Suhonen, S. J. & Kinnari-Korpela, H. (2019) Using Moodle data for early warning of dropping out. In Nagy, BV., Murphy, M., Järvinen, H-M. & Kálmán, A. (ed.) Varietas delectat… Complexity is the new normality. SEFI 47th Annual Conference Proceedings, p. 1113 - 1122.

URL: https://www.sefi.be/conference/2019-budapest-isbn-978-2-87352-018-2/

Using Moodle data for early warning of dropping out

S. J. Suhonen[1]
Tampere University of Applied Sciences
Tampere, Finland

H. Kinnari-Korpela
Tampere University of Applied Sciences
Tampere, Finland

Conference Key Areas: Gender, inclusion and ethics, Integrated learning environments for the digital native learners

Keywords: Learning analytics, Moodle, dropout

ABSTRACT

Dropout rates are high in engineering education throughout Europe. According to interviews, the main reasons were life situation, learning or studying difficulties, wrong field of studies and health problems. The interviewees would have hoped more support and guidance during their studies and 70 % of the dropouts were very interested to continue their engineering studies. Sometimes support at the right time to right students might solve the problem and studying would continue normally.

How to find those students that are at risk of dropping out? Single teachers might notice absences, but they don't have the whole picture. Study counsellors need to rely on the course completion data on the transcript of records, which is always delayed weeks or months in comparison to the studying actions and possible problems in it. Therefore, not even the study counsellors have a real time view to the student progress.

Many higher education institutions use digital learning management systems (LMS) like Moodle to deliver their online, blended and face-to-face courses. Students leave digital footprints on the platform and this tells about studying habits. Therefore, LMS data has the potential to show decrease in learning activity well before it becomes visible elsewhere. In this study, two engineering student groups are followed during one academic year covering both simultaneous and consecutive courses. With the data, a simulation is run to raise an alarm if a student is at risk of dropping out. These alarms are then compared with the real dropping out information of the groups.

[1] Corresponding Author
S. J. Suhonen
sami.suhonen@tuni.fi

1 INTRODUCTION

Education is largely seen as a key element for a successful life, better income and usefulness as a citizen in society. In the Europe 2020 strategy [1], one of the goals is to increase the percentage of higher education graduates to 40 % in the age group 30-34 years old people. There is still work to be done, as can be seen in Table 1. The table shows course completion rates in some of the EU countries according to the European Commission main report "Dropout and Completion in Higher Education in Europe" [2].

Table 1. Course completion rates in higher education in some EU countries according to 2015 report. [2] * The data from Germany is from 2005

Country	Course completion rate
UK	82 %
Denmark	81 %
France	80 %
Germany	77 %*
The Netherlands	72 %
Czech Republic	72 %
Poland	62 %
Norway	59 %

This paper focuses on the dropouts in engineering degree programs in Finnish universities of applied sciences. Some insight to the situation regarding graduation and dropout can be seen using "Vipunen" data portal. It is education administration's reporting portal operated by The Ministry of Education and Culture and the Finnish National Agency for Education. According to the data available in "Vipunen" [3], the percentage of graduates in ICT engineering after 6 years of studying is 43 %. In engineering the average is 54 % and the average of all fields of studies in universities of applied sciences is 61 %. Figure 1 shows the percentage of engineering students who has earned a degree in their first field of studies (blue line) and in any field of studies (dotted red line) as a function of study year. The nominal graduation time is 4 years in engineering. Clearly, it is worth putting an effort on the helping of the students that struggle in their studies and are in danger of dropping out. It is no wonder that European Commission in their main report suggests institutions to monitor pathways of individual students to identify students at risk of dropout. [2].

The dropout problem is not new. Already in the 50's there were studies about the influence of student's personality factors and socio-economic background on the academic success and models were built to understand the nature of dropping out [4-6]. It was pointed out that one should differentiate the dropouts according to the reasons: failure in studies, lack of motivation, voluntary change of studies, health reasons etc.

