Visualisation of key splitting milestones to support interventions

Martin Hlosta, Jakub Kocvara, David Beran, Zdenek Zdrahal
Knowledge Media institute, The Open University
martin.hlosta@open.ac.uk

ABSTRACT: The paper presents an approach to help staff responsible for running courses by identifying key milestones in the educational process, where the paths of successful and unsuccessful students started to split. By identifying these milestones in the already finished courses, this information can be used to plan the interventions in the next runs. This is achieved by finding the earliest time when the differences in behaviour or key performance metrics of unsuccessful students start to become significant. We demonstrate this approach in two case studies, one focused on a course level analysis and the latter on a whole academic year. This suggests its generic nature and possible applicability in various Learning Analytics scenarios.

Keywords: Learning Analytics, Visualisation, At-Risk Students, Intervention support

1 INTRODUCTION - SETTING THE SCENES

Identifying students at-risk of failing either the course or the whole qualification is a very topical issue of Learning Analytics. Further analysis of reasons why the student is lagging behind may suggest interventions that guide him/her to the successful completion of the course (Jayaprakash et al., 2014). Usually, two sources of data are available: data about the student and data about the course. Student data include their demographics, their study history, and activities within the course. The data related to the course are the study plan i.e. study materials, dependencies between different study resources, time allocated to each task, assignment to be completed by the student to prove that he/she has mastered the expected content and progression rules, which define criteria of student's success or failure in the course. Often student data from previous presentations of the course are available and machine learning techniques can be used for developing predictive models (Wolff et al., 2014). This problem specification applies both to classroom-based and to distance education. One of the typical issues is selecting a moment in the course to use the predictive model for interventions so that the predictions are accurate yet early enough for at-risk students to get back on track. Howard et al., (2018) selected this point based on manual inspection of decrease of the error between week 4 and 5.

In this paper, we offer a different view on the learning analytic tasks. As mentioned above, by assessing whether the student satisfied the course progression rules, we distinguish two groups of students: those who pass and those who fail. In fact, we may extend this dichotomy by an additional group of students who have not met all progression rules, but there is a reasonable chance that they can complete the missing requirements in the future and finish the course. For example, the student has not acquired all credits required to successfully pass, but then he/she has earned enough key credits and therefore may be allowed to continue and complete the missing credits in the next years. Consequently, we may distinguish three groups of students denoted as fail, continue and pass.

By analysing already completed course presentations, we have noticed that there are "points" in the study plan where the "homogeneous" cohort divides into two or three of these groups. This split can be verified by a suitable statistical test and without early intervention, it is usually persistent to the end of the presentation. Once the student starts losing pace with the study plan, the gap is likely to grow and eventually, the student may resign and fail. The same situation can apply to the continue/pass split. Such points are usually identified by the manual analysis. For example, Simpson (2004) identified different withdrawal routes of students by showing the proportion of students not submitting their assignments using the `river` diagram and only very few of them returning to submit the next ones. In the same paper, he suggested that different withdrawal types might indicate different interventions with students. Coffrin et al. (2014) presented state transition diagrams for students who completed the course and those who did not. Using these diagrams, users can observe transitions of students between the assessments and the differences between the completed and non-completed groups, although these differences are not stated explicitly. Teasley (2018) mentions this identification of important points in courses when discussing what it means to do learning analytics, referring to finding a “point of no return” when poorly performing students are likely not to succeed in the course.

The recent survey analysing 52 papers in Visual learning analytics found that most of the work focuses on Understanding Collaboration and Instructional Design, with analytics on students for instructors being most prevalent (Vieira et al., 2018). Some of the work focuses on time changes, especially students progressing in the course, e.g. a simple approach in (Breslow et al., 2013) using line plots to show different activity types used in different weeks. Chen Y. et al., (2016 October) helps to explain the behaviour of students in different clusters based on their predictions and actual results. Moreover, some papers support the identification of interesting points in time. Chen Q. et al. (2016) visualises the peaks in the videos from the clickstream to better design the videos in the future.

