Exploring student learning style and procrastination behaviours in an online programming course using machine learning methods

First Author[1][0000-1111-2222-3333] and Second Author[2][1111-2222-3333-4444]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
lncs@springer.com

Abstract. This paper investigates student behaviour in terms of learning style and procrastination tendencies, whether it changes over time and how it impacts the overall course score and drop off rates. We analyse data from a large online Python programming course (n=10,558 students) using machine learning methods. Informed by the Felder-Silverman learning style theory and previous work, we define learning style and procrastination features appropriate for our context, use clustering to find students with common behaviours, and Markov chains to analyse the transitions between clusters. We found that students with different learning styles were able to achieve high scores. The procrastination time and effort clusters provided greater insight into how learning behaviours impact the course performance compared to the learning style clusters. The Markov chain analysis allowed to identify at-risk students likely to drop out early in the course. Our methods allow teachers to better understand student behaviour, identify at-risk students and take remedial actions to improve the student online learning experience and course performance.

Keywords: Student behaviour, learning style, procrastination, assessment, feedback, automatic grading system, clustering, Markov chains.

1 Introduction

Programming is becoming an increasingly important skill in today's technological society and online courses can provide students with programming course content as well as automatic feedback on programming submissions. Whilst these systems allow for efficient grading, they are not always as effective at providing tailored feedback to students. Additionally, teachers can struggle to provide effective feedback to students due to the high number of online enrolments and the time-consuming analysis needed to manually assess each student's program and behaviour. Therefore, there is an opportunity for data driven methods to provide insights into student's learning behaviour and performance in online courses to better personalize student feedback.

In this paper, we examine the behaviour of high-school students participating nationwide in a large online beginners Python programming course. The course provides explanatory content and programming exercises for students to submit, which are graded automatically using testcases. After each submission, students receive feedback about the passed and failed tests and can repeatedly improve and resubmit their code before the deadline. Analysing this rich historical data presents an opportunity to understand student learning behaviours, how they change over time, and their relation to the overall course score and drop off rate. Specifically, in this paper we focus on two aspects of student behaviour: learning style and procrastination tendency. The two main objectives of this study are as follows:

1. To investigate similarities in student learning behaviour defined by student's learning styles and procrastination tendencies, and how these behaviours influence the students' overall course score.
2. To investigate how student behaviour changes over time and how behavioural shifts can impact course drop-off rates.

Overall, this study aims to gain greater insight into student behaviour in terms of learning style and procrastination, and what drives course engagement, to provide more personalised student feedback, remedial actions and better student outcomes.

2 Previous Work

Educational data mining has sought to identify groups of similar students to determine their behaviour and overall course performance. Various methods have been used – clustering [3-5, 8-10], classification [6-7], questionaries and statistical methods [1].

Graf et al. [1] drew from the Felder-Silverman Learning Style Model (FSLSM) [2] to gain insights into the student's navigational behaviour in an online object orientated modelling course. FSLSM classifies students into different learning styles: active/reflective, sensing/intuitive, visual/verbal and sequential/global learners. These styles determine how a student prefers to learn, e.g. sequential learners prefer to start and finish tasks in order, whilst intuitive learners prefer high level thinking and will complete multiple tasks simultaneously. In [1] the student learning style was determined by completing the LSI survey [2]. It was found that students with different learning styles used the course in different ways. For example, detail orientated sensing learners used the forum more than intuitive learners and did so primarily to clarify assessment details.

However, while Graf et al. used surveys to identify the student learning style and then linked the self-reported learning style with the student behaviour, our goal is different – we define features which are informed by the FSLSM theory, are appropriate for our programming context and can be objectively measured, and then use them to find groups of similar students by a data-driven approach, without using surveys.

Conversely, Hooshyar et al. [3] used procrastination features to develop an algorithm to predict student procrastination behaviour. Procrastination was defined by spare time and inactive time. Spare time is the time from when a student submits an assignment until the assignment is due, and inactive time is the time between when an assignment is open and a student views it for the first time. Clustering was used to group the students and the resulting clusters were categorised as non-procrastinators, potential procrastinators and procrastinators. There was a positive correlation between spare time and assignment score and a negative between inactive time and assignment score.

