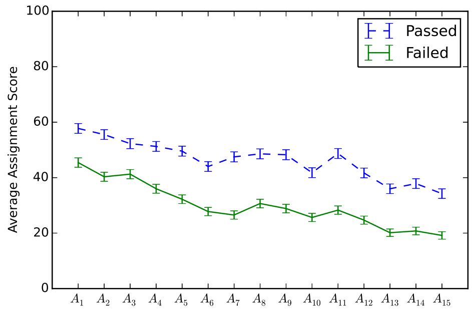


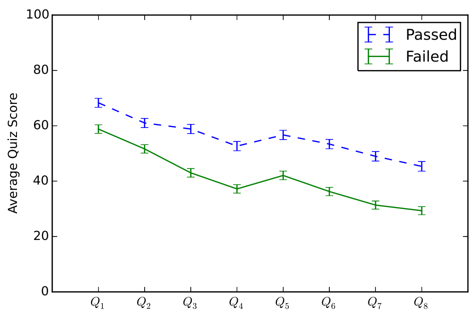
Figure 3: Students who pass are more likely to attempt assignments than students who fail.

Additionally, some assessments show a greater difference between the two groups of students, and performance on these assessments are more informative of student learning. In Figure 4(c), we observe that for both passed and failed students, we see the greatest dip in performance in the final exam. As the final exam is the most comprehensive exam, and possibly most related to the post test, analyzing why students do so poorly on this exam is a worthwhile direction of study in its own right.

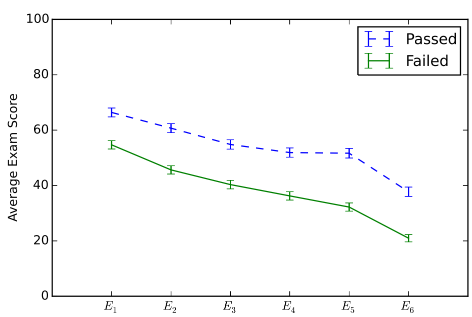
Another important dimension is considering assignment com-



(a) Average assignment scores of passed and failed students



(b) Average quiz scores of passed and failed students



(c) Average exam scores of passed and failed students

Figure 4: Passed students have higher average scores across all assessments than failed students.

pletion rate of these two groups of students. In Figure 3, we examine the relationship between attempting assignments and course performance and find that students passing the post test also attempt more assignments. This implies that the high scores of these students are not only the product of strong prior knowledge, but are also the result of learning from the course.

5. FORUM PARTICIPATION AND POST-TEST PERFORMANCE

In this section, we analyze forum participation of students and examine it's effect on course success. To do so, we answer the following questions:

- Does participation in forums impact post-test performance and learning?
- What are the key differences between participation styles of students who pass the course and students who do not?

We first look at the average score of students who use the forum compared to the average score of students who do not use the forum. Students who use the forum have a statistically higher post test performance score of **2.77**, whereas students who do not use the forum obtain a score of **2.34**, ($p < .001$). It is not clear if the forum impacts learning, or if instead, students with a high desire to learn are more likely to use the forum.

To accurately evaluate forum participation of the two subpopulations, we analyze them on different types of forum participation. Forum participation comprises of different types of student interactions: asking questions, answering other student questions, viewing posts, and contributing to conversation threads. Table 1 gives the comparison of students who pass the post test against student who do not across the various forum participation types. The different types of forum participation types are referred to as: Questions, Answers, Post Views, and Contributions. We also consider the number of days that a student was logged into the forum, which is denoted by Days Online.

On average, students who pass the course make more contributions than students failing in the course. They also answer more questions. Both groups seem to spend roughly the same amount of time online, to view the same number of posts, and to ask the same number of questions. What most distinguishes a student who passes, from one who fails is whether they are answering questions and contributing to conversations.

