



Explaining Errors in Predictions of At-Risk Students in Distance Learning Education

Martin Hlosta ^(✉) , Tina Papathoma, and Christothea Herodotou

The Open University, Milton Keynes, UK
{martin.hlosta,tina.papathoma,
christothea.herodotou}@open.ac.uk

Abstract. Despite recognising the importance of transparency and understanding of predictive models, little effort has been made to investigate the errors made by these models. In this paper, we address this gap by interviewing 12 students whose results and predictions of submitting their assignment differed. Following our previous quantitative analysis of 25,000+ students, we conducted online interviews with two groups of students: those predicted to submit their assignment, yet they did not (False Negative) and those predicted not to submit, yet they did (False Positive). Interviews revealed that, in False Negatives, the non-submission of assignments was explained by personal, financial and practical reasons. Overall, the factors explaining the different outcomes were not related to any of the student data currently captured by the predictive model.

Keywords: Learning analytics · At-risk students · Error analysis · Interviews

1 Introduction

Identifying correctly at-risk students has emerged into one of the most prevalent topics in Learning Analytics (LA) and education in general [1]. The identification of at-risk students using Predictive Learning Analytics (PLAs) and followed by a subsequent intervention targeting flagged students (e.g., phone call) could tackle this problem. Many published papers focused on achieving the highest prediction performance, often comparing several learning algorithms. Machine learning models are more likely to exhibit some sort of error hence, the need to understand and explain these errors. In a cross-disciplinary field such as LA, not having the best model could still help understand or even improve student learning. Kitto et al. [2] argued that having imperfect models does not necessarily mean that these should not be deployed. As the LA field is maturing, it becomes essential to understand how models are behaving and how errors occur [3, 4].

Only few studies have examined errors up to now. This paper aims to explain errors in predictions through 12 in-depth interviews with undergraduate online students wrongly predicted as being/not being at risk of failing their next assignment. We treated False Positive (FP) and False Negative (FN) errors separately. Following [5], we refer to FP as students predicted as being at-risk but succeeded, and FN as students that failed despite predicted to succeed. We build on the work of Calvert et al. [6] that

investigated within a single online course why some FP students passed despite predictions showing the opposite and our early quantitative results from [7].

2 Methodology

To analyse the predictive model errors, we used a mixed-methods approach (See Fig. 1). We focused on first year STEM courses and predictions for the first assignment only (A1), when dropout is more likely to happen [8, 9]. The predictions were enhanced by additional data: course context (e.g. the length of the course) and future data from the weeks following the predictions, unknown during the prediction's generation. Predictions for each course were put together in one matrix and only predictions with confidence ≥ 0.85 were selected. A Decision Tree was constructed to distinguish between (1) FP and True Positive (TP) and (2) FN and True Negative (TN).

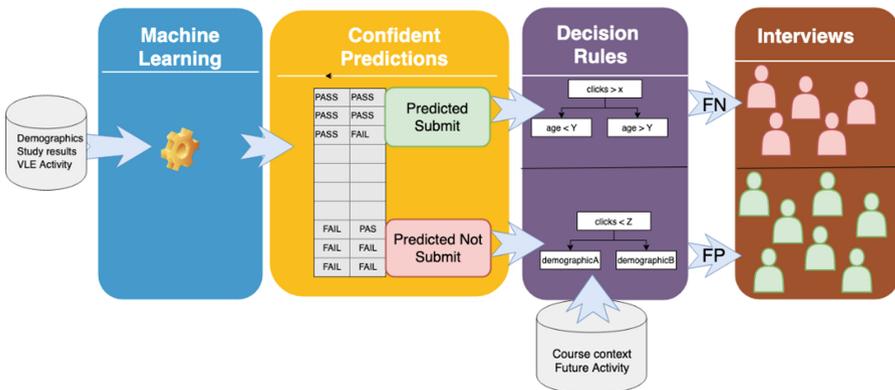


Fig. 1. The schema of the methodology - quantitative analysis followed by interviews.