Similarly, Cerezo et al. [4] used procrastination features and effort features to define student learning behaviour in online courses. They defined procrastination using six variables: time spent on quizzes, time spent viewing content, time spent on forums, number of words written on forums, number of days students took to submit tasks and number of other relevant actions. Using clustering they showed that students who spent more time on the quizzes performed better than students who did not and that the final exam score was not related to the time spent on viewing content or lecture slides.

Studies have observed that student behaviour is fluid and changes over time. McBroom et al. [5] captured student's movement between clusters by first clustering students based on behavioural attributes, and then re-clustering based on how many weeks a student's submission belonged to each of the 5 original clusters. Their results concluded behavioural clustering could occur as early as Week 3 of a 13 week-long semester and accurately predict a student's final course score. Additionally, the lowest performing students were consistently in low performing clusters throughout, but the highest performing students moved to an average of 4.2 clusters during the semester.

Other studies, e.g. [6] and [8], investigated whether at-risk students in programming courses can be detected early in the semester. In [6] information from an autograding system, discussion board and assessment marks was used to built a decision tree which achieved a high accuracy in predicting the exam mark mid-semester. In [7], the same data sources were used to define characteristics of high, average and low performing students and predict exam performance. Pereira et al [8] also used a method based on decision trees to predict if students are likely to drop out based on data from the first two weeks of the semester and information from an autograding system, e.g. number of: tests passed, logical lines for each submitted solution and student logins.

In this paper, we investigate student behaviour in terms of learning style and procrastination tendencies in a large beginner programming course by utilizing a data-driven approach. We define features informed by the FSLSM theory and prior work on procrastination, and apply machine learning methods based on clustering and Markov chains to identify distinct student behaviours and how they influence the course score and drop-off rate.

3 Data

The data used in this study is from a beginner Python course with 10,558 participating students. The course consists of 5 weeks of content (slides) and interleaved graded programming tasks (problems).

The slides introduce students to the Python concepts tested in the graded problems. The graded problems are automarked by testcases; students receive feedback about the passed and failed testcases and can correct their code and re-submit. There are 40 graded problems in total, 8 for each of the 5 weeks. The highest possible score in the course is 400 with a maximum of 10 possible marks awarded for each correct problem. Marks are deduced after every 5 incorrect submissions. If a student does not attempt a problem or is unable to pass all testcases, a 0 mark for the problem is awarded. Before submitting their solution for testing, students can run their program without feedback from the testcases and without penalty through the terminal run option.

3.1 Defining FSLSM Learning Style and Procrastination Features

We define two types of features: FSLSM learning style and procrastination features, informed by the Felder-Silverman theory [2] and the work of Cerezo's et al. [4].

FSLMS learning style features. The FSLSM theory describes eight categories of learners (active/reflective, sensing/intuitive, sequential/global and visual/verbal learners). We used six categories – all except visual/verbal, as shown in Table 1. We didn't use visual/verbal due to the difficulty in measuring such preferences in an online course without student self-assessment. For each FSLMS category, we defined a feature which is most indicative of the type of learner and appropriate for our programming course context; the defined features are shown in Table 2. It should be noted that while these features have been informed by the FSLSM learning theory, it is not expected that the clusters will strictly fit these learning styles.

Table 1. FSLSM learning style categories and how they apply to the Python course

FSLSM category	Category description	Hypothesis
Active	Learn by trying things in a hands-on way. Prefers working collaboratively.	Frequent programming submissions.
Reflective	Cerebral learners, learn through thinking and reflecting.	Takes longer on problems to reflect on feedback.
Sensing	Prefers concrete examples, practical learners, high attention to detail.	Completes majority of slide questions.
Intuitive	Prefers abstract concepts and are innovative. Attempts challenges without completing previous challenges.	Works on several problems at once.
Sequential	Learns in linear steps and follows course content sequentially.	High number of sequential slide views.
Global	Learns in large leaps and prefers a high degree of freedom in the learning process. Prefers learning from conclusions and summaries.	High number of course menu views to provide a high-level overview of the course.

