Proceedings of EMOOCs 2017:
Work in Progress Papers of the Experience and Research Tracks and Position Papers of the Policy Track

Predicting Student Participation in Peer Reviews in MOOCs

Erkan Er, Miguel Luis Bote-Lorenzo, Eduardo Gómez-Sánchez, Yannis Dimitriadis, Juan Ignacio Asensio-Pérez

GSIC/EMIC, Universidad de Valladolid, Valladolid, Spain.
erkan@gsic.uva.es, {migbot|edugom|yannis|juaase}@tel.uva.es

Abstract. Assessing and providing feedback to thousands of student artefacts in MOOCs is an unfeasible task for instructors. Peer review, a well-known pedagogical approach that offers various learning gains, has been a common approach to address this practical challenge. However, low student participation is a potential barrier to the success of peer reviews. The present study proposes an approach to predict student participation in peer reviews in a MOOC context, which can be utilized to achieve an effective peer-review activity. We attempt to predict the number of different peer works that students will review for each of four assignments based on their past activities in the course. Results show that students' preceding activities were predictive of their participation in peer reviews starting from the first assignment, and that the prediction accuracy improved considerably with the inclusion of past peer-review activities.

Keywords: MOOC, Peer review, Engagement prediction, Regression

1 Introduction

Massive open online courses (MOOCs) enable millions to receive university-level courses at no cost. However, the massiveness comes with several practical challenges. One known challenge is the assessment of thousands of student artefacts (submitted to open-ended assignments) [1]. One approach to address this challenge has been the use of peer review (or peer assessment). Peer review is an active learning process in which a student work is examined and rated by another equal-status student [2]. Besides its utility in terms of reducing the workload of instructors, which is considered a main benefit in the MOOC context, peer review offers learning gains for both those students who performed the review and those whose work was reviewed. These benefits include, but are not limited to, the development of higher-order thinking skills, problem solving skills, communication skills, and teamwork skills [2, 3]. However, conducting an effective peer review itself is a challenge in large scales. One barrier to its successful implementation is the low student participation. Considering the lack of instructor mediation and the large diversity in MOOC participants (e.g., native language, culture, etc.), there are high chances that not many students will be naturally motivated to review a peer's work [4]. Lack of participation in peer review may result in situations in which the submissions of striving students remain ungraded, leading to a decrease in their motivation to continue the course. Nevertheless, as opposed to numerous studies that are concerned about resolving the validity issues of peer reviews [5, 6], there exists scarce works that investigated student participation in peer review at large scale [7]. Thus, there is a need for further research to contribute to the solution of this problem.

The present study proposes an approach to predict student participation in reviewing peers' work in a MOOC context, and in this paper, we share the preliminary findings of this in-progress research. In particular, we attempt to predict the number of different peer works that students will review for a specific assignment based on their past activities in the course. An accurate estimation of number of times a student will perform peer review can help instructors take timely actions to achieve a successful peer-review process [8]. For example, the peer-review task might be rather challenging for some students depending on their abilities [1], and these students may need more time for completing their reviews. Therefore, instead of a firm deadline for peer reviews, an adaptive schedule based on the predicted participation levels can be used to promote participation in peer reviews. In addition, this estimation might be utilized in designing other effective collaborative learning activities. For example, using the information regarding the levels of participation, student groups can be formed in a way that maximizes the likelihood that each peer work will be reviewed by another group member. As student participation in peer reviews can be also considered an engagement indicator, other approaches that are used to foster engagement can be applied [9].

In the following section, we describe the course data at hand and the features generated for the prediction task. Next, we present the experimental study by describing the details of the method and the results regarding the performance of each prediction model employed. We conclude by presenting the follow-up research ideas.