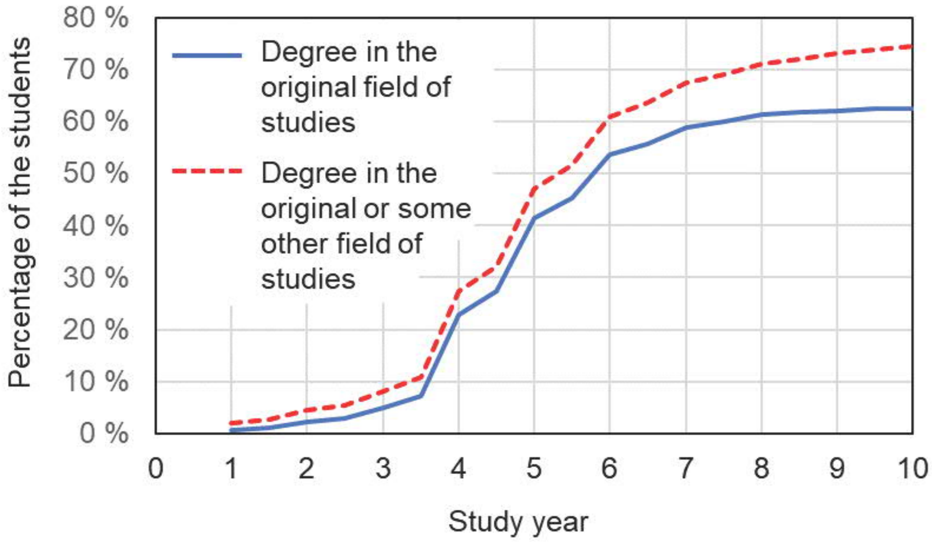

Fig.1. The percentage of engineering students who has earned a degree in their first field of studies (blue line) and in any field of studies (dotted red line) as a function of study year. The data is the average over years 2002-2015 and concerns Finnish universities of applied sciences. Data is from "Vipunen" [3].

Nowadays higher education institutions use digital learning management systems (LMS) like Moodle to deliver their online, blended and face-to-face courses. Students leave digital footprints on the platform and this tells about studying habits. There are studies in which LMS data, like number of logins, number of posts to discussion forums, material openings etc., are used to build predictive models of student success [7-9].

When trying to offer guidance and help to the students at risk of dropping out, the first challenge is the same irrespective of the underlying reason: How to find those students that are at risk? Some universities have built student dashboards and signal-lights for visualizing course attendance. The results for example from Purdue University are very promising: the course attendance has increased and the student feedback about the system is rather good [10]. In the universities of applied sciences in Finland, there is not yet such a system. Therefore, some teachers might notice student absences, but they don't have the whole picture. Study counsellors need to rely on the course completion data on the transcript of records, which is always delayed weeks or months in comparison to the studying actions and possible problems in it. Therefore, not even the study counsellors have a real time view to the students' progress.

This paper shows a case study of using Moodle log file data to recognize the students that are at risk of dropping out. The aim is not to build any predictive model of student success but rather find as easy way as possible to find those students who are at risk. After recognizing these students, it is time for interventions, help and guidance, of course, but those are not on the scope of this study.

2 COURSES AND DATA IN THIS STUDY

2.1 Courses

Two bachelor's level engineering degree programs in Tampere university of applied sciences were chosen for this study: laboratory engineering and ICT engineering. One student group from each program was investigated. The students had started their studies fall 2016 and their studying activity was tracked from the beginning of their studies until December 2018.. From their curricula, several courses, both consecutive and simultaneous, were selected for this study. All these courses are taught face-to-face but they utilize also Moodle using a blended learning approach. The aim was not to include all possible courses to the study, but rather see if it is possible to find students at risk using some key courses as indicators. The selected key courses are mostly mathematics and physics – topics that students sometimes find difficult. The degree programs, selected courses of which the Moodle data is gathered, and course timings are shown in Table 2.

Table 2. The degree programs, courses and their timings.

Program	Courses	Timing
Laboratory engineering	Mechanics	1st year autumn
Number of students: 36	Analytical Chemistry	1st year autumn
	Basics of Measuring and Reporting	1st year autumn
	Working English for Engineers	1st year autumn
	Physics lab. Course	1st year spring
	Electrostatics and Magnetism	1st year spring
	Oscillations and Wave physics	1st year spring
ICT engineering	Geometry and Vector Algebra	1st year autumn
Number of students: 47	Functions and Matrices	1st year autumn
	Differential Calculus	1st year spring
	Integral Calculus	1st year spring

2.2 Moodle data

When a student clicks an object on Moodle page at the main level of course's page hierarchy, a time stamp data is generated to the log file. Therefore, many of the actions a student take, generate a digital footprint. However, learning takes place naturally also outside Moodle. And courses differ from each other quite a lot in structure and Moodle usage. The time stamp data contains the basic information of the action: who did it, when did it happen and what was the clicked object. Perhaps the simplest possible quantity describing student's overall activity is the number of these log events. This can be easily derived from Moodle. It can be argued that the number of log events doesn't tell about learning and somebody can just play to be active by clicking randomly around in Moodle page. These claims are true, of course. Nevertheless, previous studies have shown that the number of log events really tells about the studying habits and student's intention to study [11-13]. On the other hand, the lack of log events certainly tells that a student is not doing the intended learning tasks in Moodle.