Table 2. FSLSM learning style features

FSLSM category	Feature	Description
Active	Number of terminal runs	Number of times a student runs their problem attempt.
Reflective	Average time spent on problem	Average time spent on problems for the week from first to last submission.
Sensing	Percentage of completed slides	Percentage of content slides with all questions answered correctly.
Intuitive	Problems attempted simultaneously	Represented as a binary - 0 for no problems attempted simultaneously and 1 for problems attempted simultaneously.
Sequential	Percentage of sequential slide views	Slide views, immediately before or after the current slide viewed, out of all slides viewed by a student for the week.
Global	Number of course menu views	Number of times a student viewed the course menu overview.

Procrastination features. We define two groups of procrastination features: time related and effort related, described in Table 3 and Table 4 respectively.

Table 3. Procrastination features – time related

Feature	Description
Inactive time	Amount of time [seconds] between when an assignment is released to a student's first problem submission.
Time spent on content	Amount of time [seconds] from first slide view to last slide view.
Avg. time spent on problems	Average time [seconds] between first and last problem run for each problem in a week.
% completed problems	Percentage of problems which passed all testcases.

Table 4. Procrastination features – effort related

Feature	Description
% completed problems	Percentage of problems which passed all testcases.
Number of autosaves	Number of automatic autosaves. Autosaves are generated after periods of inactivity on the website when programming problems.
Number of terminal runs	Number of times a student runs their programming problem.
% slides completed	Percentage of slide problems completed. Slide problems are not scored and are questions relating to the concepts taught each week.
% slides viewed	Percentage of slides viewed by the student for the week.

4 Results

4.1 Similarities in student behaviour in terms of learning style and procrastination tendencies

We used clustering to find groups of similar students based on their behaviour. Three separate K-means clusterings were performed: one on the FSLM features, one on the procrastination time-related features and one on the procrastination effort related features. The number of clusters was selected with the elbow method.

Each feature vector represents a student for one week, therefore each student has a total of 5 vectors, one for each of the 5 weeks.

Prior to the clustering the data was pre-processed by removing missing values and outliers and normalizing all features to the range 0 to 1. This resulted in 15,733 vectors for the FSLSM clustering, and 19,724 vectors for the procrastination clustering.

FSLSM learning style clustering. This clustering yielded k=4 clusters, with the cluster centroids shown in Table 5. Although outliers were removed, for the time features like average time spent on problems, there is still a very large range of values - from 0 to 518,400 seconds. The majority of the time values are on the lower end of this range and has resulted in small centroid values after normalisation. The cluster characteristics can be summarized as follows:

Cluster 0: This cluster is characterized by the lowest number of terminal runs. Students in this cluster also had a low number of slide problems completed, averaging 24%

of slides completed despite viewing at least 72% of the slides on average. The high number of consecutive slide viewership would be consistent with a Sequential learning style. However, the low number of slides completed may indicate that students skimmed through the slides without reading them.

Cluster 1: This cluster has the highest number of slide questions completed successfully with a 84.9% completion rate. It has the highest number of course menu views which indicates Global learning. The students also worked on more than one problem at once which can suggest an Intuitive learning style. They also work the longest on problems, which may be due to students working on problems simultaneously.

Cluster 2: This cluster has a high number of completed slides with 83.1% of slides completed. Whilst this cluster is similar to Cluster 1, the main difference is evident in Cluster 2 having the lowest amount of time spent on problems and that on average students complete one problem at a time. This cluster had the highest number of vectors and consisted of 54.64% of the overall data.

Cluster 3: Cluster 3 has a high number of terminal runs and time spent on problems, indicating high engagement with the problems. However, students were not engaged with the content information and had the lowest slide completion and slide views. They also had high course menu views which could indicate skipping to problems and not engaging with the slide content.

Table 5. FSLSM learning style cluster centroids

	Cluster				
	Full data	**0**	**1**	**2**	**3**
Terminal runs	0.122	0.087	0.199	0.108	0.154
Slides completed	0.671	0.246	0.849	0.831	0.241
Consecutive slides viewed	0.745	0.726	0.727	0.765	0.679
Menu views	0.114	0.074	0.196	0.101	0.134
Avg. time spent on problems	0.087	0.083	0.182	0.049	0.171
Simultaneous problems	0	0	1	0	1

Time procrastination clustering. This clustering also yielded k=4 clusters, with the cluster centroids shown in Table 6. These clusters can be described as follows.