After getting a favourable opinion from the university's Ethical Committee, we conducted 12 semi-structured interviews with students lasting 20 to 40 min. The interview schedule was developed by two of the authors, piloted with one student, and the analysis followed inter-rater reliability principles. Students were not new to the university, assuming that they would have devised strategies on how to successfully complete their assignment without accessing the Virtual Learning Environment (VLE). Gift vouchers were offered. We grouped participants to (a) students predicted to submit yet they did not submit (FN; $N = 7$) and (b) students predicted not to submit yet they submitted (FP; $N = 5$). Following other published work [10], we analysed students as individual case studies creating a distinct profile picture for each student. The following themes emerged from the thematic analysis [11]: motivations for taking the module, studying patterns, reasons for not submitting the assignment, factors that helped or hindered submission, tutor contact, student contact, recommendations for other students so that they submit and proposed module changes. We then plotted this

information on a table and identified similarities and differences within and between the two cohorts of students¹.

3 Results

3.1 Quantitative Analysis

Considering the predictions two weeks before the deadline of A1, we analysed 38,073 predictions in 17 courses in 62 presentations between 2017–2019, having 29,247 students (383 FP, 1,507 FN, 2,671 TP and 33,512 TN). The ROC AUC over all predictions was 0.8897. For confident Not Submit predictions, the decision tree classified correctly 50.91% of the FP errors with 75.29% precision (195 students). The strongest attribute was the number of clicks one week before the deadline of the first assignment (confidence 0.82). For the confident Submit predictions, the model distinguished 18.73% of the FN errors with precision 68.73% (1,036 students). The strongest attribute related to a dramatic decrease in students' activity in the last week before assignment 1 (A1) in courses with high activity (confidence 0.83). Both types of errors were associated with a change of student activity after the predictions were generated, and it is worth further examination.

3.2 Qualitative Analysis

Participants (see Table 1) were older than those invited to take part in the study, more successful in their previous courses, female repeating the same course.

Table 1. Invited and interviewed participants – demographic information.

	Age (avg)	Disabled (%)	Female (%)	Previous Pass (%)	Repeat Course (%)	Totals
Invited	38	22	42	64	26	131
Interviewed	45	25	57	70	33	12

FN - Predicted to Submit but did Not Submit: Participants were motivated to study either because they were driven by completing a qualification/degree or out of interest. Their studying patterns were rather random, with no strict schedule. The reasons explaining non-submission were related to family matters/issues (i.e. caring responsibilities), practical issues (i.e. no internet connection at the time of submission) or they were restricted to submit because of student financial issues. FN_2 who took the course out of interest, found it pointless to submit her assignment as it only weighted 7% of the final grade and it was too easy for her. On the contrary, FN_5 found it difficult to

¹ The table can be found here: <https://docs.google.com/spreadsheets/d/1MwT-luUSI96XYIGhjz24pXdwbnkn-nLHae9n8gLXiDk>.

submit because of her lack of digital skills and absence of detailed guidance. Further, student contact with tutors was minimal and related to requesting an extension to submitting the assignment. FN_11 mentioned that although she contacted her tutor via email, the tutor never replied. Two interviewees reported that their tutor support was helpful with the tutor proactively getting in touch and communicating with them. Interacting with other students was not common for four interviewees. FP_3 reported though that she helped other students and FN_4 used Facebook Groups and forums to communicate socially.

FN students made suggestions for future students to follow the online study guidance and plan ahead for submitting assignments on time. FN_2 who had prior knowledge suggested that assignments should have optional questions for the needs of more advanced students. Two participants would like to have online tutorials with a tutor to guide the assignment submission. FN_5 suggested that the course should be more accessible by adding detailed guidelines on technical aspects for submission. Most participants took the course for the first time apart from one interviewee.

FP - predicted Not to Submit but Submitted: All participants were motivated to take their course in order to get a qualification/degree. Their studying patterns varied mostly studying in the evenings. The reasons they managed to submit related to the fact that this was not the first time they were taking the course. Two of them took the course for the second time. FP_3 was determined to submit as it was their third time taking the course. Two interviewees took the course for the first time. FP_7 on the other hand, did not prepare for the assignment, yet answered the assignment questions as they had some prior knowledge. The other two interviewees submitted after watching videos, consulting books, or with help from external networks.

