Learning Latent Engagement Patterns of Students in Online Courses

[1]Arti Ramesh, [1]Dan Goldwasser, [1]Bert Huang, [1]Hal Daumé III, [2]Lise Getoor
[1]University of Maryland, College Park [2]University of California, Santa Cruz
{artir, bert, hal}@cs.umd.edu, goldwas1@umiacs.umd.edu, getoor@soe.ucsc.edu

Abstract

Maintaining and cultivating student engagement is critical for learning. Understanding factors affecting student engagement will help in designing better courses and improving student retention. The large number of participants in massive open online courses (MOOCs) and data collected from their interaction with the MOOC open up avenues for studying student engagement at scale. In this work, we develop a framework for modeling and understanding student engagement in online courses based on student behavioral cues. Our first contribution is the abstraction of student engagement types using latent representations. We use that abstraction in a probabilistic model to connect student behavior with course completion. We demonstrate that the latent formulation for engagement helps in predicting student survival across three MOOCs. Next, in order to initiate better instructor interventions, we need to be able to predict *student survival* early in the course. We demonstrate that we can predict student survival early in the course reliably using the *latent* model. Finally, we perform a closer quantitative analysis of user interaction with the MOOC and identify student activities that are good indicators for survival at different points in the course.

Introduction

The large number of students participating in MOOCs provides the opportunity to perform rich analysis of large-scale online interaction and behavioral data. This analysis can help improve student engagement in MOOCs by identifying patterns, suggesting new feedback mechanisms, and guiding instructor interventions. Additionally, insights gained by analyzing online student engagement can also help validate and refine our understanding of engagement in traditional classrooms.

In this paper, we study the different aspects of online student behavior in MOOCs, develop a large-scale, data-driven approach for modeling student engagement, and show that jointly measuring different aspects of student behavior early in the course can provide a strong indication of course completion. We demonstrate the construction of a holistic model incorporating content (e.g., language), structure (e.g., social interactions), and outcome data.

Predictive modeling over MOOC data poses a significant technical challenge as it requires the ability to combine language analysis of forum posts with graph analysis over very large networks of entities (students, instructors, assignments, etc.). To address this challenge, we use *probabilistic soft logic* (PSL) (Broecheler, Mihalkova, and Getoor 2010), a framework that provides an easy means to represent and combine behavioral, linguistic, and structural features in a concise manner. We analyze students' online behavior to identify how they engage with course materials and investigate how engagement can be helpful in predicting successful completion of the course. Early detection of changes in student engagement can help educators design interventions and adapt the course presentation to motivate students to continue with the course (Brusilovsky and Millán 2007). Our work is a step toward helping educators understand how students interact with MOOCs.

Our second contribution is providing a data-driven formulation that captures student engagement in the MOOC setting. As in the traditional classroom setting, assessing online student engagement requires interpretation of indirect cues. Identifying these cues in an electronic setting is challenging, but the large amounts of available data can offset the loss of in-person communication. We model engagement using *latent* variables, which take into account the observed behaviors of online students and their resulting survival in the class. Uncovering this latent information provides a better explanation of students' behavior leading to course completion.

Examining real MOOC data, we observe that there are several indicators useful for gauging students' engagement, such as viewing course content, interacting with other learners or staff on the discussion forums, and the topic and tone of these interactions. Furthermore, students often engage in different aspects of the course throughout its duration. For example, some students engage in the social aspects of the online community—by posting in forums and asking and answering questions—while others only watch lectures and take quizzes without interacting with the community. We take these differences into account and propose a model that uses the different behavioral aspects to distinguish between forms of engagement: passive or active. We use these engagement types to predict student survival and reason about their behavior over time.

We apply our model to real data collected from several Coursera[1] courses and empirically show its ability to capture behavioral patterns of students and predict course completion. Our experiments validate the importance of providing a holistic view of students' activities, combining all aspects of online behavior, in order to accurately predict the students' motivation and ability to complete the class. We show that our model is able to make meaningful class completion predictions using data obtained at an *early* stage in the class. These predictions can help provide a basis for instructor intervention at an early stage in a course, helping to improve student retention rates.