The aforementioned approach in (Corfin et al, 2014) can be used to identify points when students start to drop out and also the one in (Hlosta et al., 2014) to spot the typical patterns of students before the first assignment leading to failure.

We have demonstrated, that if the pattern of characterising that the students are approaching split point is identified before the split became persistent and the instructors intervene, the student retention or successful completion can be dramatically improved. Identifying the split points will be demonstrated and visualised in the following sections. This builds on our previous work (Zdrahal et al., 2016) and also (Wolff et al., 2014) and its aims to provide a generalisable and visual approach for early phases of Learning Analytics process.

To conclude, there is work that highlights the identification of the milestones to support the intervention. Moreover, some of existing research in visual analytics can help with this identification but to the best of our knowledge, there is a gap in automatic identification and visualisation of these milestones during the learning process. Also, the existing papers focus on a limited context, such as MOOCs, closed classroom. Providing that relevant data are available, our work aims at generalizing across different learning contexts.

First, we provide the description according to 5 questions from the workshop proposal call for paper. Then, we present two case studies from different learning scenarios showing the visualisations and concluding with the further work.

2 THE APPROACH

2.1 What kind of data is being visualized? What tools were used to clean up the data?

The visualisation expects data from the university system with the final result of students in either a course or a whole academic year. The approach expects partial measurements of students' progress towards achieving the learning goal recorded as events in time. These usually include assessment scores, optionally weighted by their importance. In addition, data of any recorded student activities can be used.

The pre-processing has been performed using SQL and Python with its common libraries for manipulating data, i.e. Pandas and Numpy and SciPy for statistical evaluation[1].

2.2 For whom is the visualization intended?

The visualisation has been designed for staff responsible for running the courses, potentially for researchers in Learning Analytics. Realising the key milestones, the course directors receive hint when to plan the interventions or where the design of the course might be updated. The users are not expected to be experts in visualisation, they should be familiar with the structure of the course.

2.3 Why: what is the goal of the visualization? What questions about the data should it answer?

The goal is to support the identification of **important milestones** in a course or academic year using visualisation. It should answer questions such as: When does the difference in measured value between successful and unsuccessful students start to be statistically significant? When is a convenient time to make interventions for poor performing group provided that a similar pattern of student behaviour will prevail the next run? What is the best splitting value of the measured value between the groups of students in time?

We expect the approach to be used for initial course analysis before building a machine learning algorithm that might be more complex and resource expensive. On a higher level of abstraction, the usage workflow consists of the following steps:

1. Identify the indicator of students' progress, e.g. assessment score, number of credits. Visualising this should provide the first insight of where the students start to split.

2. Select a behavioural characteristic of students, e.g. number of clicks, time spent in the VLE and use the visualisation on a more granular level.

2.4 How is data visualized and why? Tools, libraries, data formats used for technical implementation? What workflows and recipes can be used to develop the visualization?

The data is visualised using the line plot representing the median for each performance group, with the variance of the captured metric between the 25th and 75th percentile. The variance is shown using the same colour with added transparency level. This was a preferred variant over boxplots as they would make the graph more challenging to read, especially when shown for more performance groups. The first identified milestone is visualised using the vertical line through the whole graph,

[1] Pandas - https://pandas.pydata.org/, Numpy - www.numpy.org, SciPy - https://www.scipy.org/

with bold style in the region between the two medians, where the difference was measured. The black horizontal line denotes the best split between the performance groups.

The measurements are taken in various regular time intervals during the whole duration of either a course or the academic year, typically days or weeks. The approach provides retrospective analysis, so the results of students are required to assign students in the performance groups.

Python with its data manipulation libraries and matplotlib[2] for visualisation have been utilised. Similar results might have been also achieved with R or with some javascript library.

The approach consists of four steps:

1. Preparing the common data input format - This includes extracting the source data of student events and converting them in a time-sliced data table, where all students have records of all available measurements, i.e. not only when they change.

