Generating Actionable Predictive Models of Academic Performance

Abelardo Pardo
Faculty of Engineering and IT
The University of Sydney, Australia
Abelardo.Pardo@sydney.edu.au

Negin Mirriahi
School of Education
and Learning & Teaching Unit
University of New South Wales, Australia
Negin.Mirriahi@unsw.edu.au

Roberto Martinez-Maldonado
Connected Intelligence Centre
University of Technology Sydney, Australia
Roberto.Martinez-Maldonado@uts.edu.au

Jelena Jovanovic
Department of Software Engineering
University of Belgrade, Serbia
jeljov@fon.rs

Shane Dawson
Teaching Innovation Unit
University of South Australia, Australia
Shane.Dawson@unisa.edu.au

Dragan Gašević
School of Education and Informatics
The University of Edinburgh, UK
dgasevic@acm.org

ABSTRACT
The pervasive collection of data has opened the possibility for educational institutions to use analytics methods to improve the quality of the student experience. However, the adoption of these methods faces multiple challenges particularly at the course level where instructors and students would derive the most benefit from the use of analytics and predictive models. The challenge lies in the knowledge gap between how the data is captured, processed and used to derive models of student behavior, and the subsequent interpretation and the decision to deploy pedagogical actions and interventions by instructors. Simply put, the provision of learning analytics alone has not necessarily led to changing teaching practices. In order to support pedagogical change and aid interpretation, this paper proposes a model that can enable instructors to readily identify subpopulations of students to provide specific support actions. The approach was applied to a first year course with a large number of students. The resulting model classifies students according to their predicted exam scores, based on indicators directly derived from the learning design.

CCS Concepts
• **Applied computing → Education → Interactive learning environments.**

Keywords
Learning analytics; personalization; feedback; recursive partitioning

1. INTRODUCTION
To date, research contributions in the field of learning analytics have broadly aimed to further our understanding of the learning process [21]. Central to this work has been the development of predictive models of student learning behavior and performance [1, 2, 6, 15, 19]. The models commonly aid in identifying potential relationships among various behavioral, demographic, and performance-based factors in large data sets. A well-adopted example lies in the research associated with the early identification of students at-risk of academic performance or attrition [10]. While these types of predictive models have been well utilized for developing early indicators, there has been minimal attention investigating how such information can be best deployed to promote reflection and action among teaching staff and students. The data used to create such models requires substantial pre-processing to comply with the requirements imposed by the algorithms. Furthermore, once the predictions are obtained, they usually require non-trivial interpretations. These requirements increase the barrier for adoption by instructors and students.

In order for learning analytics to have widespread impact and uptake as a discipline, predictive models need to offer intuitive actionable insight for both instructors and students, to overcome some of the existing concerns reported in the literature. For instance, there are claims that the uptake of learning analytics has resulted in little improvements in the quality of student feedback, and only increased the frequency of such feedback [22]. Furthermore, the design and deployment of interventions derived from collected data has been relative unexplored [24]. This paper proposes how a recursive partitioning technique can offer instructors a predictive model to help them derive data-informed pedagogical interventions and personalized feedback to different student subpopulations.

The rest of the paper is organized as follows. Section 2 reviews the work in the area of predictive algorithms and their potential to drive interventions. Section 3 describes the method followed for the recursive partitioning statistical analysis. The resulting models are described in Section 4 together with some discussion about the interpretation and performance of the models. Section 0 contains the conclusions of the study and some ideas for future work.

2. RELATED WORK
There has been much research in the fields of learning analytics and educational data mining dedicated to predictive modeling. This has most commonly involved the development of models associated with the prediction of students at risk of failing a course and the prediction of students' grades [8]. These predictions are largely based on the use of data stored in institutional and learning management systems. For example, Agudo et al. [1] used classification algorithms to examine the effect of three categories of interactions with academic performance. Alstete and Beutell [2] examined student

LAK '16, April 25-29, 2016, Edinburgh, United Kingdom

DOI: http://dx.doi.org/10.1145/2883851.2883870

performance in online education in management with respect to several qualitative indicators, but did not elaborate on the actions to be derived from these relations. Barber and Sharkey [6] described one of the increasingly present models to predict when students are at risk of abandoning the institution using a Naïve Bayes algorithm. The study aimed to detecting students at risk of academic failure in order to provide timely and early support interventions. A similar application was presented by Jayaprakash et al. [15] in the context of an open source initiative across multiple institutions including a portfolio of interventions that may help address academic concerns regarding student retention. In this case, the techniques were applied at an institutional level in an attempt to reduce overall student attrition. To date, the work on predictive modeling has demonstrated promising results with respect to the level of prediction accuracy. Moreover, researchers have also shown that the application of predictive models can lead to significant educational gains and increased student retention over an extended period of time [3].