Related Work

Prior work (Kuh 2003; Carini, Kuh, and Klein 2006) has studied the relationship between student engagement and academic performance for traditional classroom courses; they identify several metrics for user engagement (such as student-faculty interaction, level of academic challenge). Carini et al. (2006) demonstrate quantitatively that though most engagement metrics are positively correlated to performance, the relationships in many cases can be weak. Our work borrows ideas from Kuh (2003), Carini, Kuh, and Klein (2006), and from statistical survival models (Richards 2012) and adapts these to the MOOC setting.

Various works analyze student dropouts in MOOCs (Kotsiantis, Pierrakeas, and Pintelas 2003; Clow 2013; Balakrishnan 2013; Yang et al. 2013). Our work differs from these in that we analyze a combination of several factors that contribute to student engagement and hence their survival in online courses. We argue that analyzing the ways in which students engage themselves in different phases of online courses can reveal information about factors that lead to their continuous survival. This will pave the way for constructing better quality MOOCs, which will then result in increase in enrollment and *student survival*. In this work, we analyze the different course-related activities and reason about important factors in determining student survival at different points in the course.

Student engagement is known to be a significant factor in success of student learning (Kuh 2003), but there is still limited work studying student engagement in MOOCs. Our work is closest to that of Kizilcec, Piech, and Schneider (2013) and Anderson et al. (2014), who attempt to understand student engagement using completely unsupervised techniques (clustering). Our work differs from the above work in that we view types of engagement as latent variables and learn to differentiate among the engagement types from data. We use quiz-submission as a measure of *student survival* and use the student survival scores to train the model. Hence, our latent engagement variables are specifically constructed for predicting *student survival*. We then use this model to predict whether the student submitted the final exam/assignments/quizzes in the course, i.e., whether the student *survived* the course. We model engagement explicitly and demonstrate that it helps in predicting student survival.

Modeling Student Survival

As students interact on a MOOC, detailed records are generated, including page and video views, forum visits, forum interactions such as voting, posting messages and replies, and graded elements such as quizzes and assignments. In this section, we describe how we model student survival, connecting it to the various behavioral, linguistic features of these student interactions.

To model the interactions between these features and student survival, we use *probabilistic soft logic* (PSL), a system for relational probabilistic modeling. PSL enables us to encode our observed features and (latent and target) variables as logical predicates and design models by writing rules over these predicates. PSL interprets these rules in a parameterized probability model and is able to perform efficient inference and parameter fitting using machine learning algorithms. The expressiveness and flexibility of PSL allows us to easily build different models for MOOC data, and we exploit this by comparing a model that represents multiple forms of latent engagement against a simpler model that directly relates the observable features to student survival.

Probabilistic Soft Logic

We briefly overview the some technical details behind PSL. For brevity, we omit many specifics, and we refer the reader to (Broecheler, Mihalkova, and Getoor 2010; Bach et al. 2013) for more details. PSL is a framework for collective, probabilistic reasoning in relational domains, which uses syntax based on first-order logic as a templating language for continuous graphical models over random variables representing soft truth values. Like other statistical relational learning methods (Getoor and Taskar 2007), PSL uses weighted rules to model dependencies in a domain. However, one distinguishing aspect is that PSL uses continuous variables to represent truth values, relaxing Boolean truth values to the interval [0,1]. Triangular norms, which are continuous relaxations of logical connectives AND and OR, are used to combine the atoms in the first-order clauses. The underlying probabilistic model defined by PSL's semantics is a *hinge-loss Markov random field* (HL-MRF) (Bach et al. 2013).