2. Identifying the important milestones - starting at the beginning of the measured period, the algorithm continuously examines the difference between the successful and unsuccessful students. In each time slice, a statistical test is performed to detect if the difference in the observed metric is statistically significant. If the conditions for unpaired t-test are met (normality of both group distributions and homoscedasticity), it is used as a preferred variant. Otherwise, the Wilcoxon rank sums test is used.

3. Best Splitting values - starting in the identified milestone, for each following time the best splitting values in the measurement is computed by minimising the error of that split, i.e. proportion of wrong predictions to all predictions. It represents the quality of the predictions that would be achieved if this splitting point was used to classify students into good and at-risk student groups.

4. Visualising the lines, variance bands, early milestone and splitting points. The graph can be enhanced by adding manually annotated events, e.g. the start of Christmas break, dates of the assessments, etc.

2.5 How has the approach been evaluated or how could it be evaluated?

The quality of each milestone split can be evaluated in terms of statistical significance. The approach counts with taking the data distribution into consideration. In each point, we can also compute the error of the split that is made based on this factor.

The goal of the visualisation is to convey a clear message to either researchers or course designers to help them understand when the intervention should happe. Understanding this and acceptance of this information can be viewed as one of the key evaluation strategies. As the next step we want to run a user study with 10-15 participants and various types of roles - i.e. tutors, course designers and researchers. We plan to use a combination of a questionnaire designed by the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, 2003) combined with open ended questions. The UTAUT uses four constructs (performance expectancy, effort expectancy, social influence and facilitating conditions) to explain the users' technology acceptance and use. The open ended questions will focus on providing information about current actions around the identified points in time, about perceived importance and potential interventions that might be possible to plan. We have conducted a similar procedure during the case study 1 in the past but in more informal way.

[2] Matplotlib - https://matplotlib.org/

2.6 Encountered problems and pitfalls during the visualization process?

One of the problems was examining zero values, i.e. if zero measurement should be included/excluded from the statistical tests. The other challenge was making the approach generic enough to cope with various x-axis unit, in our case days relatively counted from the start of the course or calendar dates.

3 TWO CASE STUDIES

3.1 Classroom-based university – Progress through the academic year

A faculty from a classroom-based university with the face to face teaching had poor progression rates of their first-year students. The students acquire credits by completing one-semester long courses prescribed in the study plan. Acquiring credits is stored as events. Typically, there are 6-8 courses per semester, the number of credits earned in the course depends on its difficulty and importance for the study program. Based on the number of credits earned at the end of the two-semester academic year, a student falls into one of the four groups (fail, fail-winter, continue, pass). Groups "pass" and "continue" progress to next study year, students in the "fail" group are deregistered, "fail-winter" fail even before the end of the winter semester. The trajectory of students is shown in Figure 1. We are interested in the difference between the "fail" and "continue" groups. There are 943 students in total, i.e. 245 pass, 198 continue, 54 fail, 446 fail-winter. The number of credits within the groups is not normally distributed, neither the homoscedasticity has been satisfied, hence Wilcoxon sum rank test was used. The groups start splitting before the Christmas break, meaning that students who have not collected enough credits at that time are already at risk. By the end of the winter exam period, the inter-group differences are very noticeable. The flat part that follows, corresponds to the period of lectures in the summer term usually without credit-earning exams. Next opportunity for earning further credits is in the summer exam period. Though the winter and summer exam periods are well-defined, the examiners may offer a few "early exam terms" up to 4 weeks before the start of the exam periods. It is visible in Figure 1, that the "pass" students take this opportunity more often than the students of the "continue" group. Moreover, Figure 1. shows, that the students in the "fail" group do not earn significant (if any) credits before the start of exam period.

This visualisation triggered a conversation with the faculty management and led to designing a precaution intervention strategy, reminding this to all the students and then repeating this to the ones that haven't collected enough credits. To the great surprise of university academics and ourselves, this has resulted to the increase of students progressing to the second year by 49%. Specifically, comparing with the best year so far, 49% of students expected to fail progressed to another year. The letter of recognition of the faculty dean is available on 3.