Table 6 Time procrastination cluster centroids

	Cluster				
	Full data	**0**	**1**	**2**	**3**
Inactive time	0.043	0.077	0.044	0.024	0.031
Time spent on content	0.264	0.166	0.088	0.701	0.357
% completed submissions	0.876	0.349	0.980	0.968	0.976
Avg. time spent on problems	0.054	0.012	0.003	0.006	0.005

Cluster 0 – High procrastination: Students had a high inactive time which meant students waited longer to start problems after they were released. They also had the lowest number of completed submissions with only 34.9% of submissions completed. The high amount of time spent on problems can indicate these students struggled with the content which is why they had a low submission completion rate. Therefore, we can consider this group to have high procrastination tendencies.

Cluster 1 – Low procrastination: This cluster shows little evidence of procrastination because of the lowest amount of time spent on content and problems whilst achieving a 98% problem completion rate. This could indicate this group had prior programming experience which allowed them to complete submissions quickly without heavy reliance on content or problem feedback.

Cluster 2 – Low procrastination: Students started problems shortly after the release time and had the highest amount of time spent on content. This cluster also had a high amount of completed submissions with 96.8% of submissions completed.

Cluster 3 – Moderate procrastination: Students had a high number of completed sub-missions and a medium amount of time spent on content.

Effort procrastination clustering. The elbow method indicated k=5 clusters and the cluster centroids are shown in Table 7. These clusters are described below.

Table 7 Effort procrastination cluster centroids

		Cluster				
	Full data	**0**	**1**	**2**	**3**	**4**
Number autosaves	0.089	0.998	0.089	0.039	0.081	0.067
Terminal runs	0.057	0.064	0.058	0.025	0.050	0.048
Completed problems	0.876	0.983	0.973	0.276	0.461	0.947
Slides completed	0.647	0.839	0.267	0.292	0.727	0.176
Slides viewed	0.842	0.962	0.864	0.412	0.823	0.429

Cluster 0 – High problem effort, high content effort: Cluster 0 has the highest amount of engagement on all measures, with 97% of submissions completed, 83% of slides completed and 96% of slides viewed. The high number of terminal runs and autosaves also suggests students spent a significant amount of effort on the course problems.

Cluster 1 – High problem effort, medium content effort: This cluster is characterised by a low number of slides completed despite having a high number of slides viewed with 86% of slides viewed. This suggests students in this cluster may skim through the content without engaging with the practice questions.

Cluster 2 – Low problem effort, low content effort: Cluster 2 has very low activity on all features and likely didn't engage with the course.

Cluster 3 – Low problem effort, high content effort: This group of students had a low number of completed submissions, averaging 46% problem completion. However, they did have high slide viewership and the highest slide completion rate of 72%.

Cluster 4 – High problem effort, low content effort: These students had a low number of autosaves and terminal runs but had a high completion rate of submissions with an average of 94% of problems completed. The high number of completed problems but low number of autosaves and terminal runs indicates these students may have been quicker and more efficient at solving problems, perhaps because of prior programming knowledge. This cluster can also be characterized by a very low amount of slide content interactions, with only 17% of slide questions completed and 42% of slides viewed.

4.2 Impact of learning style and procrastination on student's overall course score

To assess the impact of the learning style and procrastination tendency on the student overall course score, the clusters are further analysed in terms of student score distribution for each cluster. The score was discretized into 4 ranges: from low [0,100] to high [301-400].

FSLSM learning style. Fig. 1 shows the score distribution across the FSLSM clusters as a percentage. It is evident that each cluster has a similar distribution of students of all performance types including low, medium and high performing students. Therefore, the FSLSM features would not be a strong predictive measure of the final student score, consistent with the literature on the FSLSM learning theory [1] which explains that all learning groups are able to perform highly in online learning courses.

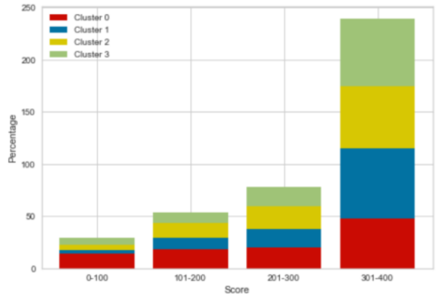

Fig. 1. Percentage of FSLSM clusters in each score range

Procrastination time. From the scores in Fig. 2, we can see that all clusters except Cluster 0 show a similar distribution of masks in the low, medium and high range.