In this study, the sum of log events of courses listed in Table 2 was calculated for each student. Due to sickness, holidays, work or other reasons, students sometimes are absent from the university for short periods of time. In this study, a three-week sliding average of the log events was calculated to even out the effect of such short absences on the log data. The weeks included in the data were 35-49 for autumn semester and 2-17 for spring semester. This sliding average of the sum of log events works as a warning signal: if it goes to zero for somebody, that student is hypothetically at risk of dropping out and an intervention would be launched.

The data in this study looks backwards to 2016 and thus the analysis works as a simulation to find out how this chosen parameter works. An example of the data is shown in Fig. 2 showing the Moodle activity in individual courses and also the three-week sliding average for A) a normally studying student and B) a student who has difficulties in studies.

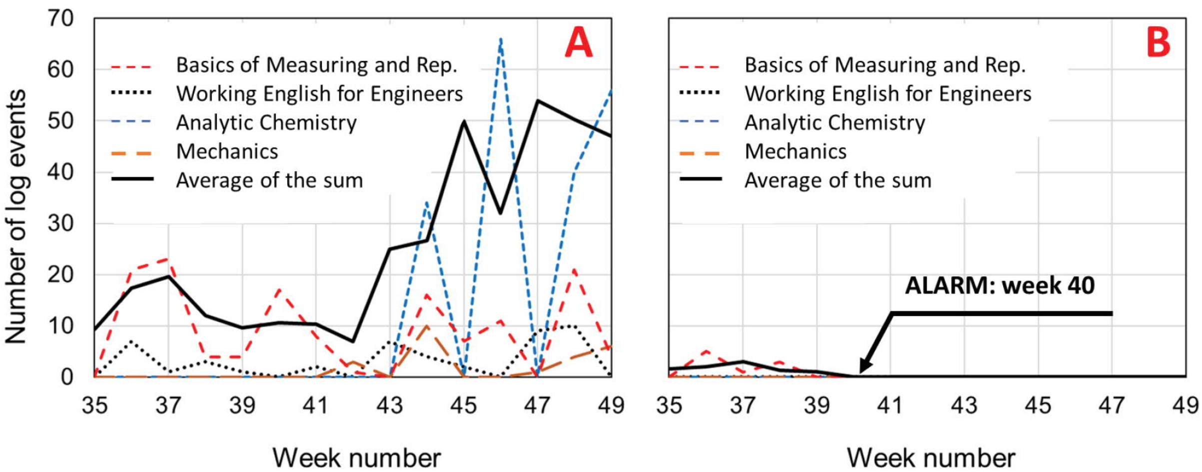

Fig. 2. Moodle activity in individual courses and the calculated three-week sliding average for A) a normally studying student and B) a student who has difficulties in studies. In the simulation student B would have risen an alarm at week 40, 2016.

The courses differ in their Moodle usage and the individual course curves thus differ from each other quite strongly as can be seen in Fig. 2. Nevertheless, the sliding average of the sum of the log events doesn't go to zero for the normally proceeding student. The figure 2B shows data of a student who has difficulties in studying. Clearly, even the overall activity is small, and the sliding average goes to zero at week 40. For this student, the simulation rises an alarm at week 40, 2016. All the students are analysed in the same way and categorized either as "studies normally" or "At risk, alarm: week X"

3 RESULTS

The results of the simulation are summarized in Table 3. The number of alarms in laboratory engineering student group was 10 (26 % of the students). When the progress of these students was examined, it turned out that seven of those 10 were actually dropped out of the university, one was temporarily non-attending and two were studying normally. The coverage of the alarm was 100 % since all the actual dropouts of the group launched an alarm in the simulation. On the other hand, there were two false alarms.

In the ICT engineering group, the number of alarms was much higher, 22 (47 %). This was expected since it was well known that the dropout problem was the biggest in ICT among engineering degree programs. In this "at risk" group, 10 students had actually dropped out, one was temporarily non-attending, one raised a false alarm and five were incoming students. These five were coming from outside the original student group and they had changed either institution or study program. Because of their prior studies they didn't participate in the courses chosen for this simulation and were thus categorized as "at risk". Then there are still five others, who are marked in the table with "other reason". It turned out that these students were present in the university, but they were not actually studying, and their amount of credit units earned was very low in comparison to the nominal amount of credits. This is visualized in Fig. 3. Again, the coverage of the alarm was 100 %. In this student group, of those 22 who rise an alarm, "only" 10 had dropped out. Of the remaining 12 alarms 11 were relevant since those students were special in some way: either non-attending, institution changers or slow in studies. One of the alarms was totally false.