Forum Behavior	Failed Mean	Passed Mean	Failed Median	Passed Median
Questions	3	4	0	1
Answers	1	4	0	0
Post Views	147	140	73	62
Contributions	9	16	1	2
Days Online	19	21	11	13

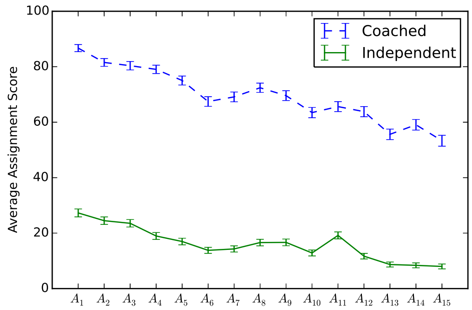
Table 1: The average forum participation is significantly more for students that pass the course. The behavior for which there was a statistical significance difference between the groups are highlighted in bold.

This analysis further demonstrates the importance of forums to MOOCs. Answering questions and contributing to conversations are two behaviors indicative of strong post-test performance. We hope that MOOC designers can use this information to create appropriate intervention and incentive strategies for students.

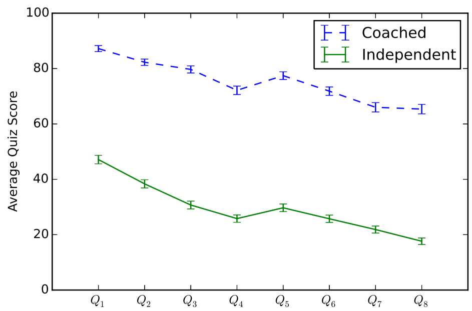
6. COACHING

In this section, we evaluate the effect of coaching on student learning. We compare coached students to independent students using their participation in course assessments and forums. We conclude this section by looking at the subset of students who have only one coach, in order to isolate the effect of coaching from other classroom effects.

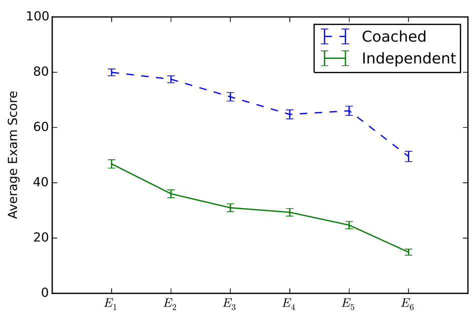
6.1 Course Behavior



(a) Average assignment scores of coached and independent students



(b) Average quiz scores of coached and independent students



(c) Average exam scores of coached and independent students

Figure 5: Coached students have higher average scores than independent students.

We inspect the average assessment scores of coached and independent students in Figure 5. Observing scores across assignments, quizzes, and exams in Figures 5(a), 5(b), and 5(c), respectively, we find that coached students perform better than independent students across all assessments.

Such differentially high performance in the course should indicate higher performance in the AP exam for coached students. However, we see that coached students fail to get a high post-test score. The average post-test score for a coached student is 2.43, while it is 2.59 for an independent student. We test statistical significance using a t-test with a rejection threshold of $p < 0.05$. In Section 6.2, we analyze forum participation of students to understand this difference in scores.

6.2 Forum Participation of Coached and Independent Students

Analyzing forum participation of coached and independent students, we find that there is a significant difference in forum participation between coached and independent students. Table 2 gives the comparison between coached and independent students in forum participation. On average, coached students ask more questions and answer fewer questions on the forums when compared to independent students. Coached students exhibit more passive behavior by predominantly viewing posts rather than writing posts, when compared to independent students. This can be particularly dangerous if the posts which are viewed contain assignment code.

Forum Behavior	Coached Mean	Independent Mean
Questions	2.81	1.90
Answers	1.45	1.72
Post Views	145.49	81.50
Contributions	8.10	7.33
Days Online	20.64	12.55

Table 2: Coached students view more posts and ask more questions. The behavior for which there was a statistical significance difference between the groups are highlighted in bold.

