How Does Student Behaviour Change Approaching Dropout? A Study of Gender and School Year Differences

Jessica McBroom, Irena Koprinska and Kalina Yacef
School of Computer Science, University of Sydney, Sydney, NSW 2006, Australia
{jmcb6755, irena.koprinska, kalina.yacef}@sydney.edu.au

ABSTRACT

In the context of online education, one important consideration is ensuring learning is equitable for a diverse range of students. In particular, understanding how factors such as gender and age can affect student behaviour is crucial for adapting courses to better suit the needs of these students. In this paper, using data from an online introductory programming course, we apply hierarchical clustering to identify changes in student behaviour as students approach dropping out from a course. By considering how these behavioural trends differ based on gender and school year level, we then discuss how this information can lead to insights to assist in improving equity and educational outcomes.

Keywords

gender equity, age equity, student dropout, student behaviour, computer programming education, behavioural trends, hierarchical clustering

1. INTRODUCTION

In an educational setting, ensuring that all students receive equitable learning opportunities is a challenge of great significance. This is particularly important in the context of online education, where teachers may be unable to personally monitor all students due to large cohort sizes, and where the increased accessibility of course materials can allow for very diverse ranges of students. It is also particularly important in areas where certain groups are under-represented, since inequitable education may discourage students from these groups from entering the field. Recent work on improving educational equity has often focused on improvements at an organisational level, such as through teacher training [11], frameworks for addressing equity challenges [7] or analyses of funding distributions [2].

A particularly promising avenue for improving educational equity is the analysis of student behaviour. In particular, educational data mining and analysis techniques can be utilised to understand how students from different backgrounds respond differently to a course, thereby providing insight into how the course can be modified to improve learning outcomes and equity. Previous work employing such techniques has considered, for example, differences in behaviour at school and home [4], course enrolment and completion rates [10], social behaviour [3], online participation and activity [6, 9], debugging techniques used [8] and motivation for study [5] for different student groups.

In contrast to previous work, this paper uses a hierarchical clustering technique to analyse the evolution of behaviour of students who drop out of an introductory programming course. In particular, samples of student behaviour are taken at different stages during a course (e.g. when they first begin, at points midway through and just before they drop out). These samples are then clustered to detect changes over time and to compare students of different gender and grade groups.

2. DATA

Our data come from a beginner-level Python programming course run in 2018. The course was run online for school students primarily in Australia over a 5 week period, and consisted of weekly exercises interleaved with notes on different topics. In total, this amounted to 40 exercises. Of the 6516 students who attempted the first exercises of the course, 82% dropped out before completing the last exercise.

3. METHODOLOGY

To observe how student behaviour changed as students came closer to dropping out, we analysed the behaviour of all students who completed at least 10 exercises in the first four weeks of the course but still ended up dropping out. These students were selected because they were more likely to be seriously attempting the course (10 exercises constituted 25% of the course), so their dropout was particularly significant. In addition, there would have been more opportunities for interventions to assist these students, so insights from their behaviour could potentially have a larger impact on similar students in future. We considered the first four weeks of the course since all the exercises were comparable (i.e. structured similarly with a similar time limit to complete them). In total, 3677 students were selected.

For each of these students, we then selected a sample of evenly spaced out exercises from the set of exercises they completed during this time, which would represent their

Jessica Mcbroom, Irena Koprinska and Kalina Yacef "How Does Student Behaviour Change Approaching Dropout? A Study of Gender and School Year Differences" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 643 - 647

behaviour at different stages during their interaction with the course (e.g. when they first began, midway through or just before dropping out). Since all of the selected students had completed at least 10 exercises, we used a sample of 10 evenly spaced out exercises for each student, which was the maximum we could select without including missing data for some students. Since the first and last completed exercise are particularly important when analysing dropout, we wanted to include both of these for each student. As such, we selected exercises using the following process: for each student let $e_1, e_2...e_n$ be the n exercises they completed ($n \geq 10$). Then, define $f(k) = round(1 + \frac{k(n-1)}{9})$ and select $e_{f(0)}, e_{f(1)}, ..., e_{f(9)}$. For example, if a student completed 20 exercises, the selected exercises would be e_1, e_3, e_5, e_7, e_9, e_{12}, e_{14}, e_{16}, e_{18} and e_{20}.