Educational data mining [4] has specialized in the study of algorithms to discover new relations or knowledge that are not evident from large data sets. Although data mining algorithms are being applied to numerous areas, the applications have mostly focused on detecting adequate relationships rather than on the implications stemming from the use of such algorithms or how to make the resulting models available to instructors to derive learning support actions (see for example [5, 16]). A comprehensive description of how to use data mining techniques was presented by Romero et al [19, 20]. However, the interventions are only discussed as potential actions to be decided by the instructor.

Timely and effective feedback has been identified as a key factor to influence student learning [13, 14]. It has also been considered as an inherent part of self-regulated learning as it affects cognitive engagement with tasks [7]. Formative assessment provides opportunities to give students feedback as they progress in their learning experience, helping them identify areas needing attention or improvement [12]. The increasing use of student-centered and active-learning teaching methodologies to improve learning outcomes [9, 23] provide additional opportunities to offer highly personalized feedback. Data-supported predictive models, such as the one presented in this paper, can offer instructors with valuable insights on the type of feedback and how to deliver it.

While the work on predictive modeling shows promise for advancing teaching practice, there is much less research available on how to actually make use of such predictive models to improve feedback strategies. The study conducted by Tanes et al. [22] explored feedback messages sent to students as a response to early warning alerts triggered by a predictive model of the *Signals* software tool. The study showed that while the early warning system increased the frequency of summative feedback (i.e., informing students how they stand with respect to meeting certain criteria and standards), it did not increase the quality of feedback (i.e., sending formative and instructional feedback to help students identify how they could improve their learning). A likely reason for the low influence on instructional practice can be attributed to the lack of actionable insight these models provide [11]. Although the traffic light metaphor used in *Signals* is intuitive, there is insufficient transparency related to the reasons why and how certain predictions are made. This transparency in risk calculations is essential in order for instructors and students to understand how best to act upon the predictions.

Further, predictive models are often created without taking into account the nuanced factors found in different learning situations. With the aim to maximize scalability and reusability, predictive models are frequently constructed to generalize across different courses and subject domains. However, research shows that such generalized models cannot detect factors that can inform teaching and learning if the instructional conditions of individual courses are not considered [10]. Likewise, assumptions are frequently made about learners on average rather than taking into account student agency and individual differences (e.g., metacognitive skills and prior knowledge) in order to tailor personalized feedback for individual students and student sub-populations [17].

In order to explore the need for new approaches in presenting the results of predictive models, this paper reports the findings of a study that examined the use of a recursive partitioning technique to produce a predictive model. The model is suitable to inform instructors about the estimated future academic performance of the students based on engagement with course material. The algorithms based on recursive partitioning provide a good match between their requirements and the factors typically available in a learning environment [18]. The algorithms of this type do not require the specification of any model in advance, and can handle an arbitrarily large number of numeric features with different scales. This is particularly useful in multimodal environments in which multiple data sources are obtained from a variety of tasks. Conventional algorithms require data to be categorized, normalized, or transformed into the required format. These algorithms are capable to handle variables with no restrictions. The algorithms also perform automatic selection of the most relevant features from a potentially large collection. This is also an advantage when a rich set of features is obtained from a learning experience but only some of them are relevant for predictive purposes. Additionally, the resulting model, a tree, provides instructors with a visualization that can help them translate the findings into appropriate learning support actions. The following sections outline a case study describing the adopted technique used to predict student academic performance and identify subpopulations of students for personalized feedback.

3. METHOD

3.1 Study Design

The case study described in this section collected data from a 13 week first year course at a large research intensive university in Australia. The course design involved students engaging in weekly activities comprised of both formative and summative assessment items over eleven core weeks of the course. All course components were offered in electronic form and made available following a blended learning strategy mediated by the corporate Learning Management System. The data for the study was derived from the interactions of the students with the course components. The study was carried out in the context of a natural experiment as the participation of the students was beyond the control of the researchers.

