Student Drop-out Modelling Using Virtual Learning Environment Behaviour Data

Conference or Workshop Item

Student drop-out modelling using Virtual Learning Environment behaviour data

Jakub Kuzilek[1], Jonas Vaclavek[1], Viktor Fuglik[1,2] and Zdenek Zdrahal[1,3]

[1] CTU in Prague, CIIRC, Jugoslavskych partyzanu 1580/3, 160 00 Prague, Czech Republic
[2] Charles Univ, Fac Edu, Magdaleny Rettigove 4, 116 39 Prague, Czech Republic
[3] Open University, KMi, Walton Hall, MK7 6AA, United Kingdom
jakub.kuzilek@cvut.cz

Abstract. With the rapid advancement of Virtual Learning Environments (VLE) in higher education, the amount of available student data grows. Universities collect the information about students, their demographics, their study results and their behaviour in the online environment. By applying modelling and predictive analysis methods it is possible to predict student outcome or detect bottlenecks in course design. Our work aims at statistical simulation of student behaviour in the VLE in order to identify behavioural patterns leading to drop-out or passive withdrawal i.e. the state when a student is not studying, but he has not actively withdrawn from studies. For that purpose, the method called Markov chain modelling has been used. Recorded student activities in VLE (VLE logs) has been used for constructing of probabilistic representation that students will perform some activity in the next week based on their activities in the current week. The result is an instance of the family of absorbing Markov chains, which can be analysed using the property called time to absorption. The preliminary results show that interesting patterns in student VLE behaviour can be uncovered, especially when combined with the information about submission of the first assessment. Our analysis has been performed using Open University Learning Analytics dataset (OULAD) and research notes are available online[1].

Keywords: Student Drop-out, Modelling, Virtual Learning Environment, Markov Chains.

1 Introduction

In the past decade, higher education experiences a massive boom of ICT based education. At present, educators and students extensively use Virtual Learning Environments such as Moodle platform [1]. The ICT based education is further boosted by the introduction of Massive Open Online Courses (MOOCs) platforms such as Coursera [2]. With all these platforms the amount of information about students grows. The possibilities of student data usage for improvement of the education have been investigated in over 200 studies in past years [3].

[1] https://bit.ly/2JrY5zv

In 2014 Hlosta et. al. [4] proposed two methods for activity analysis: General Unary Hypothesis Automaton and Markov chains. The first method produces set of rules that describe the data. The second generates state transition probabilities from state to state, which represents chances that student change behaviour based on his previous behaviour. The main disadvantage of both methods is the complexity of achieved results.

The idea of previously mentioned work is further extended by Okubo et. al. [5]. The authors employed the Markov chain-based method using data from Kyushu University and provided the method as a Moodle analysis module.

Later on, Davis et al. [6] employed Markov chains in the analysis of MOOC data from edX and Coursera courses with over 100,000 students.

Our research focused on the exploration of student behaviour using VLE logs in order to uncover behaviour leading to withdrawal or passive withdrawal of the student. For that purpose, we employed Markov chain modelling [7] on behavioural data available in Open University Learning Analytics dataset (OULAD) [8], which contains the data from a Moodle-like system used at the Open University[2]. Furthermore, the previously used approach [4] has been simplified and the state space of student activities was reduced to 7 possible states, which will be further discussed in section 3

2 Data

The OULAD [8] contains information about 32,593 students visiting 22 Open University courses in years 2013 and 2014. The Open University is largest distance learning institution in the United Kingdom with more than 170,000 students. The typical course has one or more assignments, final exam and has the length of approximately 9 months. OU uses the Moodle-like platform (VLE) to deliver content to students. Usually, course VLE provides a plan of activities for the whole course and it is recommended for students to follow it. For more details see the original paper [8].

The dataset includes data about both students and courses. We focused on data from one course-presentation namely course *FFF* and presentation *2014J*. The course is focused on STEM subject more than 1/3 of the students withdrawn during the semester.

In the following text logs of student VLE activities, the information about first assessment submission and the date of de-registration of the student from the course will be used.

3 Methods

In this section, the process of Markov chain model construction will be presented. This can be divided into a transformation of log data to student state data and Markov chain construction itself.

[2] http://www.open.ac.uk/

3.1 Transforming VLE logs to states

At first, VLE logs were aggregated on a weekly basis. Next, by combining with course plan (available in OULAD dataset) the student state for every study week has been estimated as follows.