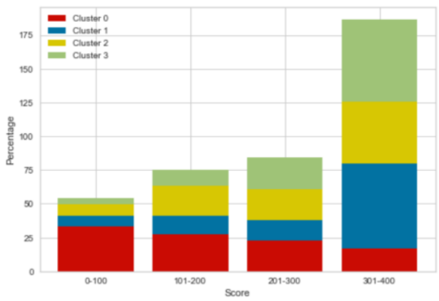

Fig. 2. Percentage of procrastination time clusters in each score range

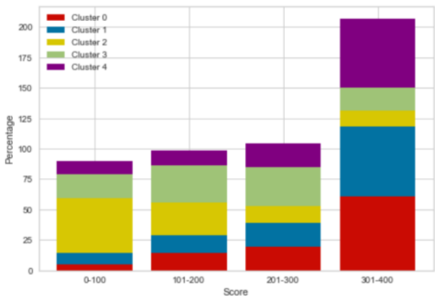

Fig 3. Percentage of procrastination effort clusters in each score range

Cluster 0 (high procrastination) has the lowest problem completion rate, with only 16% of scores achieved in the 301-400 range. Whilst Cluster 2 (low procrastination) and Cluster 3 (moderate procrastination) had similar centroids, Cluster 3 had a higher percentage of students in the high score range, with 61% of students obtaining 301-400 score, compared to 46% of students in Cluster 2. The predominant difference between these two cluster's centroids is students in Cluster 3 spent less time on content and slides. This can indicate Cluster 3 had students with previous programming knowledge who didn't have to rely on the slide content to achieve high marks.

Additionally, Cluster 3 (moderate procrastination) had 39% of students in the 0-300 range which is similar to Cluster 1 (high procrastination) which had 36% of students in this range. Therefore, despite Cluster 3 having higher procrastination tendencies, Cluster 3 students achieved similar results to students with low procrastination tendencies.

Procrastination effort. Fig. 3 shows that three clusters dominate the highest score range (301-400): Cluster 0 (high problem effort, high content effort) with 61% of students achieving this score, Cluster 1 (high problem effort, medium content effort) and Cluster 4 (high problem effort, low content effort) both have 56% of students obtaining high scores. Cluster 0 has strong problem and content effort metrics which likely contributed to higher marks.

Students with a low content engagement were still able to achieve high marks with Cluster 1 and Cluster 4 having low levels of slide completion but high course scores.

4.3 Changes in student behaviour over time and impact on drop off

Markov chains were used to analyse the student learning behaviour changes over time and show the student's transitions from the clusters in Week 1 to the clusters in any of the proceeding weeks. Two new groupings were added to indicate student drop-off rates: a No Attempt cluster for students who did not engage in the course for the week and a Partial Attempt cluster for students who completed some features measured.

FSLSM learning style Markov Chain. The Markov chain for the FSLSM clusters (Fig. 4) has 44.5% of vectors in Cluster 0 in Week 1 transitioning to No Attempt in the next weeks. This is the highest percentage transition to No Attempt of any of the FSLSM clusters, indicating that Cluster 0 students are of high risk of dropout.

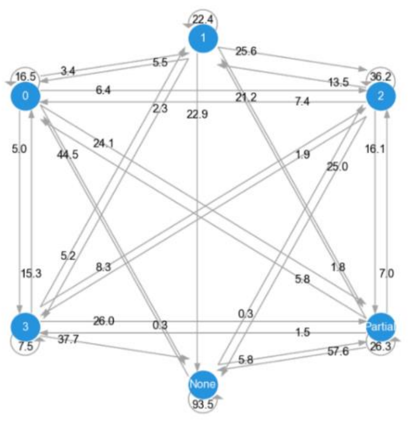

Fig. 4. FSLSM learning style Markov chain

Procrastination time Markov Chain. Fig. 5 shows a high drop off rate for Cluster 0 (high procrastination) with 92% of students in Cluster 0 in Week 1 moving to the No Attempt cluster in the proceeding weeks. Cluster 2 (low procrastination) and Cluster 3 (moderate procrastination) show a relatively smaller number of drop off rates but a high number of transitions to other clusters. Both Cluster 2 and 3 are similar, where students start questions early shortly after they are released, but with Cluster 2 spending almost double the average time on content than Cluster 3, therefore showing student's spending more or less time on content throughout the course.