3.2 Materials

The study used several data sources derived from the interaction of the students with the course components, namely four types of activities. The first type (VID) consisted of an interactive HTML page with a video clip introducing new course concepts. All *play* and *pause* events in the video clips were recorded. The second type of activities (VEQ) were included immediately next to the video clip and consisted of a formative assessment in the form of multiple-choice questions related to the concepts covered in the adjoining video clip. Students could answer each question, review whether the answer was correct, or request to see all the answers. This last option was offered only if an answer was provided. The events *correct, incorrect* and *show* were recorded denoting the three possible actions. The third type of activity (EQT) required students to read text in an HTML document with additional formative assessment in the form of multiple-choice questions (identical to the ones previously described). The fourth set of activities (EXC) required students to solve a sequence of exercises and provide their answer through a multiple-choice question. The platform would select an exercise randomly from a pre-defined sequence. If the exercise was answered correctly, it no longer appeared in the sequence. If the answer was incorrect, the exercise remained as part of the sequence. The students advanced through the sequence (and therefore repeated exercises that were answered incorrectly) and a score was calculated as the percentage of exercises answered correctly. This exercise was a summative assessment and the score was directly added to the course marks. The server registered the events *correct* and *incorrect* for each exercise attempted by each student. Each event generated in any of the assessment items has a unique identifier that allows its differentiation from the rest of items. The three types of assessment (VEQ, EQT, and EXC) were treated as separated categories in the analysis. These activities were made available to the students gradually every week and remained available for the remainder of the semester. Two additional data sources were considered in the study: the results of the midterm examination (scheduled in Week 6 of the semester) and final examinations. These exams contributed 20% and 40% respectively towards the final course mark.

3.3 Procedure

The data was collected in the 2014 offering of a large first year engineering course (n = 272). The weekly schedule included one 2-hour lecture, one 2-hour tutorial and a 3-hour laboratory session. The data was extracted from the server logs and fully anonymized. The course design included the same pattern of activities for the core weeks 2-5 and 7-13. The activity sequence was comprised of an initial set of formative assessment activities of types VID, VEQ and EQT followed by a summative assessment activity of type EXC with submission deadlines before the start of the lecture to encourage in-class preparation. Additionally, another set of EQT activities and an additional EXC activity were scheduled with deadlines before the start of the tutorial sessions. Students were provided with real-time indicators of their participation in these activities showing the percentage of activities that were attempted (if formative) and their score (if summative).

3.4 Variables and Measures

The variables used as predictors to build the models were derived from the interaction students had with the various learning design resources available in an online platform. Since the course has tasks and assessments due every week, the variables used for the analysis are the weekly counts for the activity events across eleven core course weeks. More precisely, for each week and each student, the following counts were extracted from the server logs:

- VID.PL/VID.PA: Number of *play/pause* events in the videos for the week.
- VEQ.CO/VEQ.IN/VEQ.SH: Number of times a question next to a video clip was answered correctly, incorrectly or the answer was shown.
- EQT.CO/EQT.IN/EQT.SH: Number of times a question in the course notes was answered correctly, incorrectly or the answer was shown.
- EXC.CO/EXC.IN: Number of exercises answered correctly/incorrectly.

For each week of the course, a set of variables was extracted and used to calculate the predictive models. For example, the model for week 3 was calculated with the set of variables reflecting the number of events recorded during that week (ignoring the rest). These models (eleven of them) were divided into two categories: those predicting student performance on the midterm exam (for weeks 2 to 5) and those predicting performance on the final exam (weeks 7 to 13). The midterm and final exam scores are numerical variables with values in the range [0-20] and [0-40] respectively.

3.5 Data Analysis

The study aimed to show the feasibility of obtaining a predictive model of the performance of the students in the midterm and final examination using the data about their interaction with the online activities. Essentially, the study explored the suitability of predictive techniques that can handle numeric indicators and provide a simple interpretation of the results. The performance of the model was measured using the Mean Absolute Error (MAE) defined as the mean of the differences between the predicted and real scores, and the Root of the Mean Squared Error (RMSE) defined as the square root of the mean of the square differences between the predicted and real scores (MSE). The model validation was performed following a *leave one out* cross-validation strategy. For each of the samples a model was calculated leaving one sample out and its error in the estimation (on the left-out sample) was obtained. The errors were then combined to calculate MAE and RMSE.

4. RESULTS

The descriptive statistics (mean and standard deviation) for each of the variables during the eleven core weeks of the course are shown in Table 1. The mean and standard deviation of the midterm and final scores were 13.3 (4.1) and 19.1 (8.8), respectively. Predictive models were calculated for all eleven weeks of the semester using recursive partitioning. For each model, a collection of rules and a tree structure was produced. The rules were stated as conditions on the input variables (event counts). Each node of the tree was labeled with a condition and two outgoing edges. As an example, the resulting predictive tree for Week 10 is shown in Figure 1.