Each activity in VLE has been classified as planned or not based on the course plan. Next, summarization of the planned and non-planned activities for each student and each week has been computed. From the summarized data weekly states have been estimated. Student state in planned activities can fall into the three possible categories: student did nothing (*0*), student did something from the plan (*E*), and student did everything from the plan (*A*). Similarly, unplanned activities can be categorized to: student did nothing (*0*), and student did something out of the plan (*E*). When combined 6 possible states emerged: *00*, *E0*, *A0*, *0E*, *EE*, *AE*. For example, state *00* means that student did nothing at all – nothing from a plan and nothing from other (not planned) activities.

Finally, state $Withdrawn$, which represents the fact that student has actively withdrawn from studies, has been added to the set of states resulting in seven possible states, in which every student can be in each week.

3.2 Markov chains

For the construction of Markov chain, we will consider simplifications in order to reduce the problem to the most simple one: 1) the length of a course is infinite; 2) the probability of transition from state in one week to state in another week does not change over time (homogeneity condition of Markov chain); 3) student cannot return to a course when withdrawn; 4) the probability of changing the student state depends only on current week (this is called Markov property [7]). All above leads to the construction of so-called homogeneous absorbing Markov chain [7].

Markov chain is specified by the set of states S. In our case, these are defined by student states $S = \{00, E0, A0, 0E, EE, AE, Withdrawn\}$. From the set of states S and weekly student states, we can construct the state transition matrix $\boldsymbol{P}$, where the entry in i-th row and j-th column represents the probability p_{ij} that a student moves from state s_i in current week to state s_j in following week. In addition, the computed transition matrix is reorganized in order to be in the canonical form [7].

Clearly, state $Withdrawn$ is absorbing state, that means the student (the process) in this state cannot leave it. Since this state is of the interest we can analyse the resulting transition matrix of Markov chain by means of absorption time [7], which represents the average number of weeks needed to end up in the $Withdrawn$ state for the student starting in state s_i.

4 Results

The Markov chain has been constructed for the three cases: 1) the whole cohort of students; 2) students who submitted the first assessment; 3) students who did not submit the first assessment. Following subsections present the results.

4.1 Markov chain of the whole cohort

As depicted above, the transition matrix of the whole cohort of students has been constructed. Before the estimation of transition probabilities, the students with states containing a small number of samples ($E0$ and $A0$) have been filtered out. The resulting model has 5 states and its transition matrix follows:

$$\boldsymbol{P_1} = \begin{matrix} \\ 00 \\ 0E \\ EE \\ AE \\ Withdrawn \end{matrix} \begin{matrix} \begin{matrix} 00 & 0E & EE & AE & Withdrawn \end{matrix} \\ \begin{pmatrix} 0.66 & 0.29 & 0.02 & 0 & 0.02 \\ 0.13 & 0.75 & 0.09 & 0.01 & 0.01 \\ 0.05 & 0.45 & 0.37 & 0.11 & 0.01 \\ 0.03 & 0.24 & 0.63 & 0.09 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

Since the complexity of graphical representation is high, we decided to work with the transition matrix only. From the matrix $\boldsymbol{P_1}$the vector of absorption times t_1 is then computed: $t_1 = (78 \quad 81 \quad 81 \quad 82)^T$.

4.2 Markov chain of submitting students

Same as in case of the whole cohort the students with states containing a small number of samples ($E0$ and $A0$) have been filtered out. Then the students who did submit the first assessment has been selected and the transition matrix was constructed:

$$\boldsymbol{P_2} = \begin{matrix} \\ 00 \\ 0E \\ EE \\ AE \\ Withdrawn \end{matrix} \begin{matrix} \begin{matrix} 00 & 0E & EE & AE & Withdrawn \end{matrix} \\ \begin{pmatrix} 0.62 & 0.35 & 0.02 & 0 & 0.01 \\ 0.13 & 0.77 & 0.08 & 0.01 & 0 \\ 0.06 & 0.59 & 0.33 & 0.01 & 0 \\ 0.01 & 0.031 & 0.59 & 0.07 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

Based on the transition matrix the absorption times vector is computed: $t_2 = (142 \quad 145 \quad 146 \quad 146)^T$.

4.3 Markov chain of non-submitting students

Lastly the Markov chain for those who did not submit the first assessment has been computed. The students with states containing a small number of samples ($E0$, $A0$ and AE) have been filtered out and the transition matrix has been constructed:

$$\boldsymbol{P_3} = \begin{matrix} \\ 00 \\ 0E \\ EE \\ Withdrawn \end{matrix} \begin{matrix} \begin{matrix} 00 & 0E & EE & Withdrawn \end{matrix} \\ \begin{pmatrix} 0.95 & 0.03 & 0 & 0.02 \\ 0.51 & 0.41 & 0.03 & 0.05 \\ 0.38 & 0.38 & 0 & 0.25 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

From the matrix $\boldsymbol{P_3}$ the absorption times vector has been computed: $t_3 = (50 \quad 47 \quad 37)^T$.