Inference of the most probable explanation in HL-MRFs is a convex optimization problem, which makes working with PSL very efficient in comparison to many relational modeling tools that use discrete representations. HL-MRFs admit various learning algorithms for fully-supervised training data and are amenable to expectation maximization using point estimates for partially-supervised data with latent variables. In our model, we use this capability to represent student engagement as latent variables.

Modeling MOOC Student Activity

The MOOC online environment mainly consists of two resources: video lectures and forums. Students can watch lec-

[1]https://www.coursera.org

tures multiple times and respond to on-demand quizzes during the lectures[2]. Students can interact by asking and responding to questions in the forums. There are typically multiple forums organized by topics, each consisting of multiple threads, and each thread consisting of multiple posts. Students can respond, vote (up or down) on existing posts and subscribe for updates to forums threads. Each student is given a reputation score based on the votes on posts created by the student. These activities are depicted in Figure 1.

We quantify these activities by defining a set of PSL predicates, which are used to create features. We categorize these predicates as either behavioral, interaction-based, or temporal, and describe them in the following sections.

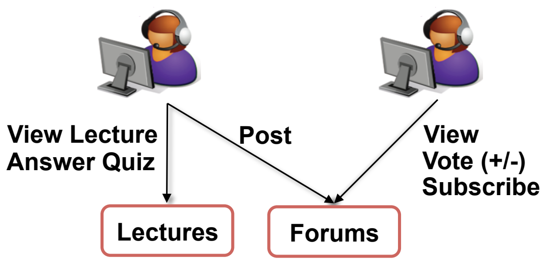


Figure 1: Structure of MOOC student activity.

Behavioral Features Behavioral features are attributes that the student exhibits while on the MOOC website. We consider two types of behavioral features: aggregate and non-aggregate.

Aggregate features describe the student's behavior, relative to others. The predicates *postActivity*(USER), *voteActivity*(USER) and *viewActivity*(USER) capture user activity in the forums. The student's reputation is captured using *reputation*(USER). These are calculated for each user by assessing if the value of the feature is more than the median value considering all users. The aggregate predicates take Boolean values.

Non-aggregate features directly quantify student's behavior. The predicates *posts*(USER, POST) and *votes*(USER, POST) capture an instance-level log of users posting and voting on the discussion forums. The predicates *posts* and *votes* are true if the USER posts or votes on POST. Predicate *upvote*(POST) is true if the post has positive votes and false otherwise, and predicate *downvote*(POST) is true if a post has been down-voted.

Forum Content and Interaction Features MOOC forums are rich with relevant information, indicative of the students' attitudes toward the course and its materials as well as the social interactions between students. We capture this information using two types of features, *linguistic* features capturing the sentiment of the post content, and *structural* features capturing the forum structure, organized topically into threads and forums types.

[2]These quizzes are generally not used to calculate the final evaluation.

The attitudes expressed by students on the forums can be captured by estimating sentiment polarity (positive or negative) and identifying subjective posts. Since MOOC forums contain thousands of posts, we use an automated tool, *OpinionFinder* (Wilson et al. 2005) to avoid manual annotation. The tool segments the forums posts into sentences, and assigns subjectivity and polarity tags for each sentence. Based on its predictions, we define two predicates, *subjective*(POST) and *polarity*(POST). Both predicates are calculated by normalizing the number of subjective/objective tags and positive/negative polarity tags marked by OpinionFinder. The normalization keeps these values in the $[0, 1]$ interval.

Forums are structured entities, organized by high-level topics (at the forum level) and specific topics (thread level). Including these structural relationships allows our model to identify structural relations between forum posts and connect them with students participating in the forum discussions. The predicates representing forum structure are *sameThread*(POST_1, POST_2) and *sameForum*(THREAD_1, THREAD_2), which are true for posts in the same thread and threads in the same forum, respectively. These predicates are used in rules to propagate *survival* and *engagement* values among students interacting on the forums.