After selecting 10 representative exercises for each student, we then generated features to describe their behaviour during each of these exercises, as shown in Table 1. These features related to the number of times particular events occurred, such as viewing the exercise or failing it, and the timings of these events. Note that these features did not need to be independent due to the clustering technique used.

Table 1: Features used to perform the clustering

Feature	Description
num views	the number of times the student viewed the exercise page
num autosaves	the number of times the student's work was autosaved (this was triggered if they had unsaved work that was not modified for 10 seconds)
num failed	the number of times the student submitted their work for marking but did not pass all automated tests
earliest: view, autosave, failure and pass	the time of the first view, autosave, failure or pass respectively (in seconds, relative to the deadline)
average time between fails	if the student failed the exercise two or more times, the average time between these failures, in seconds
time from first failure to completion	if the student ever failed the exercise, the time in seconds from this point until these passed.

After preparing the features, we applied the temporal hierarchical clustering algorithm DETECT [1] to find clusters of student behaviour that changed over time. This algorithm produces hierarchical clusters defined by decision rules (e.g. a cluster could be all cases where the number of views was ≤ 3 and the number of fails was > 2). To do this, it performs a search over many different options for clustering the data, each time observing the resulting distribution of clusters over time. It then chooses the option that maximises an objective function based on this distribution. In this case, the objective was to find clusters that changed the most between the student's first two sample exercises and their last two. Since the algorithm selects only the best features in the final clustering, the method is robust to dependencies between features. The resulting clusters are shown in Figure 1 and discussed in more detail in the next section.

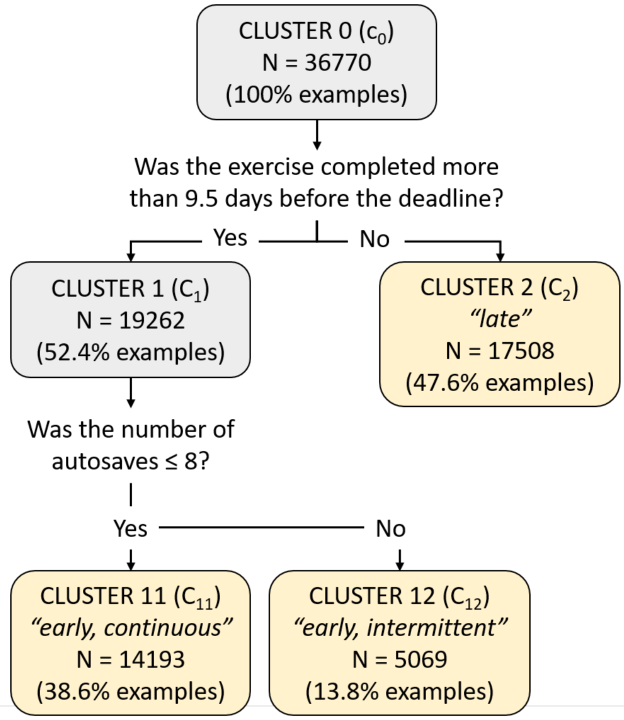

Figure 1: Hierarchical clusters of student behaviour. The cluster names show the hierarchical nature of the clusters. Cluster 2 represents cases where students completed the exercise late. In contrast, Clusters 11 and 12 represent cases where students completed the exercise early and worked continuously and intermittently respectively. This is discussed in further detail in the text.