Table 3. The summary of the simulation results.

Laboratory engineering student group:	
Number of students:	37
Studies normally	27
Students a risk:	10
Dropped out	7
Non-attending	1
False alarms	2
Warning coverage:	100 %
ICT engineering student group:	
Number of students:	47
Studies normally:	35
Students a risk:	22
Dropped out	10

Non-attending	1
Incoming students	5
Other reason	5
False alarms	1
Warning coverage:	100 %

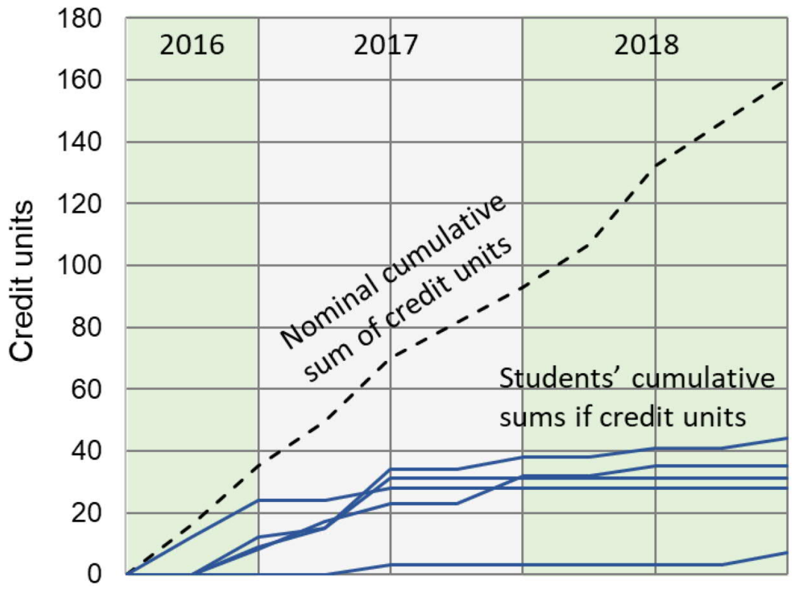

Fig. 3. The cumulative sum of credit units from autumn 2016 to the end of 2018 for those five students ("other reason"), who triggered an alarm.

As an output, the simulation/analysis gave those weeks when students were identified to be at risk of dropping out. Of those students, who had actually left university, the drop-out date was also taken from the transcript of records. The time differences of these dates are shown in Fig. 4. for all 17 students.