Temporal Features Student activity level changes over the span of the course. Students are often active at early stages and lose interest as the course progresses. To include signals of how student activity changes over time, we introduce a set of temporal features. We divide the course into three time periods: *start*, *mid*, and *end*. The time period splits are constructed by dividing the course by duration into three equal chunks. The temporal features *lastQuiz*, *lastLecture*, *lastPost*, *lastView* and *lastVote* indicate the time-period in which each last interaction of the user occurred. These features measure to what lengths the user participated in different aspects of the course.

Constructing Complex Rules We use the features above to construct meaningful PSL rules using logical connectives, as demonstrated in Table 1[3]. The PSL model associates these rules with student survival, either directly or indirectly using latent variables. We explain this process in the following section.

• Behavioral Feature
$\textit{postActivity}(U) \wedge \textit{reputation}(U)$
• Forum Content Features
$\textit{posts}(U, P) \wedge \textit{polarity}(P)$
• Forum Interaction Feature
$\textit{posts}(U_1, P_1) \wedge \textit{posts}(U_2, P_2) \wedge \textit{sameThread}(P_1, P_2)$
• Temporal Features
$\textit{lastQuiz}(U, T_1) \wedge \textit{lastLecture}(U, T_1) \wedge \textit{lastPost}(U, T_1)$

Table 1: Constructing complex rules in PSL

[3]Full model available at https://github.com/artir/ramesh-aaai14.

PSL Student Survival Models

Probabilistic relational modeling is a popular approach for capturing structural dependencies such as the one above, and have been applied to a wide range of problems. We are interested in predicting if each student *survives* the course, i.e., whether the student took any of the quizzes/assignments near the end of the class. Student survival is calculated as a Boolean value—*1* if the student takes the last few quizzes/assignments and *0* if the student does not.

We construct two different PSL models for predicting student survival in a MOOC setting—first, a flat model (denoted DIRECT) that directly infers student survival from observable features, and second, a latent variable model (LATENT) that infers student engagement as a hidden variable to predict student survival. By building both models, we are able to evaluate the contribution of the abstraction created by formulating engagement patterns as latent variables.

• $\textit{postActivity}(U) \wedge \textit{reputation}(U) \rightarrow \textit{survival}(U)$
• $\textit{voteActivity}(U) \wedge \textit{reputation}(U) \rightarrow \textit{survival}(U)$
• $\textit{posts}(U,P) \wedge \textit{polarity}(P) \rightarrow \textit{survival}(U)$
• $\textit{posts}(U,P) \wedge \neg\textit{polarity}(P) \rightarrow \neg\textit{survival}(U)$
• $\textit{posts}(U,P) \wedge \textit{upvote}(P) \rightarrow \textit{survival}(U)$
• $\textit{posts}(U_1,P_1) \wedge \textit{posts}(U_2,P_2) \wedge \textit{survival}(U_1)$ $\wedge \textit{sameThread}(P_1,P_2) \rightarrow \textit{survival}(U_2)$

Table 2: Rules for the DIRECT model.

• $\textit{postActivity}(U) \wedge \textit{reputation}(U) \rightarrow \textit{eActive}(U)$
• $\textit{voteActivity}(U) \wedge \textit{reputation}(U) \rightarrow \textit{ePassive}(U)$
• $\textit{posts}(U,P) \wedge \textit{polarity}(P) \rightarrow \textit{eActive}(U)$
• $\textit{votes}(U,P) \wedge \textit{polarity}(P) \rightarrow \textit{ePassive}(U)$
• $\textit{posts}(U,P) \wedge \textit{upvote}(P) \rightarrow \textit{eActive}(U)$
• $\textit{posts}(U_1,P_1) \wedge \textit{posts}(U_2,P_2) \wedge \textit{eActive}(U_1)$ $\wedge \textit{sameThread}(P_1,P_2) \rightarrow \textit{eActive}(U_2)$
• $\textit{eActive}(U) \wedge \textit{ePassive}(U) \rightarrow \textit{survival}(U)$

Table 3: Rules for the LATENT model.