[3] The letter of recognition of the faculty dean available on our website: https://analyse.kmi.open.ac.uk/resources/documents/letter_of_recognition.pdf

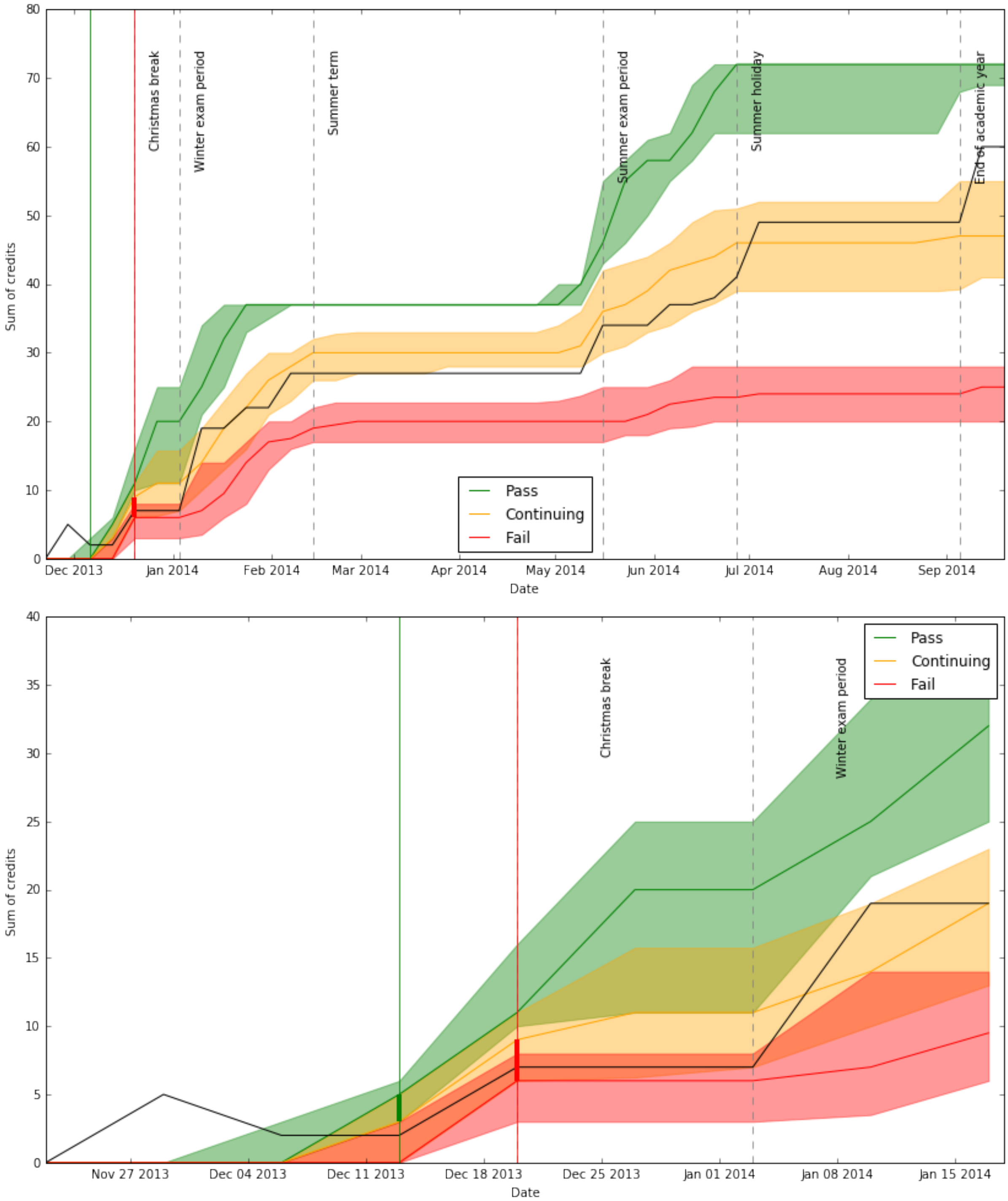

Figure 1: a) case study - students achieving credits in face to face university b) zoomed view highlighting the first statistically significant difference between the groups.