Procrastination effort Markov Chain. In Fig. 6, Cluster 2 (low problem effort, low content effort) and Cluster 3 (low problem effort, high content effort) have students with the highest risk of dropout with 95% and 81% of students moving to No Attempt in previous weeks respectively. This is higher than the students who began in the No Attempt group in Week 1. The high drop-out rate is surprising since Cluster 3 had 72% of slides viewed on average, which would normally indicate high course engagement. In this example, the Markov chain highlighted at-risk students who may have gone unnoticed due to high content engagement.

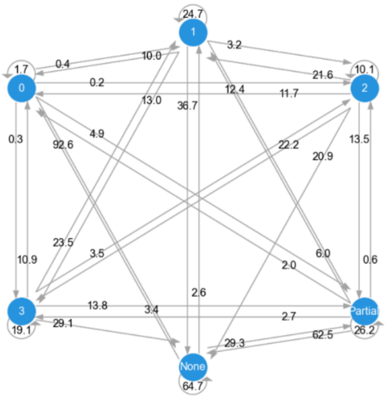

Fig. 5. Procrastination time Markov chain

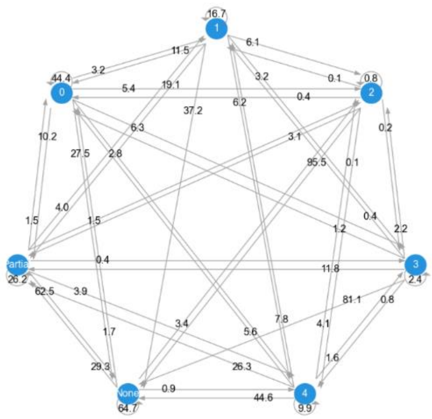

Fig. 6. Procrastination effort Markov chain

Changes in student learning week by week. Student's cluster transitions between consecutive weeks highlight student learning behaviour changes. Fig. 7 and Fig. 8 show the transitions for the procrastination effort clusters to the Partial Attempt and No Attempt clusters. Previously in Fig. 6, Cluster 2 (low problem effort, low content effort) consisted of students with the highest risk of course disengagement. However, calculating the transitions to No Attempt for Cluster 2 between Week 1 and Week 2 (Fig. 7), showed a 91.5% drop-off rate. Therefore, these students would require intervention as early as the first week of the course to prevent student disengagement.

Calculating transitions between consecutive weeks can also indicate course difficulty. From Fig. 7 we can see that students had a decrease in partial attempts from Week 3 to Week 4, indicating they found Week 4 easier and were able to complete more problems. Students from Cluster 2 and Cluster 3 who saw an increase in partial attempts in Week 4 (Fig. 7) had a drop in No Attempt in the same week (Fig. 8). Therefore, we can assume these students who were previously not attempting questions were able to attempt questions in Week 4 due to the easier content. However, the rise of Partial Attempts in Week 5 could indicate that Week 5 tasks were particularly challenging, or students were disengaging from the course towards the final week.

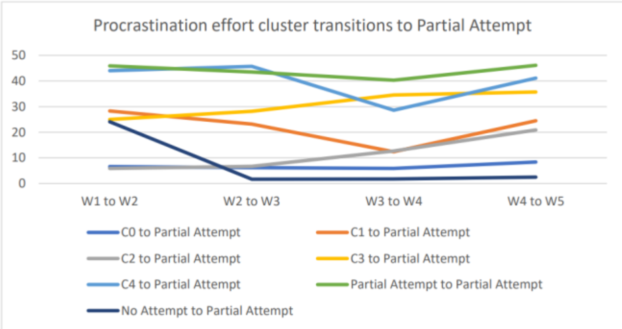

Fig. 7. Procrastination effort clusters to Partial Attempt clusters between consecutive weeks

Conversely, the rates of No Attempt are relatively consistent within each cluster, as shown in Fig. 8, and after Week 2, there is an almost 100% retention rate in the No Attempt cluster. Additionally, procrastination effort Cluster 2 which had a reduced percentage of No Attempt representation in Week 5 (Fig. 8), also had an increased rate of Partial Attempt representation for that week (Fig. 7). Therefore, as question difficulty increases some clusters such as procrastination effort Cluster 2, may be more prone to partial course completion rather than completely disengaging from the course.