PSL-DIRECT In our DIRECT PSL model, we model student survival by using the observable behavioral features exhibited by the student, linguistic features corresponding to the content of posts, and structural features derived from forum interactions. Meaningful combinations of one or more observable behavioral features (described in the Features section) are used to predict *survival*. Table 2 contains a subset of rules used in this model (*U* and *P* in tables 2 and 3 refer to USER and POST respectively). As evident in these examples, the simple model contains rules that allow observable features to directly imply student survival.

PSL-LATENT In our second model, we enhance this type of reasoning by including latent variables semantically based on concepts of student engagement. These variables cannot be directly measured from the data, and we therefore treat student engagement as a latent variable and associate various observed features to one or more forms of engagement.

We define three types of engagement variables, denoted *engagement_active*, *engagement_passive* and *disengagement* to capture three types of student engagement in MOOCs. *engagement_active* represents students actively engaged in the course by participating in the forums, *engagement_passive* represents students following the class materials but not making an active presence in the forums, and *disengagement* represents students discontinuing from engaging with the course both actively and passively. We associate different features representing MOOC attributes relevant for each engagement type to the latent engagement variables.

• **Active Engagement** Submitting lectures, posting on discussion forums and giving quizzes are signs of active engagement by the student.

• **Passive Engagement** Submitting lectures, viewing/voting/subscribing to posts on discussion forums are signs of passive engagement.

• **Disengagement** Temporal features, indicating the last point of user's activity, capture signs of disengagement.

We connect the latent engagement variables to student survival by introducing PSL rules. In this model, some of the observable features (e.g, *postActivity*, *voteActivity*, *viewActivity*) are used to classify students into one or more forms of engagement or disengagement. Then, the engagement predicates—conjuncted with other observable features that are not used to imply user engagement, such as *reputation*—are modeled to imply student survival. For example, in Table 3, conjunction of *postActivity* and *reputation* implies *engagement_active*; conjunction of *voteActivity* and *reputation* implies *engagement_passive*; while *engagement_active* and *engagement_passive* implies *survival*. Note that *engagement_active* and *engagement_passive* are abbreviated to *eActive* and *ePassive* in the table.

We train the weights for the model by performing expectation maximization with *survival* as the target variable. The resulting model with latent engagement suggests which forms of engagement are good indicators of student survival. Thus, not only does the latent model produce better predictive performance, but it can provide more insight into MOOC user behavior than a simpler model.

In our experiments, we consider some meaningful combinations of data from different phases. The following section provides more details about the student survival experiments.

Empirical Evaluation

We conducted experiments to answer the following questions. First, how effective are our models at predicting student survival? Second, what are the key factors influencing student survival in an online setting?

Datasets and Experimental Setup

We evaluate our models on three Coursera MOOCs at University of Maryland: *Surviving Disruptive Technologies*, *Women and the Civil Rights Movement*, and *Gene and the Human Condition*. In discussion below, we refer to these courses as DISR-TECH, WOMEN-CIVIL, and GENE, respectively. Our data consists of anonymized student records, grades, and online behavior recorded during the seven week duration of each course.

Figure 2 plots the number of participants in different course-related activities. Of the total number of students registered, around 5% of the students in DISR-TECH and WOMEN-CIVIL , and around 14% in GENE, complete the final exam. We use this to define course *survival*. In all the three courses, the most prominent activity exhibited by students while on the site is viewing lectures. Hence, we rank students based on number of lectures viewed, as a baseline (denoted LECTURE-RANK in our tables) for comparison. The other prevalent activities include quiz submission and viewing forum content. Observing the statistics, DISR-TECH and WOMEN-CIVIL have a higher percentage of total registered students participating in forums compared to GENE.