3.2 Distance education course

The second example comes from a publicly available OULAD dataset from the Open University (Kuzilek, 2014). Using this dataset allows better reproducibility of this approach. We selected a level-one course that is fully online - EEE/2014J. The rest of the courses in 2014J can be found in the GitHub repository [4]. Students gain a score after submitting their assessments, which enable them to pass the course. Their final result is either Distinction, Pass, Fail or Withdrawn. Moreover, student

[4] Github repository with figures – https://github.com/hlostam/milestone-vis

sum of clicks per days is captured. We used the weighted assessment score to account for the importance of the assessment. Figure 2 shows that the first important difference is just after the submission of the first assessment with the best splitting point for score 13.

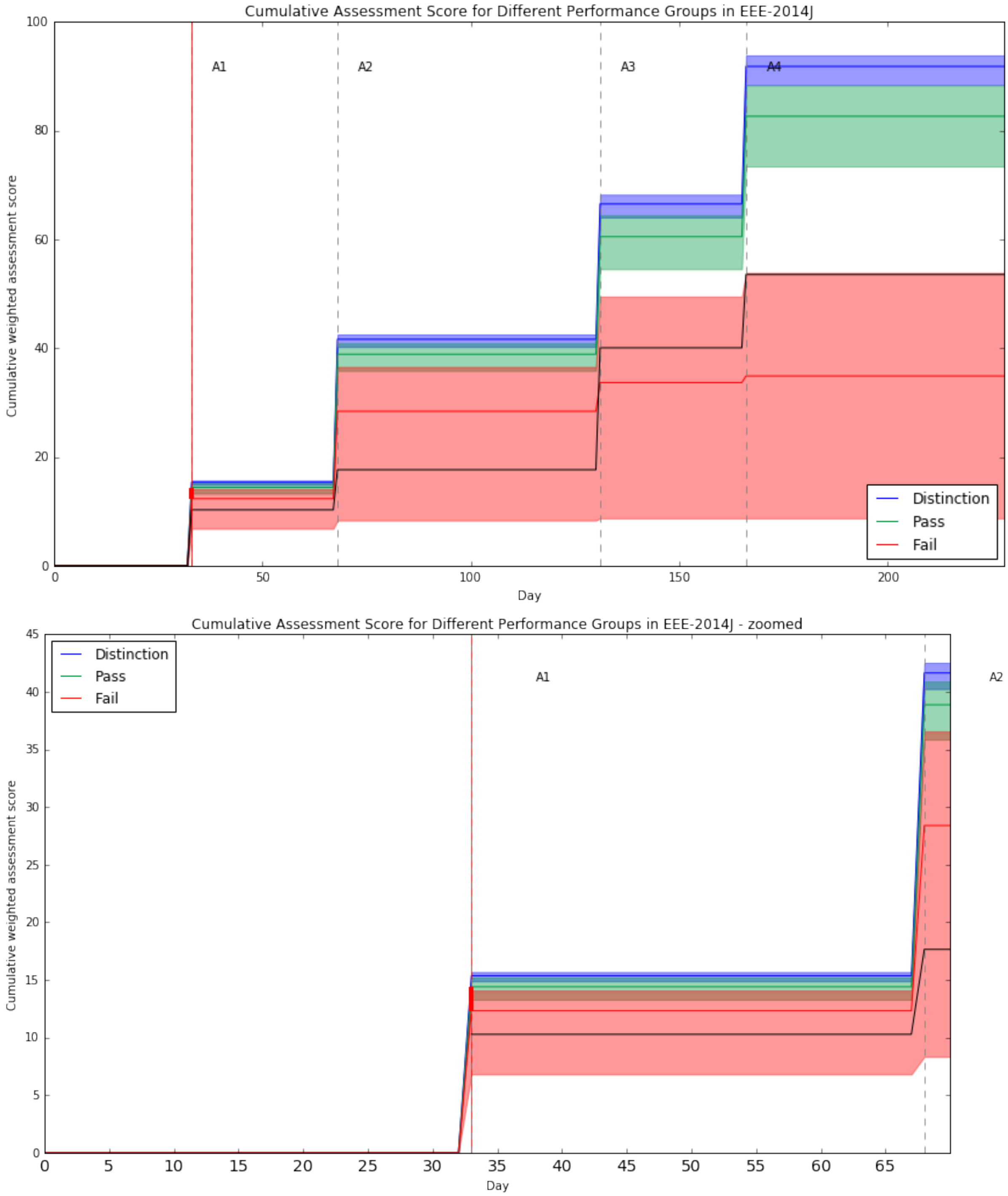

Figure 2: a) Distance education course and acquiring the weighted assessment score b) the detail highlighting the first statistically significant difference between the groups.

This might justify focusing on intervening even before the first assessment. Focusing on more detailed student online behaviour, Figure 3 shows that for the sum of the clicks in VLE, the first observed difference between both Failed and Passed and between Passed and Distinction is in the first day of the course.