After producing the clusters, we then analysed the differences in cluster distributions for students from different backgrounds. In particular, we considered the relationship between student gender and school year level on the cluster distributions over time in order to gain insight into equity issues. The results are discussed in the next section.

4. RESULTS AND DISCUSSION

4.1 Behavioural Clusters

The behavioural clusters produced by DETECT are shown in Figure 1. The cluster names indicate the hierarchical nature of the clusters: c_0 at the root contains all examples, c_1/c_2 are mutually exclusive subsets of c_0 and c_{11}/c_{12} are mutually exclusive subsets of c_1. Since there were 3677 students in total and 10 representative examples for each student, this made $N = 3677 \times 10 = 36770$ clustered examples of behaviour in total, with 38.6%, 13.8% and 47.6% in c_{11}, c_{12} and c_2 respectively.

In this work, we focus on the three final clusters, c_{11}, c_{12} and c_2, which we label for convenience as "*early, continuous*", "*early, intermittent*" and "*late*" respectively. We label c_2 as "*late*" since it represents cases where students completed the exercise close to the deadline (i.e. within 9.5 days of it). In contrast, c_{11} and c_{12} are labelled as "*early*" since here students completed the exercise more than 9.5 days before the deadline. In addition, c_{11} is labelled as "*continuous*" because there were ≤ 8 autosaves, and these were triggered when a student with unsaved work paused for more than 10 seconds. As such, a student with a small number of autosaves did not pause very often and worked continuously. In contrast, we label C_{12} as "*intermittent*" since there were a large number of pauses.

Since we used a sample of 10 exercises for each student, this meant that students could move from cluster to cluster over time. As such, in order to understand how the cluster distributions changed over time, we plotted the number of students in each cluster for each exercise, as shown in Figure 2. Note that the total number of students in the graph is constant over time (3677 students). Note also that the exercises are relative to the students, not the course. For example, Exercise 1 and Exercise 10 represent the first and last exercise that each student completed before dropping out, not the first and tenth exercise in the course.

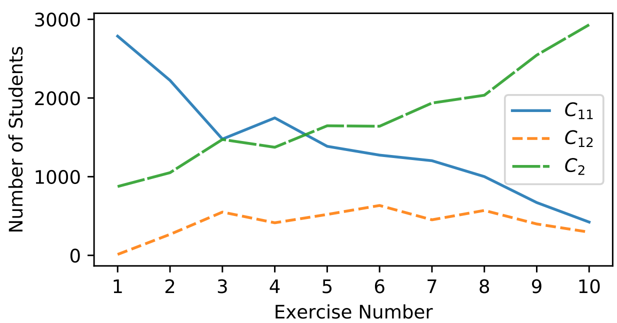

Figure 2: Cluster distributions over time

When submitting their first exercise, most students were in c_{11} (early, continuous), where they completed the exercise more than 9.5 days before the deadline and had ≤ 8 autosaves, indicating that the students worked continuously on the exercise. As such, dropping out students tended to complete the exercises early and continuously when they first began the course. Over time, however, the number of students in the other clusters increased. In particular, by the time they were close to dropping out, most students were in c_2 (late), where they were no longer completing the exercises early. In addition, the increase in c_{12} (early, intermittent) indicates that the students who did complete the exercise early paused more, possibly due to difficulty or distraction.

In summary, by clustering evenly spaced-out samples of student work over time, it is possible to observe how student behaviour develops as students approach dropping out. In the next sections, we will filter these students based on grade level and gender to observe differences in these trends for different student groups.

4.2 School Year Level Differences

In order to analyse the differences between students of different school year levels, we divided students into four groups based on school year, as shown in Table 2. Using the same clusters as before, we then observed the differences in cluster distributions with respect to these groups. The results are shown in Figure 3.

Table 2: Grade groups used in the analysis. N is the number of students in each group who dropped out but completed at least 10 exercises in the first four weeks of the course.