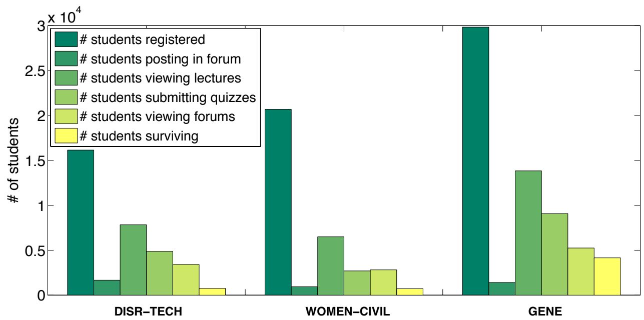


Figure 2: Comparison of number of users participating in course-related activities in three courses.

We evaluate the model on the following metrics: area under the precision-recall curve for positive and negative labels, and area under the ROC curve. We use ten-fold cross-validation, leaving out 10% of the data for testing and revealing the rest for training the model weights.

Student Survival Analysis Our experiments in the student survival models are aimed at measuring student survival by understanding factors influencing students' survival in the course, engagement types and changes in engagement, and the effectiveness of prediction using different time periods of the course.

In our first set of experiments, we consider all student activity during the entire course to predict whether each student takes the final quiz. We choose a baseline by ranking students on number of lectures viewed. The scores for our baseline and the two models are listed in Table 4. The baseline using just the lecture submission feature can predict dropouts reasonably well, but its comparatively low precision and recall for positive survival (AUC-PR pos.) indicates that using this feature alone is suboptimal. Because of the class imbalance and the high proportion of students who drop out, models that can identify students who will complete the course are more valuable. The strength of our model comes from combining behavioral, linguistic, temporal, and structural features for predicting student survival. Our models DIRECT and LATENT significantly improve on the baseline, and the LATENT model outperforms the DIRECT model.

Early Prediction Predicting student survival can provide instructors with a powerful tool if these predictions can be made reliably before the students disengage and drop out. We simulate this scenario by training our model over data collected early in the course. The student survival labels are the same as for the complete dataset (i.e., whether the student submitted the final quizzes/assignments at the end of the course), but our models are only given access to data from the early parts of the course. We divide the course into three parts according to the duration of the course, as mentioned in the *Modeling MOOC Student Activity* section under *Temporal Features*.

Table 5 lists the performance metrics for our two models using different splits in the data. Similarly to the results in Table 4, the change in the AUC-PR (Neg.) scores are negligible and close to optimal for all models because of class imbalance. To highlight the strength our models, we only report the AUC-PR (Pos.) scores of the models. We refer to the three phases of each course by *start*, *mid*, and *end*. *start-mid* refers to data collected by combining time spans *start* and *mid*, and *start-end* refers to data collected over the entire course.

Early prediction scores, in Table 5 under *start*, *mid*, and *start-mid* (i.e., survival prediction using partial data), indicate that our model can indeed make these predictions reliably. As the data available is closer to the end of the course, models make better predictions. Just as in the previous experimental setting, the latent engagement model achieves the highest prediction quality. The LATENT model for *start* outperforms DIRECT model on all time-periods in WOMEN-CIVIL, including the ones which contain more data (*mid*, *end*, and *start-mid*).

From the results, it appears that the middle phase (*mid*) is the most important phase to monitor student activity for predicting whether the student will survive the length of the course. Our model produces higher AUC-PR values when using data from the *mid* phase, compared to the settings where we use data from the *start* phase, and an almost equal value when compared to *start-mid*. We hypothesize that this is due to the presence of a larger student population in the *start* that fails to remain engaged until the end. This phenomenon is typical in both traditional and online classrooms where students familiarize themselves with the course and then decide whether to stay or drop out. Eliminating data collected from this population helps improve our prediction of student survival, as indicated by an increase in performance values for *mid*.