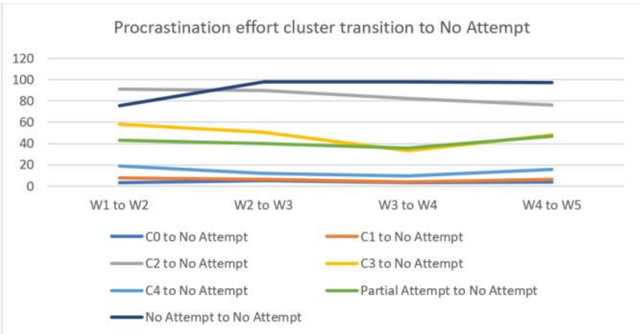

Fig. 8. Procrastination effort clusters to No Attempt clusters between consecutive weeks

5 Conclusion

This paper investigates student behaviour in terms of learning style and procrastination tendencies, how it changes over time and how it impacts the overall course score and drop off rates in an introductory programming course. Informed by the Felder-

Silverman learning style theory and previous work, we define three sets of features appropriate for our context: FSLSM, time procrastination and effort procrastination, and use them to find clusters of similar students. The FSLSM learning style clusters showed a similar distribution of low, medium and high performing students, highlighting that students with different learning styles were able to achieve high scores. Conversely, both the procrastination time and procrastination effort clusters had distinct distributions of student's scores. The low procrastinators and the 'high problem effort, high content effort' clusters contained the highest achieving students. Therefore, although the FSLSM clusters can identify different learning styles, the procrastination time and effort clusters provide greater insight into how learning behaviours impact course performance.

The Markov chain analysis allowed to identify at-risk students likely to drop out of the course. It showed that students with low problem effort were most at-risk as early as after Week 1. Surprisingly, this group included both students with low and high content effort (slides viewed and completed). Our results also showed that for all clusters, once a student disengages for a week, it is unlikely they will reengage in any proceeding weeks. Therefore, Markov chain analysis and comparison with previous cohorts can inform teachers when to provide assistance to at-risk students to prevent drop off rates.

Our methods allow teachers to understand student behavior in terms of learning style and procrastination, identify groups of poorly performing students and when these students are most at-risk of disengaged from the course. These insights can help teachers tailor remedial action and feedback to improve student learning outcomes and course performance.

References

1. Graf, S., Lui, T. C., Kinshuk.: Analysis of learner's navigational behaviour and their learning styles in an online course. Journal of Computer Assisted Learning 26, 2, 116-131 (2010)
2. Felder, R. M., Silverman, L. K. (1988). Learning and teaching styles in engineering education. Engineering Education, 78, 7, 674-681.
3. Cerezo, R., Esteban, M., Sanchez-Santillan, Nunez, J. C.: Procrastinating behaviour in computer-based learning environments to predict performance: A case study in Moodle. Frontiers in Psychology, 8, 1403 (2017)
4. Hooshyar, D., Pedaste, M., Yang, Y.: Mining educational data to predict student's performance through procrastination Behaviour. Entropy, 12, 1-24. (2019)
5. McBroom, J., Jeffries, B., Koprinska, I., Yacef, K.: Mining behaviours of students in autograding submission system logs. In: International Conference on Educational Data Mining, 159-166 (2016)
6. Koprinska, I., Stretton, J. and Yacef, K.: Students at risk: detection and remediation, In: International Conference on Educational Data Mining, 512-515.
7. Koprinska, I., Stretton, J. and Yacef, K.: Predicting student performance from multiple source. In: Int. Conference on Artificial Intelligence in Education, 678-681 (2015)
8. Pereira, F. D., Oliveira, E., Cristea, A., Fernandes D., Silva, L., Aguiar, G., Alamri, A., and Alshehri, M.: Early dropout prediction for programming courses supported by online judges. In International Conference on Artificial Intelligence in Education, 67-72 (2019)