4.1 Instructional Interpretation

As illustrated in Figure 1, the predictive tree for Week 10, obtained through recursive partitioning, contains 13 nodes and 6 rules. The recursive structure of the tree represents how the entire population of students is recursively partitioned into subsets. The final partition is represented by the leaf nodes at the bottom of the figure. The top node represents the entire cohort and shows the condition used to divide it into two subsets represented by the sub-trees. The left branch includes 108 students who provided more than 21 incorrect answers to the exercise sequence (condition EXC >=22 and node number 2 in the tree). The right branch of the top node contains 167 students who did not satisfy the condition (i.e. less than 22 errors in the exercise sequence) and with an average predicted score of 9.6 over 40. The partitioning is then performed recursively at the level of the sub-nodes. At each point in the tree, a different feature is chosen to divide the group.

The final partitioning of the cohort is represented by the leaves at the bottom of the tree. This model has divided the group into

Table 1: Descriptive statistics of the event count variables of the study

Week	VID.PL	VID.PA	VEQ.CO	VEQ.IN	VEQ.SH	EQT.CO	EQT.IN	EQT.SH	EXC.IN	EXC.CO
W2	8.9(16.1)	8.9(18.5)	6.4(6.7)	3.7(4.5)	2.1(2.9)	6.8(12.6)	4.6(9.6)	2.3(6.7)	16.7(11.0)	15.2(14.6)
W3	9.4(16.8)	9.5(20.9)	9.1(10.3)	6.1(7.9)	2.5(5.0)	11.9(14.5)	7.8(10.4)	3.6(7.1)	25.5(13.5)	38.6(25.3)
W4	13.2(26.1)	12.7(24.7)	8.6(11.5)	5.7(9.3)	2.0(4.1)	6.8(12.8)	4.4(8.6)	1.8(5.3)	25.6(12.4)	37.9(28.7)
W5	6.1(14.9)	5.9(15.9)	4.2(6.9)	2.7(4.7)	1.0(2.9)	7.3(13.3)	5.4(11.1)	2.1(5.2)	22.9(11.0)	33.5(25.4)
W7	7.3(21.2)	7.2(24.0)	3.8(6.3)	3.2(5.5)	1.2(3.2)	1.6(4.3)	1.0(2.9)	0.4(1.8)	19.9(7.4)	26.4(21.1)
W8	8.1(21.0)	9.1(27.8)	3.0(5.7)	2.5(4.9)	0.7(2.3)	2.3(4.5)	1.5(3.2)	0.4(1.3)	18.2(8.5)	26.8(21.6)
W9	6.9(15.7)	7.2(19.1)	3.0(5.8)	2.7(5.4)	0.9(2.6)	2.4(5.6)	1.6(3.7)	0.7(2.5)	19.0(8.3)	33.6(26.2)
W10	6.0(16.4)	6.8(21.0)	2.5(5.5)	2.6(5.3)	0.8(2.4)	1.8(5.3)	1.3(3.3)	0.3(1.4)	11.7(18.7)	18.2(21.1)
W11	5.5(11.9)	5.5(15.1)	1.9(3.5)	1.9(3.9)	0.6(1.7)	2.1(7.0)	1.8(5.4)	0.7(3.3)	15.4(8.4)	23.5(23.8)
W12	4.2(13.0)	4.3(15.1)	2.4(5.2)	2.3(5.1)	0.9(2.9)	1.0(3.6)	0.7(2.7)	0.4(2.1)	15.5(7.3)	23.0(19.2)
W13	44.4(95.8)	38.6(94.8)	14.2(26.8)	11.4(20.7)	3.8(8.4)	30.1(39.5)	17.0(20.5)	6.7(10.9)	71.1(96.6)	52.0(63.7)

seven subpopulations with a different number of students per subgroup and with predicted scores on the final examination ranging from 6 (left most node) to 15 (right most node). An equivalent interpretation of the model is that every path from the root node to a leaf provides the set of conditions satisfied by that subset of students. The conjunction of these conditions would create the rule used to identify the students. For example, the 90 students represented by node number 4 are identified by the property that EXC.IN >=22 AND VID.PL < 8.5, or in other words, the number of incorrect answer to one of the test is larger than 21, and the number of play events in the videos less than 8.5.

The main advantage of this model is that the inputs for the algorithm (features) are indicators derived directly from the learning design. Additionally, the result divides the cohort into a manageable number of subpopulations. An instructor may now use the predicted score to provide personalized feedback to each subset of students. For example, the suggestions given to the 20 students with a predicted score of 15 (out 40) would be different from those given to the 90 students with a predicted score of six.