In Table 3, we compare coached students who pass to coached students who fail and see the same differences as those observed between all students who pass, and all students who fail. Students who pass are more likely to answer questions, and contribute to conversations.

Forum Behavior	Passed Mean Coached	Failed Mean Coached
Questions	3.97	2.87
Answers	3.04	0.56
Post Views	141.56	164.14
Contributions	14.19	5.93
Days Online	22.71	21.53

Table 3: The differences in forum behavior between coached students who pass and who fail follow the same trends in forum behavior exhibited by the general population, and shown in Section 5. The behavioral features for which there was a statistical significance difference between the groups are highlighted in bold.

6.3 Coaches with Only One Student

To examine the effect of coaching class size on coached students' post-test performance, we examine coached students in a classroom size of one. Comparing average post test scores of coached students who are singly advised by their coaches (classroom size of one) with independent students, we find that the average post-test score for the coached students is 3.6, while it is 3.2 for independent students. We hypothesize that the lower score of coached students in classroom size greater than one is due to the possibility of sharing answers when students study together. This explains their high overall score but lower post-test scores. This analysis further suggests that the effect of coaching is confounded by the effects of learning in a classroom with peers. To fully understand the effect of a coach guiding a student through the learning process, the peer-effects of classmates should be better understood and isolated. In Section 7, we take first steps in this direction by proposing student types.

7. INSPECTING UNEXPECTED STUDENT TYPES

In this section, we identify and analyze various types of students in the course based on their performance in the assessments. We classify students into two broad types based on whether the overall scores and post-test scores are correlated. Figure 6 gives the relationship between overall score and post test score for all students. Two groups of students emerge, students who exhibit a correlation between overall scores and post test scores, and students who do not. These two groups can be further broken down based on whether they obtain a high score on the post test, yielding four groups of students.

- *Low learners*: These students have low values for both overall scores and post test scores.
- *High learners*: These student obtain high values for both overall scores and post test scores.
- *Unexpected low learners*: These students obtain high overall scores, but low post test scores.
- *Unexpected high learners*: These students obtain high post test scores, but low overall scores.

Among these, the unexpected low learners and unexpected high learners deviate from the rest of the students. To analyze these two groups, we delve deeper into other aspects of the course such as forum participation and coaching.

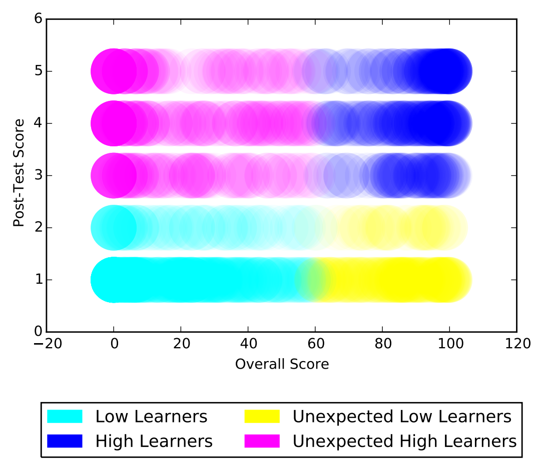


Figure 6: Four groups of students emerge: low learners, high learners, unexpected low and high learners. For high course performance we choose a threshold of 60% as a passing grade.

7.1 Unexpected Low Learners

Unexpected low learners are those students who perform well on the course assessments (with an overall score of over 60%) but who do not earn a passing post-test score. We hypothesize that this might be due to their not retaining information from the course, or not arriving at high overall course scores on their own. To understand their low post-test performance, we examine their forum behavior and coaching environment.

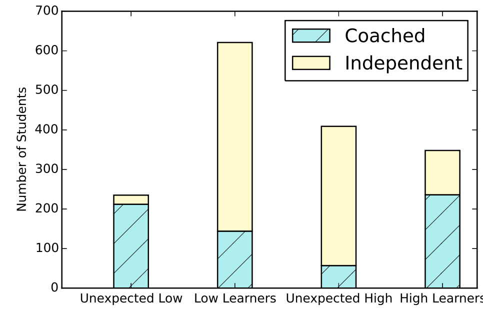


Figure 7: The majority of unexpected low learners are coached, while the majority of unexpected high learners are independent.