Contact with tutors was minimal. FP_7 only contacted their tutor for an extension. No interactions with other students were reported. In terms of recommendations, FP_3 suggested that asking for support from their tutor is important although they did not initiate that. FP_8 and FP_9 suggested looking at the VLE material in a timely manner and prepare early on. They proposed more contact with teachers and suggested that audio recordings would be a good addition. Interestingly, the interviewee who was taking the course for the third time, mentioned that assignments should be given more weight towards the final grade. FP_7 suggested that students with prior knowledge or expertise on a topic should be allowed to skip an assignment.

4 Conclusions

None of the predictive errors could be fully explained by only looking at the course data. Errors were explained by factors not currently captured by the university data sets, including personal, technical and financial issues students faced before submission. The factors reported are rather hard to capture automatically and in a timely manner to support students with difficulties. Hence, the role of teachers becomes critical; pastoral and proactive care could identify and resolve such issues on time and enable students to succeed. Existing studies already showcased the significance of teachers' monitoring and intervening with students at risk for better learning outcomes [12]. A university-

wide policy accompanied by relevant teachers' training as to when and how teachers should get in touch with their students would ensure that academic connection and social presence are established [13, 14]. Given that we do not gather data from external systems, errors might be hard to prevent in the future. Yet, we could add error explanations especially for students submitting their assignment (e.g. taking the course for a second time).

References

1. Ochoa, X., Merceron, A.: Quantitative and qualitative analysis of the learning analytics and knowledge conference 2018. *J. Learn. Anal.* **5**(3), 154–166 (2018)
2. Kitto, K., Shum, S.B., Gibson, A.: Embracing imperfection in learning analytics. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, ACM (2018)
3. Gardner, J., et al.: Modeling and experimental design for MOOC dropout prediction: a replication perspective. In: *International Educational Data Mining Society* (2019)
4. Taylor, C., Veeramachaneni, K., O'Reilly, U.-M.: Likely to stop? predicting stopout in massive open online courses. arXiv preprint [arXiv:1408.3382](https://arxiv.org/abs/1408.3382) (2014)
5. Archer, E., Prinsloo, P.: Speaking the unspoken in learning analytics: troubling the defaults. *Assess. Eval. Higher Educ.* **1**, 1–13 (2019)
6. Calvert, C., Hilliam, R.: Student feedback to improved retention: using a mixed-methods approach to extend specific feedback to a generalisable concept. *Open Learn. J. Open Dist. e-Learning* **34**(1), 103–117 (2019)
7. Hlosta, M., et al.: Why predictions of at-risk students are not 100% accurate? showing patterns in false positive and false negative predictions. In: *Companion Proceedings of the 10th International Learning Analytics & Knowledge Conference*. Frankfurt, Germany 2020 (Accepted)
8. Walker-Gibbs, B., Ajjawim, R., Rowe, E.: Success and failure in higher education on uneven playing fields (2019)
9. Hlosta, M., Zdrahal, Z., Zendulka, J.: Ouroboros: early identification of at-risk students without models based on legacy data. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, ACM (2017)
10. Herodotou, C., et al.: Implementing predictive learning analytics on a large scale: the teacher's perspective. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (2017)
11. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qualit. Res. Psychol.* **3**(2), 77–101 (2006)
12. Herodotou, C., et al.: Empowering online teachers through predictive learning analytics. *British J. Educ. Technology* (2019)
13. Milen, J.F., Berger, J.B.: A modified model of college student persistence: Exploring the relationship between Astin's theory of involvement and Tinto's theory of student departure. *J. College Student Dev.* **38**(4), 387 (1997)
14. Lukosius, V., Pennington, J.B., Olorunniwo, F.O.: How students' perceptions of support systems affect their intentions to drop out or transfer out of college. *Rev. High. Educ. Self-Learn.* **6**(18), 209–221 (2013)