Group	N	Description
Year 11+	263	Senior students in Year 11 or above
Years 9-10	2090	Intermediate students in Year 9 or 10
Years 7-8	1081	Junior students in Year 7 or 8
Primary	238	Primary students in Year 6 or under

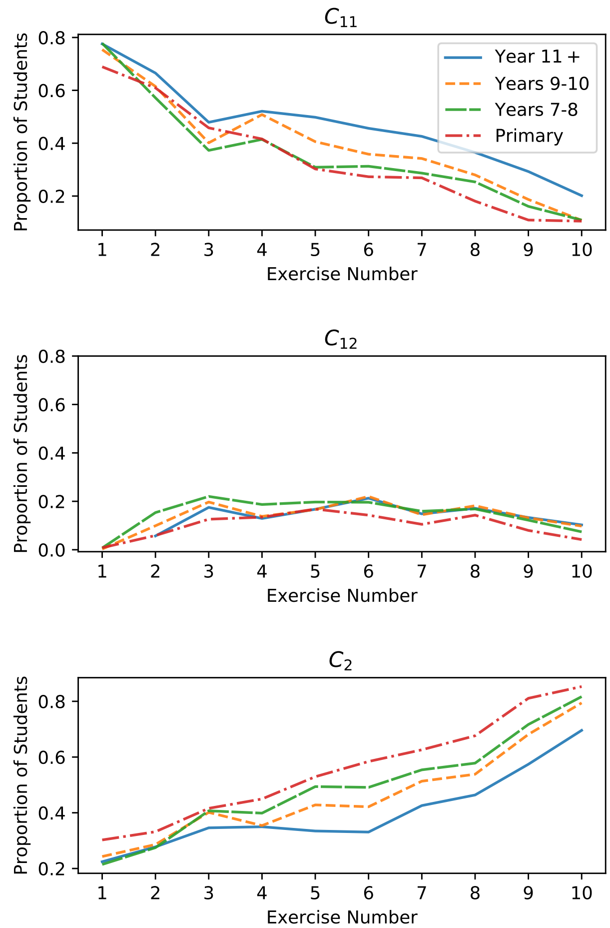

Figure 3: Differences in cluster distributions across school grade groups

From the graphs in Figure 3, one important observation is that increasing age is correlated with a decrease in the proportion of students in c_2 (late), and this difference increases over time. This suggests that younger students tend to complete exercises later than older students as they approach dropping out. This could suggest that younger students are more likely to drop out because they are falling behind and possibly having difficulties with time management, whereas older students may be dropping out for other reasons, such as losing interest or the exercises being too easy.

Another interesting observation is that, while most students were similar at the beginning, older students were more likely to be in c_{11} over time than younger students. Since c_{11} represents behaviour where students complete the exercise early and work continuously, this suggests that older students may have been more organised than younger students, or found the exercises easier immediately before dropping out. This supports the idea that older students may have been dropping out because the exercises were too easy, while younger students may have done so due to difficulty.

These observations are important for informing future course development, since they can provide insight into how courses can be made more equitable. For example, if younger students are more likely to drop out from a course because it is difficult and they are falling behind, then interventions could be developed to help support these students. For example, they could be given extra practice questions or time to complete the exercises. In contrast, if older students were dropping out because the exercises were too easy, then more advanced content or optional extension exercises could be added for these students. This could then help to make the course more equitable by addressing the needs of different student groups.

4.3 Gender Differences

In addition to analysing differences based on school grade, we also considered how student behaviour differed based on gender. In total, we analysed data from 2334 male students and 1124 female students who dropped out after completing at least 10 exercises from the first four weeks of the course. The differences in cluster distributions over time are shown in Figure 4.