5 Discussion of results

When observing resulting transition matrix $\boldsymbol{P_1}$of the whole student cohort, one can notice that the probability of student withdrawing from the studies is twice larger for students with no activity in VLE than for student with at least some activity in VLE.

Another interesting observation is that students with no planned activity tend to do nothing from the plan next week (states 00 and $0E$) and those who did nothing will do nothing next week in 2/3s of cases. On the other hand, students doing everything from the plan do not tend to withdraw their studies and with high probability will do at least something from the plan next week. Also, they will interact with the VLE with probability 0.96. If we compare the average time to withdraw from the course (time to absorption) students starting in state 00 (doing nothing in the first week) has the lowest time to withdraw.

When we split the data to students who did submit and who did not submit the first assessment, which has been proven to be a good predictor of student success [9], we can observe dramatic changes in the structure of a Markov chain. First, students who submitted the first assignment (transition matrix $\boldsymbol{P_2}$) do not tend to withdraw from studies if they have at least minimal contact with VLE. Second, those who did everything planned tend to do at least something from a plan in the next week. Finally, only those who submitted the first assessment, but then did nothing in VLE have a small probability to withdraw.

What is much more interesting that students who did not submit the first assessment (transition matrix $\boldsymbol{P_3}$) but still interacted with the planned activities in the VLE, tend to withdraw from the studies with probability 0.25. Those, who did not submit the first assessment and did nothing in the VLE tends to do nothing next week (the probability is 0.95). They can be understood as passive withdrawal students– they do nothing, do not actively withdraw and fail the course at the end.

What is important is the fact of homogeneous Markov chains meaning transition probabilities are not changing over time. Of course, it is important to say that in real situation transition probabilities changes over time, but the model called non-homogenous Markov chain is much harder to interpret. For that purpose, we stayed with the simple model, which can be further extended.

6 Conclusion

In this paper, we employed Markov chain modelling for the analysis of student behaviour in VLE and its influence on student drop-out from the course. For the purpose of reproducibility, we used OULAD dataset and all the results and codes are available at https://bit.ly/2JrY5zv . The preliminary results showed that we can uncover interesting patterns of behaviour, which might help tutors to uncover conditions leading to student withdrawal. Results also indicated a pattern for passive withdrawal students. Since this is still work in progress we plan, for example, to include Monte Carlo simulation using computed Markov chains to simulate the behaviour of a single student.

7 Acknowledgement

This work was supported by junior research project by Czech Science Foundation GACR no. GJ18-04150Y.

References

[1] Moodle HQ, "Moodle statistics," Moodle HQ, 2018. [Online]. Available: https://moodle.org/stats/. [Accessed 25 04 2018].

[2] Coursera Inc., "Coursera," Coursera Inc., 2012. [Online]. Available: https://www.coursera.org/. [Accessed 10 April 2018].

[3] Z. Papamitsiou and A. A. Economides, "Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence," *Educational Technology & Society,* vol. 17, pp. 49-64, 2014.

[4] M. Hlosta, D. Herrmannova, L. Vachova, J. Kuzilek, Z. Zdrahal and A. Wolff, "Modelling student online behaviour in a virtual learning environment," in *Proceedings of the 4th International Conference on Learning Analytics and Knowledge*, Indianapolis, 2014.

[5] F. Okubo, A. Shimada, Y. Taniguchi and S. Konomi, "A Visualization System For Predicting Learning Activities Using State Transition Graphs," in *Proceedings of 14th International Conference on Cognition and Exploratory Learning in Digital Age*, Vilamoura, 2017.

[6] D. Davis, G. Chen, C. Hauff and G.-J. Houben, "Gauging MOOC Learners' Adherence to the Designed Learning Path," in *Proceedings of 9th International Conference on Educational Data Mining*, Raleigh, 2016.

[7] J. R. Norris, Markov Chains, Cambridge: Cambridge University Press, 1997.

[8] J. Kuzilek, M. Hlosta and Z. Zdrahal, "Open University Learning Analytics dataset," *Scientific Data,* vol. 4, 2017.

[9] A. Wolff, Z. Zdrahal, D. Herrmannova, J. Kuzilek and M. Hlosta, "Developing predictive models for early detection of at-risk students on distance learning modules," in *Proceedings of the 4th International Conference on Learning Analytics and Knowledge*, Indianapolis, 2014.