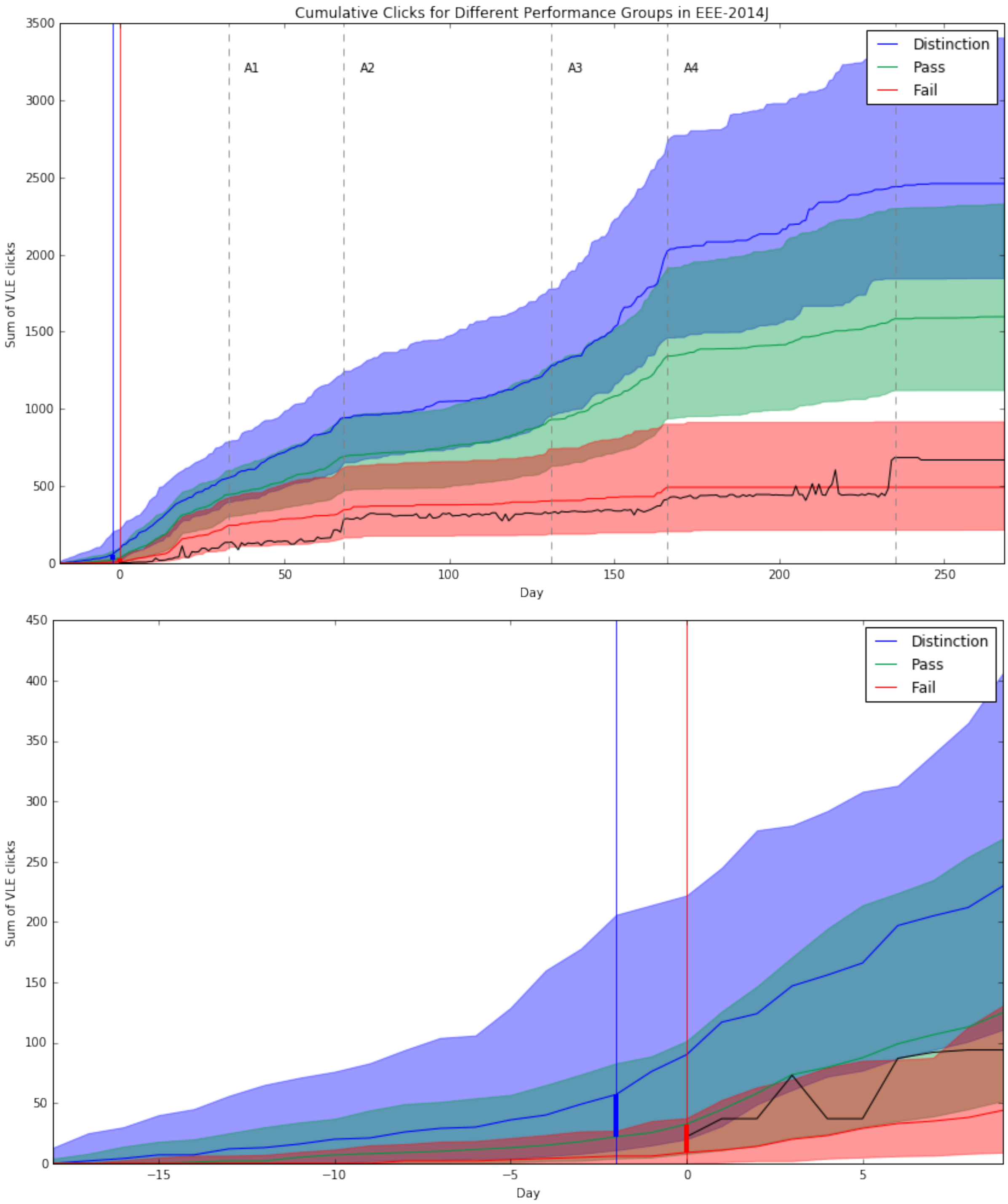

Figure 3: a) Number of clicks in distance education course for different performance groups b) the detail highlighting the first statistically significant difference between the groups.

It should be mentioned that the identified key milestones do not mean that potential predictive models would be accurate enough to split between the successful and unsuccessful students. It gives only a signal that starting this point, the differences between the behaviour of these two groups in terms of the measured variable started to be statistically significant.

4 CONCLUSIONS AND FURTHER WORK

Until now we have deployed this framework in three case studies, two of which we share here. In the case of the conventional university, the usefulness and impact of the approach have been demonstrated by successfully improving the retention by about 49% in two consecutive years. In both cases we compare results with the lowest retention achieved in 2013/14 i.e. before the described predictions and interventions have been deployed. Our current focus is to include the study history of the students, which might help to identify groups of students where interventions might have higher impact.

REFERENCES

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom research into edX's first MOOC. *Research & Practice in Assessment*, *8*, 13-25.

Chen, Q., Chen, Y., Liu, D., Shi, C., Wu, Y., & Qu, H. (2016). Peakvizor: Visual analytics of peaks in video clickstreams from massive open online courses. *IEEE Transactions on Visualization & Computer Graphics*, (10), 2315-2330.

Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., & Qu, H. (2016, October). DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. In *Visual Analytics Science and Technology (VAST), 2016 IEEE Conference on* (pp. 111-120).

Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). Visualizing patterns of student engagement and performance in MOOCs. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 83-92). ACM.

Hlosta, M., Herrmannova, D., Vachova, L., Kuzilek, J., Zdrahal, Z., & Wolff, A. (2018). Modelling student online behaviour in a virtual learning environment. In: *Machine Learning and Learning Analytics Workshop at The 4th International Conference on Learning Analytics and Knowledge (LAK14)*, 24-28 Mar 2014, Indianapolis, Indiana, USA.

Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*, *37*, 66-75.

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, *1*(1), 6-47.

Kuzilek, J., Hlosta M., Zdrahal Z. (2017). Open University Learning Analytics dataset In: *Sci. Data* 4:170171 doi: 10.1038/sdata.2017.171.

Klerkx, J., Verbert, K., Duval, E. (2017). Learning analytics dashboards. In: *Handbook of Learning Analytics*, Chapt. 12 Society for Learning Analytics Research.

Simpson, O. (2004). The impact on retention of interventions to support distance learning students. *Open Learning: The Journal of Open, Distance and e-Learning*, *19*(1), 79-95.

Teasley, S. D. (2018). Learning analytics: where information science and the learning sciences meet. *Information and Learning Science*.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, *122*, 119-135.

Wolff A., Zdrahal Z., Herrmannova D., Kuzilek J., Hlosta M. (2015). Developing predictive models for early detection of at-risk students on distance learning modules In: *Machine Learning and Learning Analytics Workshop at The 4th International Conference on Learning Analytics and Knowledge (LAK14)*, 24-28 Mar 2014, Indianapolis, Indiana, USA.

Zdrahal, Z., Hlosta, M., & Kuzilek, J. (2016). Analysing performance of first year engineering students. In: Data Literacy for Learning Analytics Workshop at Learning Analytics and Knowledge (LAK 16), Edinburgh, UK.