Feature Analysis

We evaluate the contribution of each feature by leaving each feature out and observing the resulting change in prediction performance values. The features considered are: posting in forums (*post*), viewing forum content (*view*), time period of last quiz submitted by user (*quiz*), temporal features (*temporal*), and viewing lectures (*lecture*). The model with all the features included is given by *all*. For each of the five

COURSE	MODEL	AUC-PR Pos.	AUC-PR Neg.	AUC-ROC
DISR-TECH	LECTURE-RANK	0.333	0.998	0.957
	DIRECT	0.393	0.997	0.936
	LATENT	0.546	0.998	0.969
WOMEN-CIVIL	LECTURE-RANK	0.508	0.995	0.946
	DIRECT	0.565	0.995	0.940
	LATENT	0.816	0.998	0.983
GENE	LECTURE-RANK	0.688	0.984	0.938
	DIRECT	0.757	0.985	0.939
	LATENT	0.818	0.985	0.944

Table 4: Performance of LECTURE-RANK, DIRECT and LATENT models in predicting student survival

COURSE	MODEL	*start*	*mid*	*end*	*start-mid*
DISR-TECH	LECTURE-RANK	0.204	0.280	0.324	0.269
	DIRECT	0.304	0.400	0.470	0.372
	LATENT	0.417	0.454	0.629	0.451
WOMEN-CIVIL	LECTURE-RANK	0.538	0.518	0.415	0.533
	DIRECT	0.593	0.647	0.492	0.596
	LATENT	0.674	0.722	0.733	0.699
GENE	LECTURE-RANK	0.552	0.648	0.677	0.650
	DIRECT	0.647	0.755	0.784	0.692
	LATENT	0.705	0.755	0.789	0.778

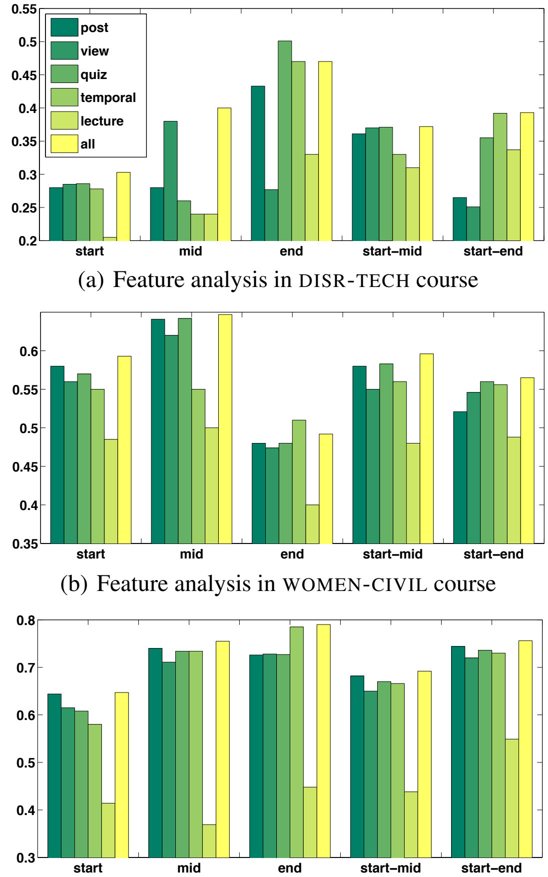
Table 5: Early prediction performance of LECTURE-RANK, DIRECT and LATENT models in time-periods *start*, *mid*, *end*, and *start-mid*

features above, we construct a PSL model by omitting the relevant feature from all PSL rules. Figure 3 plots the results from these tests for phases—*start*, *mid* and *end*. The *lecture* feature is consistently important for predicting student survival. Discussion forums serve as a platform connecting students worldwide enrolled in the course, hence activity in the discussion forums also turns out to be a strongly contributing feature. Since, the concentration of forum posts in the courses analyzed is more in the *mid* and *end* phases, posting in forums is accordingly more important during the *mid* and *end* phases. Simply viewing content on the forums is also a strong feature, contributing consistently in all phases across all courses. In fact, from Figure 3, we can see that the feature strength of forum views is second only to lecture views. While *quiz* is a strong feature in most phases, it can be observed that it is not a strong feature in the *end* phase of the course. The data suggests that this effect is because quiz taking gradually drops as the course progresses, leading to fewer quiz takers in the *end* phase. Hence, *temporal* and *quiz* are not very predictive features in the *end* phase.