2 Predicting Participation in Peer Reviews

2.1 Course Data

The course data for this study was retrieved from a public dataset published by Canvas Network [1]. No contextual information was available (e.g., whether the peer review was mandatory or not), but we attempted to make some inferences about the course design based on the available log data, since such contextual information may help us explain better the prediction results. The course had 3620 enrolments and contained four main assignments (each worth 25 points) for which students needed to upload a specific artefact. These assignments were reviewed by peers, and they were scheduled starting from the second week of the course with a one-week interval between each one of them.

The course data contains fine-grained information regarding students' content visits as well as their various activities in discussions, assignments, and quizzes (e.g., create, view, or subscribe to a discussion topic, submit or view an assignment, etc.). Moreover,

[1] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/XB2TLU
The id of the course is 770000832960949.

we identified the number of peer submissions reviewed by each student (at each assignment), which is the outcome (or dependent) variable in this study. Given that most students reviewed three different peer works at each assignment (see Figure 1), it is likely that students were suggested to perform at least 3 reviews by the course instructors. Descriptive statistics regarding the outcome variable are given in the figure below.

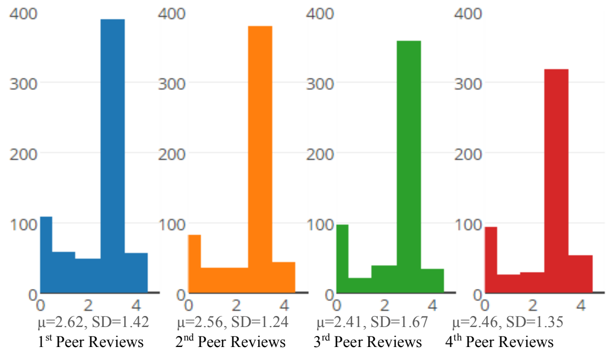

Fig. 1. Histograms of peer works reviewed along with the mean and standard deviation scores.

2.2 Feature Generation

In this subsection, we briefly discuss the rationale for the features generated to be used in the prediction of student participation in peer reviews. Active MOOC learners are likely to perform well as a result of their consistent participation in most activities of the course including the peer reviews [1]. Such active students may probably achieve a good understanding of the course content as a result of their engagement (e.g., viewing course content pages, participating in discussions, completing quizzes) [10], and therefore they are more likely to feel confident reviewing a peer's work. Accordingly, in the present study, we hypothesize that students' preceding engagement in the course is associated with their subsequent participation in peer-review activities. For this purpose, we built a set of predictors (or features) based on various student activities in the course (e.g., discussions, assignments, and quizzes) and used them to predict students' participation in peer-review activities. Based on the overview of the data at hand and the previous research [11], a set of features (see Table 1) was generated to characterize the student engagement in the course. These features considered only student activities during the last 6 days *before* the deadline of the corresponding assignment (since there was a one-week interval between assignments).

3 Experimental Study

3.1 Method

Considering the large set of features, we preferred to use regularized regression methods, which can penalize the weak predictors and eliminate them to improve the model performance. Three regularized regression methods were chosen: least absolute shrinkage and selection operator (LASSO), elastic net, and ridge regression since these methods incorporate an internal feature-selection mechanism [12]. These three methods were applied to make a prediction regarding the number of different peer works that were reviewed by students at each assignment. To evaluate the model performance, the mean absolute error (MAE) scores were used [13]. Since the sample size was small, 10-fold cross validation method was used.