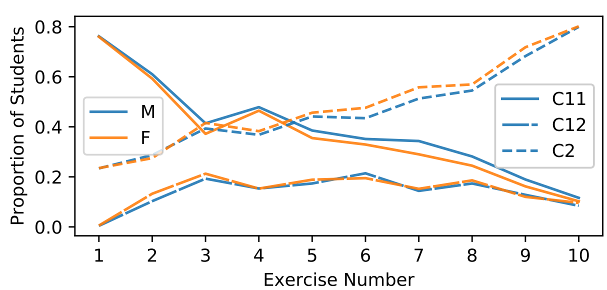

Figure 4: Differences in cluster distributions across gender groups

Interestingly, there is very little difference in the cluster distributions for male and female students over time for this course. This suggests that male and female students behaved similarly with respect to these clusters as they approached dropping out - they started the exercises at similar times and worked roughly as continuously as each other. The dropout rates for male and female students were also similar. Such information is highly valuable for improving gender equity, since it highlights where to focus attention. In particular, instead of comparing the behaviour of male and female students who drop out in order to introduce different types of interventions, perhaps focusing on reducing dropout in general for both groups, or focusing on improving the balance in enrolment rates, could assist in improving gender equity for this course.

5. CONCLUSION

In this paper, we have discussed how the behaviour of students who drop out from a course can be analysed in order to improve equity. In particular, representative samples of work from dropout students can be clustered to identify changes over time. Differences and similarities in trends as students approach dropout can then be observed for different student groups (e.g. male and female students or students from different grade levels). This comparison can lead to insights into the potential reasons for why students dropout, helping to inform further course development to improve equity.

6. REFERENCES

[1] Anonymous. DETECT: A hierarchical clustering algorithm for behavioural trends in temporal educational data. *Unpublished, submitted to AIED2020. Will be made available for reference after double blind review process.*

[2] B. Baker and J. Levin. Educational equity, adequacy, and equal opportunity in the commonwealth: An evaluation of pennsylvania's school finance system. Technical report, American Institutes for Research, 2014.

[3] J. Bayer, H. Bydzovská, J. Géryk, T. Obsivac, and L. Popelinsky. Predicting drop-out from social behaviour of students. In *the 5th International Conference on Educational Data Mining*. International Educational Data Mining Society, 2012.

[4] G. Ben-Zadok, M. Leiba, and R. Nachmias. Comparison of online learning behaviors in school vs. at home in terms of age and gender based on log file analysis. *Interdisciplinary Journal of E-Learning and Learning Objects*, 6(1):305–322, 2010.

[5] S. Chopra, H. Gautreau, A. Khan, M. Mirsafian, and L. Golab. Gender differences in undergraduate engineering applicants: A text mining approach. In *the 11th International Conference on Educational Data Mining*. International Educational Data Mining Society, 2018.

[6] D. A. Fields, Y. B. Kafai, and M. T. Giang. Youth computational participation in the wild: Understanding experience and equity in participating and programming in the online scratch community. *ACM Transactions on Computing Education (TOCE)*, 17(3):1–22, 2017.

[7] P. Gorski. Rethinking the role of "culture" in educational equity: From cultural competence to equity literacy. *Multicultural Perspectives*, 18(4):221–226, 2016.

[8] V. Grigoreanu, L. Beckwith, X. Fern, S. Yang, C. Komireddy, V. Narayanan, C. Cook, and M. Burnett. Gender differences in end-user debugging, revisited: What the miners found. In *Visual Languages and Human-Centric Computing (VL/HCC'06)*, pages 19–26. IEEE, 2006.

[9] J. McBroom, B. Jeffries, I. Koprinska, and K. Yacef. Mining behaviours of students in autograding submission system logs. *the 9th International Conference on Educational Data Mining*, 2016.

[10] J. McBroom, I. Koprinska, and K. Yacef. Understanding gender differences to improve equity in

computer programming education. In *Proceedings of the Twenty-Second Australasian Computing Education Conference*, pages 185–194. ACM, 2020.

[11] M. Wilson. Pre-service teachers' emerging views on educational equity. Master's thesis, Eastern Michigan University, 2019.