As can be seen in Figure 7, approximately 91% of unexpected low learners are coached students. Most of these students are part of large classrooms coached by the same coach, increasing the possibility of getting answers from their peers/coach. Plagiarism is a significant challenge in online courses as proctoring students online is not as efficient as in classroom courses.

Further, analyzing forum performance, we find that approximately 76% of unexpected low learners use the forum. Of those who use the forum, 91% are coached. Table 4 gives the forum participation of coached and independent unexpected low learners. The forum participation of these students have a strong similarity to failing students in Table 1, participating passively in the course by viewing forum posts and contributing to less answers. The coached students are less active than the independent students on the forum in every way, even in post views. While it was posited before that active forum participation is indicative of learning and high AP exam performance, this may not be the case in all groups. For example, the small number of independent students may be using the forum for social, rather than learning purposes.

Forum Behavior	Coached Mean	Independent Mean
Questions	3.5	9.2
Answers	0.5	15.0
Post views	195.0	293.0
Contributions	7.1	67.0
Days Online	25.6	35.2

Table 4: Forum behaviors for which there is a statistical significance between groups are highlighted in bold.

7.2 Unexpected High Learners

Unexpected high performers earn an overall course score of less than 60% but pass the AP exam with a 3 or above. Approximately 86% (357 out of 409) of unexpected high leaners are independent and approximately 80% of the unexpected high learners (323 out of 409) are not on the forums. That this group can do so well on the post test, without either a high amount of course or forum participation strongly suggests that either these students have prior knowledge in computer science or that they are not being primarily exposed to computer science through this course but are instead using it to supplement another mode of instruction. A pre test of students' prior computer science knowledge would provide further clarity.

8. PREDICTING PERFORMANCE FROM STUDENT BEHAVIOR

In Sections 4 and 5, we see that students' post-test performance is affected by their course and forum behavior. We construct features with which to model these different characteristics of student behavior. These student models are then used to predict post-test scores. By discovering the relative rank of the student model features, we draw insights about student behavior relevant to learning, and to course design.

8.1 Student Model Features

We group the course features from student interactions into four broad categories: 1) course behavior, 2) forum behavior, 3) coaching environment, and 4) topic analysis of forum posts. We extract features from student course behavior and forum behavior, which we describe in Sections 4 and 5. The two other feature categories are described below.

8.1.1 Coaching Environment

Students in the online course are either coached or independent. Coaches are provided a separate discussion forum, apart from the student forum, where they can interact with other coaches and instructors of the course. We extract features that capture coaches' prior knowledge and their involvement in guiding students. Table 5 gives the list of coaching related features extracted from the discussion forum for coaches.

Feature	Explanation
Coached	Boolean feature capturing whether a student is coached or independent
Coach Views	# posts viewed by the coach
Coach Questions	# questions posted by the coach
Coach Answers	# answers posted by the coach
Coach Contributions	# contributions in the forum

Table 5: Coaching related features

8.1.2 Posts Topic Distribution

For extracting topics of the post, we explore the topic modeling framework using Latent Dirichlet Allocation (LDA) [3]. Before using LDA we clean the text data by removing stop words, stemming certain words, and removing all common course words, such as code. To obtain the topic distribution of posts, we use the Machine Learning for Language Toolkit (MALLET) [7]. We use the following parameters for the topic model: number of topics = 150, and optimize-interval = 100, where the hyper-parameters required by LDA, α and β, are set to the default values.