Table 1. Features generated based on students' overall engagement in the course

{*a*}_{*b*}_count	Total number of requests made.
{*a*}_{*b*}_avg_p_day	Average requests per day.
{*a*}_{*b*}_count_li	Total number of requests made when later requests were given a higher weight {1/6, 1/5, 1/4, 1/3, 1/2, and 1}.
{*a*}_{*b*}_count_ei	Total number of requests made when earlier requests were given a higher weight {1, 1/2, 1/3, 1/4, 1/5, and 1/6}.
{*a*}_{*b*}_days	Total number of days with at least one request made.
{*a*}_{*b*}_days_li	Number of days with at least one request when later requests were given a higher weight {1/6, 1/5, 1/4, 1/3, 1/2, and 1}.
{*a*}_{*b*}_days_ei	Number of days with at least one request when earlier requests were given a higher weight {1, 1/2, 1/3, 1/4, 1/5, and 1/6}.
{*a*}_{*b*}_{*n*}x_times	Runs of *n* consecutive days with at least one request ($1 < n < 7$).
{*a*}_count_qs	Total number of quiz submissions.
{*a*}_unc_qs	Total number of quizzes that were not completed.
{*a*}_ttl_time_taken	Total time spent on quizzes (in minutes).
{*a*}_avg_time_taken	Average time spent on quizzes (in minutes).
{*a*}_ttl_qs_score	Total quiz scores.
{*a*}_avg_qs_score	Average quiz scores.
{*a*}_ttl_qs_attempt	Total number of quiz attempts.
{*a*}_avg_qs_attempt	Average number of quiz attempts.
{*a*}_de_count	Total number of discussion entries.
{*a*}_de_msg_cc	Average character of the discussion entries posted.
{*a*}_pr_ count	Total number of peer reviews performed.
{*a*}_pr_subms_count	Total number of student submissions reviewed.

a denotes the assignment period (1, 2, 3, or 4) and ***b*** denotes the type of the request (cr – content requests, qr –quiz requests, ar –assignment requests, and dr –discussion requests).

3.2 Preliminary Findings

Using only those features that are available prior to the peer review activity, the number of peer reviews performed was predicted for each assignment period separately. In each prediction, only students who submitted the corresponding assignment were included since only those students can review others' submissions. The MAE scores of all models are provided as a boxplot in Figure 2. Based on Figure 2, for all methods, the accuracy seemed to increase at each subsequent prediction and levelled at the 4th peer-review activity. This increase was expected as the previous peer-review activities were considered starting from the 2nd set of submissions. That is, features derived from student activities in the course were relevant when predicting the subsequent participation in peer reviews starting from the first assignment, and features regarding the past peer-review participation were the strongest predictors. These results support our hypothesis regarding the predictive potential of students' overall engagement in a course in their subsequent peer-review participation. Moreover, LASSO and Elastic Net yielded a higher accuracy at each assignment period compared to ridge regression. Although majority of the predictions were within an acceptable error range, particularly for LASSO and Elastic Net, there were some outliers, which are to be further examined.

Based on our examination on the coefficients of each feature (as determined by the regression methods), students' previous peer-review participation was in general the strongest predictors (e.g., _pr_count, _pr_subms_count). Many of the retained features were related to students' discussion activities (e.g., _dr_count, _de_msg_cc, dr_avg_p_day) and assignment activities (e.g., ar_{x}_days_bef, _ar_count). Features related to students' course content views were not strong predictors in overall.

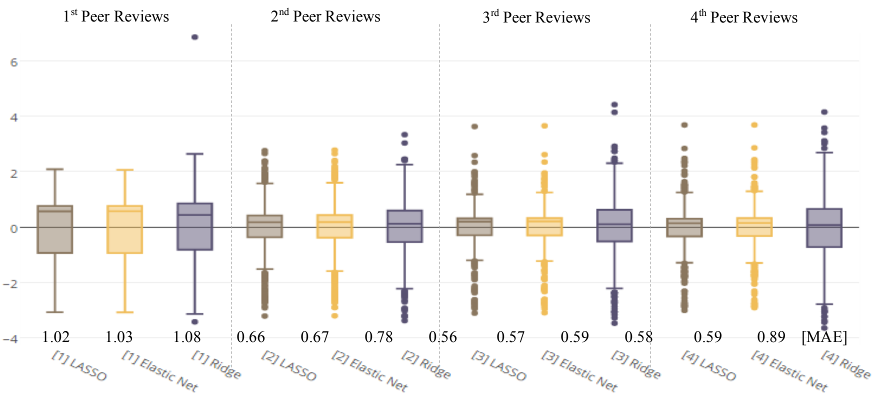

Fig. 2. The MAE scores and their distributions for each model at each peer-review activity.