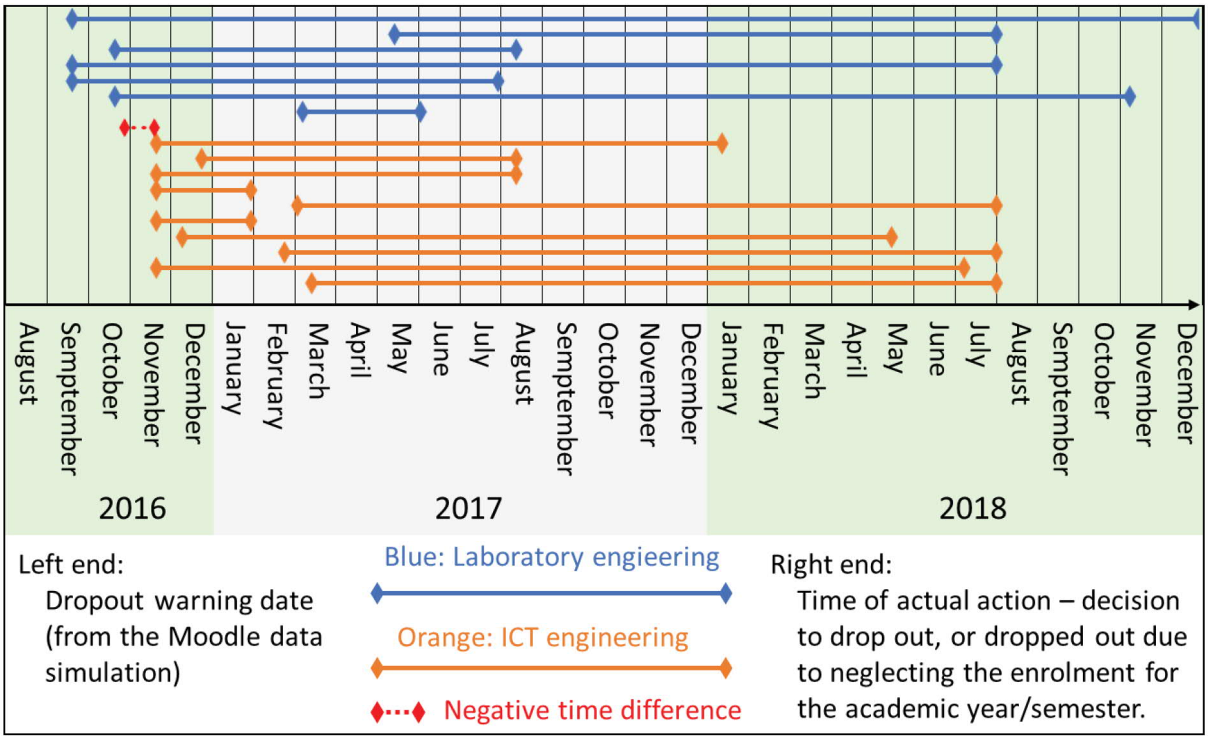

Fig. 4. The time differences between the alarm date and actual dropout date for all 17 students who triggered an alarm and had later dropped out.

It is clear that the studying activity goes down to zero in Moodle much earlier than the students take any actions or decide to stop studying and leave university. Such actions typically happen/occur before these students are detected in the university guidance process. Some of them "drift away" due to neglecting the enrolment for the semester. In this study, the decrease in studying activity was detected in Moodle data averagely 13 months earlier than actual drop-out happened.

4 DISCUSSION AND CONCLUSIONS

The results clearly suggest that the number of events in Moodle log can be used as a measure of student's activity and to identify those at risk of dropping out. The decrease in activity can be noticed much earlier in Moodle than in transcript of records or as student's actual dropout decision. The time difference in this case study was 13 months on average from the "at risk"-warning to actual dropping out.

Dropout rates are high in engineering education throughout Europe. In Tampere university of applied sciences, students were interviewed to find out the underlying reasons. According to interviews, the main reasons were life situation, learning or studying difficulties, wrong field of studies and health problems. The interviewees would have hoped more support and guidance during their studies and 70 % of the dropouts would have been interested to continue their engineering studies. This emphasises the need to find those students who are at risk as soon as possible.

Then appropriate interventions need to be launched for example by the study counsellor or tutor. However, the interventions were not in the scope of this study.

5 ACKNOWLEDGMENTS

The APOA project and the Finnish Ministry of Education and Culture are acknowledged for supporting this study. The aim of the project is to study and pilot the use of learning analytics in the Universities of Applied Sciences (UAS) in Finland. The project is funded by the Finnish Ministry of Education and Culture.

REFERENCES

[1] European Commission (2010), EUROPE 2020 - A European strategy for smart, sustainable and inclusive growth.

[2] European Commission main report (2015), Dropout and Completion in Higher Education in Europe.

[3] Vipunen - Education Statistics Finland. The education administration's reporting portal. Retrieved 20th April 2019. Available at: https://vipunen.fi/en-gb/

[4] Grace, H. (1957) Personality factors and college attrition. Peabody Journal of Education, vol. 35, pp. 36-40.

[5] Iffert, R. E. (1958) Retention and withdrawal of college students. Washington, D.C.: U. S. Department of Health, Education and Welfare, Bulletin No. 1, 1958.

[6] Tinto, V. (1975) Dropout from Higher Education: A Theoretical Synthesis of Recent Research. Review of Educational Research, Vol 45, No. 1, pp 89-125

[7] Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). Predicting student performance: an application of data mining methods with an educational web-based system. In 33rd Annual Frontiers in Education, 2003. FIE 2003. (Vol. 1, pp. T2A-13). IEEE.

[8] Kabakchieva, D. (2013). Predicting student performance by using data mining methods for classification. Cybernetics and information technologies, 13(1), 61-72.

[9] Osmanbegovic, E., & Suljic, M. (2012). Data mining approach for predicting student performance. Economic Review: Journal of Economics and Business, 10(1), 3-12.

[10] Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. Proceedings of the 2nd international conference on learning analytics and knowledge (pp. 267-270). ACM.

[11] Suhonen, S. (2016). Students’ experiences of different types of (distance) learning. In Proceedings of The Online, Open and Flexible Higher Education Conference, Rome, Italy, pp. 380-389.

[12] Suhonen, S. (2016) Learning Analytics View to Students' Homework Activity in Engineering Physics. INTED2016 Proceedings, pp. 3998-4005.

[13] Suhonen, S., & Tiili, J. (2015) Students’ Online Activity on a Fully Online Introductory Physics Mechanics Course, Proceedings of SEFI2015 43rd Annual Conference, Orleans, France