8.2 Predictive Model

We incorporate extracted features in a linear kernel Support Vector Machines (SVM) model, using the python package Scikit-learn [10]. Comparing this model with other machine learning algorithms such as logistic regression, decision trees, and Naive Bayes we found the results to be comparable. We filter our student pool to those who participated in the forums and took the post test (approximately 16% of all students who completed the post test). A subset of features that are predictive of post-test performance were selected using recursive feature elimination in Scikit-learn [10]. Recursive feature elimination works by training a classifier which weighs features and then trims all features with the lowest weights; this trimming allowed us to obtain the best predictions, and to understand which features are most predictive of student success.

8.3 Empirical Results

In this section, we present empirical results using the SVM model defined above to predict post-test performance. To evaluate the effectiveness of this model we compute the F-measure, which is the harmonic mean of precision and recall. F-measure is an optimal metric for a setting with unbalanced classes such as ours, where accuracy may appear to be deceptively high if a classifier reliably predicts the majority class. Our model gives an F-measure of 0.81 for predicting post-test performance. We validate our results with 10-fold cross validation. In the next sections, we analyze the attributes of student behavior which are most predictive of performance.

8.3.1 Topics and Performance

The topics discovered by the topic model fall into four broad categories: help requests, assignments, course material, and course activities. In Table 6, we present the ten topics which are most predictive of post-test performance. The first three topics in the table fall into the help requests category. They include words such as *trouble*, *help*, and *fail*. Four of the top ten topics correspond to assignments, with top words which are descriptive of assignments from the course. For example, in assignment A_4 students are asked to write a program to count the number of hashtags, links, and attributions in a tweet, and in the topic associated with this assignment we see the words: *hashtag*, *tweet*, *attributions*, *mentions*, and *links*. Two topics represent the concepts discussed in the course: object oriented programming, and hash maps. The hash maps topic is particularly interesting as hash maps are not introduced in the course, but students still use them in their projects, and discuss them on the forum. The other prominent topics are topics related to course activities. For example, the activity topic in the the table is an activity given to students to print the location of a vehicle. This is the most elaborate activity that students undertake in the course, hence it appears in the top predictive topics for predicting post-test performance.

Topic Label	Top Words
Help requests	trouble, don't, perfectly, won, updated
Help requests	hope, helps, change, find
Help requests	fail, expected, updated, supposed
Assignment content (A_4)	hashtag, tweet, attributions, mentions, links
Lecture (hashmaps)	Map, key, Getvalue, Hashmap, entry
Course Activity	vehicle, location, backward, forward, GetLocation
Assignment content (A_6)	ArrayList, words, remove, equals, size
Assignment content (A_{10})	strand, size, TurnOn, green, BurntOut
Assignment content (A_{14})	sort, insertion, swap, insert, algorithm
Lecture (OOP and Methods)	object, constructor, methods, parameter, returns

Table 6: Top predictive topics and the words in these topics

Figure 8 gives the distribution of passed and failed students across the different ten most predictive topics given in Table 6. We observe that passing students post about the course activity on vehicles more than failing students. Since activities only contribute to a small portion of their grade,

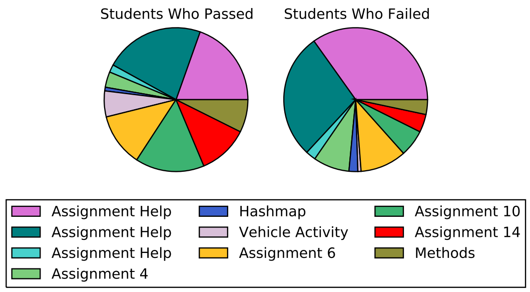


Figure 8: Students who pass post about different topics than students who fail.

participation in activities is a good measure for students' level of motivation and learning.

Additionally, we observe that failing students are far more likely to write posts which fall in the help category. Looking at some of the posts in this category, we find that these posts are often short and use help words, but do not contain detailed information about the specific assignment problem in question. This finding suggests that analyzing the posts for linguistic cues is helpful in understanding students' motivation.

The third important take away from this analysis is that this topic distribution can help discover patterns in student behavior. For example, passing students post about assignment A_{10} more than failing students. But, failing students post more about assignment A_4. As assignments tend to get harder as the course progresses, the difference in behavior can be attributed to failing students needing help on the easier assignments, while the savvier students focus on the harder assignments.