Discussion

In this work, we take a step toward understanding student engagement in MOOCs using data-driven methods. We formalize, using PSL, intuitions that student engagement can be modeled as a complex interaction of behavioral, linguistic and social cues, and we model student engagement types as latent variables over these cues. Our model constructs an interpretation for latent engagement variables from data and predict student course completion reliably, even at early stages in the course. These results are a first step toward facilitating instructors' intervention at critical points, thus helping improve course retention rates.

The latent formulation we present can be extended to more sophisticated modeling by including additional latent factors that affect academic performance, such as motivation, self-regulation and tenacity. These compelling directions for future interdisciplinary investigation can provide a better understanding of MOOC students.



(a) Feature analysis in DISR-TECH course

(b) Feature analysis in WOMEN-CIVIL course

(c) Feature analysis in GENE course

Figure 3: Bar graph showing AUC-PR (Pos.) value upon removal of each feature from the DIRECT model

Acknowledgements This work is supported by National Science Foundation under Grant Nos. IIS0746930 and CCF0937094. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Anderson, A.; Huttenlocher, D.; Kleinberg, J.; and Leskovec, J. 2014. Engaging with massive online courses. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, 687–698. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Bach, S. H.; Huang, B.; London, B.; and Getoor, L. 2013. Hinge-loss Markov random fields: Convex inference for structured prediction. In *Uncertainty in Artificial Intelligence*.

Balakrishnan, G. 2013. Predicting student retention in massive open online courses using hidden Markov models. Master's thesis, EECS Department, University of California, Berkeley.

Broecheler, M.; Mihalkova, L.; and Getoor, L. 2010. Probabilistic similarity logic. In *Uncertainty in Artificial Intelligence (UAI)*.

Brusilovsky, P., and Millán, E. 2007. User models for adaptive hypermedia and adaptive educational systems. In Brusilovsky, P.; Kobsa, A.; and Nejdl, W., eds., *Lecture Notes In Computer Science, Vol. 4321*. Berlin, Heidelberg: Springer-Verlag. 3–53.

Carini, R.; Kuh, G.; and Klein, S. 2006. Student engagement and student learning: Testing the linkages. *Research in Higher Education* 47(1):1–32.

Clow, D. 2013. MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, 185–189. New York, NY, USA: ACM.

Getoor, L., and Taskar, B. 2007. *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

Kizilcec, R. F.; Piech, C.; and Schneider, E. 2013. Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, 170–179. New York, NY, USA: ACM.

Kotsiantis, S.; Pierrakeas, C.; and Pintelas, P. 2003. Preventing student dropout in distance learning using machine learning techniques. In Palade, V.; Howlett, R.; and Jain, L., eds., *Knowledge-Based Intelligent Information and Engineering Systems*, volume 2774 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 267–274.

Kuh, G. D. 2003. What were learning about student engagement from nsse: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning*.

Richards, S. J. 2012. A handbook of parametric survival models for actuarial use. *Scandinavian Actuarial Journal* 2012(4):233–257.

Wilson, T.; Hoffmann, P.; Somasundaran, S.; Kessler, J.; Wiebe, J.; Choi, Y.; Cardie, C.; Riloff, E.; and Patwardhan, S. 2005. OpinionFinder: A system for subjectivity analysis. In *Proceedings of HLT/EMNLP on Interactive Demonstrations*.

Yang, D.; Sinha, T.; Adamson, D.; and Penstein, R. 2013. Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *NIPS Workshop on Data Driven Education*.