4 Conclusions and Future Work

In this study, we presented the preliminary findings of our ongoing research on predicting student participation in peer reviews. The results suggest that students' preceding activities in a MOOC might be useful in predicting their participation in peer-review activities. The strongest predicters were not among the features associated with course content views, while they were among those associated with discussion and assignment activities. This finding is actually not surprising in a MOOC context since many MOOC learners may only view course content without active participation [14].

Among the regression methods, LASSO and Elastic Net performed better than the ridge regression. That is, the methods with more extreme penalization (e.g., LASSO) yielded a higher prediction accuracy, suggesting the presence of some irrelevant features. Therefore, we plan to perform a deeper analysis of the feature space and generate more features related to those with stronger predictive ability. We also plan to take a closer look at the outliers and explore the possible reasons, which may also inform our analysis on the feature space. Once the prediction model is refined and finalized, we plan to evaluate its performance in a real course.

Acknowledgements

Access to the data used in this paper was granted by Canvas Network. This work has been partially funded by research projects TIN2014-53199-C3-2-R and VA082U16.

References

1. Meek, S., Blakemore, L., Marks, L.: Is peer review an appropriate form of assessment in a MOOC? Student participation and performance in formative peer review. Assess. Eval. High. Educ. 1–14 (2016).
2. Topping, K.: Peer assessment between students in colleges and universities. Rev. Educ. Res. 68, 249–276 (1998).
3. Wen, M.L., Tsai, C.C., Chang, C.Y.: Attitudes towards peer assessment: a comparison of the perspectives of pre- service and in- service teachers. Innov. Educ. Teach. Int. 43, 83–92 (2006).
4. Suen, H.: Peer assessment for massive open online courses (MOOCs). Int. Rev. Res. Open Distrib. Learn. 15, (2014).
5. Kulkarni, C., Wei, K., Le, H., Papadopoulos, K., Cheng, J., Koller, D., Klemmer, S.: Peer and self assessment in massive online classes. ACM Trans. Comput. Interact. 20, 1–31 (2013).
6. Luo, H., Robinson, A., Park, J.-Y.: Peer grading in a MOOC: Reliability, validity, and perceived effects. J. Asynchronous Learn. Netw. 18, 1–14 (2014).
7. Estevez-Ayres, I., Crespo-García, R.M., Fisteus, J.A., Delgado-Kloos, C.: An algorithm for peer review matching in Massive courses for minimising students' frustration. J. Univers. Comput. Sci. 19, 2173–2197 (2013).
8. Gikandi, J.W., Morrow, D., Davis, N.E.: Online formative assessment in higher education: A review of the literature. Comput. Educ. 57, 2333–2351 (2011).
9. Neubaum, G., Wichmann, A., Eimler, S.C., Krämer, N.C.: Investigating incentives for students to provide peer feedback in a semi-open online course. In: The International Symposium on Open Collaboration. pp. 1–7 (2014).
10. Tseng, S.-F., Tsao, Y.-W., Yu, L.-C., Chan, C.-L., Lai, K.R.: Who will pass? Analyzing learner behaviors in MOOCs. Res. Pract. Technol. Enhanc. Learn. 11, 1–11 (2016).
11. Bote-Lorenzo, M.L., Gómez-Sánchez, E.: Predicting the decrease of engagement indicators in a MOOC. In: Seventh International Conference on Learning Analytics and Knowledge. pp. 143–147 (2017).
12. Hastie, T., Tibshirani, R., Friedman, J.: The elements of statistical learning: Data mining, inference, and prediction. (2009).
13. Sawyer, R.: Sample size and the accuracy of predictions made from multiple regression equations. J. Educ. Stat. 7, 91–104 (1982).
14. Kizilcec, R.F., Piech, C., Schneider, E.: Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In: Third International Conference on Learning Analytics and Knowledge. pp. 170–179 (2013).