8.3.2 Critical Assessments

Here, we describe the most predictive assignments, quizzes and exams that we use in the predictive model. We find that assignments A_4, A_8, A_9, and A_{10} are the most predictive assignments. These assignments are on core concepts and hence may be the most critical assignments in the course. This observation is bolstered by the fact that these assignments are referenced in the forums more than other assignments. Two of these assignments feature in the top ten predictive topics given in Table 6. Pinpointing the moment when a student needs help is not only predictive of their success, but also critical in maintaining engagement and understanding. Understanding which assignments are discussed more in the forums can reveal important information for initiating instructor interventions.

9. CONCLUSION

From this analysis we conclude that MOOCs are a viable option for high school students. Forty-seven percent of students who took the post test passed it. Four hundred and sixty four of these students were to the best of our knowledge self-directed. While we can say that MOOCs work for some high school students, the particularities of this group must be understood. It is not clear, for example, how the students who achieve high course scores, but low AP exam scores are able to do so. Are they receiving answers from other students, or have they truly mastered the course content, but lack the ability to demonstrate this mastery on a test? High school MOOC students are a unique group with particular modeling demands.

We have developed models of these students, characterizing high and low learners by their course and forum behavior, as well as by the topics that they post about. These models have allowed us to differentiate the behavior of students who pass from that of students who fail. In this case study post-test performance was correlated with course-performance, such that students who earned a high course score also earned a high post-test score. Students who performed well on the post test were more likely to contribute to conversations, and to answer questions on the student forum. They were also more likely to post about ungraded activities, and less likely to write posts asking for help. Coached students were more likely to perform well in the course, and spent more time on the forum. Understanding the differences between students who excel and those who do not is crucial in developing the courses that students, and particularly high school students need.

References

[1] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. Engaging with massive online courses. In *International Conference on World Wide Web (WWW)*, 2014.

[2] Ryan Baker, Albert Corbett, Kenneth Koedinger, and Angela Wagner. 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*, 2004.

[3] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3: 993–1022, 2003.

[4] Gregor Kennedy, Carleton Coffrin, Paula de Barba, and Linda Corrin. Predicting success: How learners' prior knowledge, skills and activities predict MOOC performance. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK)*, 2015.

[5] René F. Kizilcec, Chris Piech, and Emily Schneider. Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Conference on Learning Analytics and Knowledge (LAK)*, 2013.

[6] Jaakko Kurhila and Arto Vihavainen. A purposeful MOOC to alleviate insufficient cs education in finnish schools. *Transactions of Computing Education*, 15(2):10:1–10:18, 2015.

[7] Andrew McCallum. MALLET: A machine learning for language toolkit. 2002.

[8] Bock Mike and O'Dea Victoria. Virtual educators critique value of MOOCs for K-12. 2013.

[9] Hedieh Najafi, Rosemary Evans, and Christopher Federico. MOOC integration into secondary school courses. *The International Review of Research in Open and Distributed Learning*, 15(5), 2014.

[10] Fabian Pedregosa et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[11] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume III, and Lise Getoor. Learning latent engagement patterns of students in online courses. In *AAAI Conference on Artificial Intelligence*, 2014.

[12] Beth Simon, Julian Parris, and Jaime Spacco. How we teach impacts student learning: Peer instruction vs. lecture in cs0. In *ACM Technical Symposium on Computer Science Education (SIGCSE)*, 2013.

[13] Conrad Tucker, Barton K. Pursel, and Anna Divinsky. Mining student-generated textual data in MOOCs and quantifying their effects on student performance and learning outcomes. In *Conference of American Society for Engineering Education*, 2014.

[14] Lorenzo Vigentini and Andrew Clayphan. Pacing through MOOCs: course design or teaching effect? In *Conference on Educational Data Mining (EDM)*, 2015.