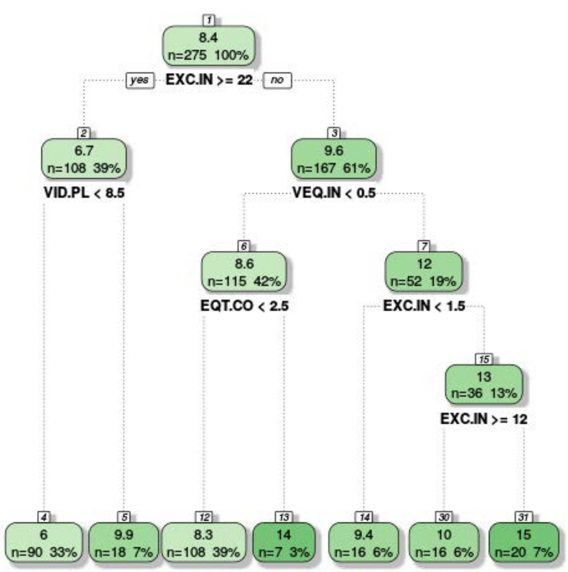

Figure 1: Decision tree for Week 10.

A more in-depth analysis of the resulting decision tree may also guide the instructor towards a higher level of personalization. For example, students in the subpopulation with the lowest predicted score have a high number of incorrect exercise submissions (node 1) and a low number of *play* video events (node 2). An instructor may refer to this information when customizing the feedback inviting them to increase their engagement with the videos, which may help reduce the number of their incorrect exercise submissions.

4.2 Model Performance

The performance of the resulting model is assessed based on the accuracy of the score predictions in the leave nodes. This information can be conveyed to the instructors so that they have an estimate of the error in the predictions. The performance analysis is shown in Table 2.

Table 2: Performance of the model for the midterm

	MSE	RMSE	MAE
W2	15.826	3.978	3.14
W3	15.476	3.934	3.053
W4	15.14	3.891	3.013
W5	14.469	3.804	3.007

The MAE for all four models calculated is stable at approximately 15% of the midterm score (3 out of 20). The RMSE, on the other hand, offers a higher value (approximately 4) but still below 20% (4 out of 20). These figures have a simple interpretation with respect to the model. The score predicted for a subgroup may have an average of 15 or 20% error. Although this type of error would not be accepted in other disciplines, when applied to the provision of feedback in a learning context it helps to reduce the uncertainty when trying to identify the appropriate sets of students.

Table 3: Performance when estimating the final exam score

	MSE	RMSE	MAE
W7	29.182	5.402	4.4
W8	32.358	5.688	4.726
W9	28.905	5.376	4.481
W10	32.101	5.666	4.665
W11	30.436	5.517	4.482
W12	25.476	5.047	4.097
W13	32.037	5.66	4.503

The performance of the models improves when predicting the score for the final exam. Table 3 shows the performance for the remaining seven weeks of the semester. In this case, since the score is in the range [0-40] the MAE is below 11.8% (4.726 over 40) whereas the RMSE is below 14.22% (5.688 over 40).

5. CONCLUSION

The predictive models derived from data captured from learning environments usually require complex data-preparation processes and produce results that are difficult to interpret by instructors and students. This paper presented a case study that used recursive partitioning to process a large number of numeric features derived from the student engagement with the learning environment and automatically select those that provide a robust classification with respect to their predicted academic performance. The estimated error in the prediction is within reasonable bounds. Although the results should not necessarily be used in the exact form as shown in this paper, they provide a transparent characterization of a student cohort based on indicators extracted from the learning environment thus facilitating its translation into actions. These models may serve as the basis to explore more efficient sense-making solutions to support instructors in the provision of frequent and effective formative and personalized feedback.

6. REFERENCES

[1] Agudo-Peregrina, Á.F., Iglesias-Pradas, S., Conde-González, M.Á. and Hernández-García, Á. 2014. Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior*. 31, (Feb. 2014), 542–550 doi:10.1016/j.chb.2013.05.031.

[2] Alstete, J.W. and Beutell, N.J. 2004. Performance indicators in online distance learning courses: a study of management education. *Quality Assurance in Education*. 12, 1 (Mar. 2004), 6–14 doi:10.1108/09684880410517397.

[3] Arnold, K.E. and 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* (New York, NY, USA, 2012), 267–270 doi:10.1145/2330601.2330666.

[4] Baker, R.S.J. 2010. Data Mining. *International Encyclopedia of Education*. Elsevier. 112–118.

[5] Baker, R.S.J. d., D'Mello, S.K., Rodrigo, M.M.T. and Graesser, A.C. 2010. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*. 68, 4 (Apr. 2010), 223–241 doi:10.1016/j.ijhcs.2009.12.003.

[6] Barber, R. and Sharkey, M. 2012. Course Correction: Using Analytics to Predict Course Success. *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (New York, NY, USA, 2012), 259–262 doi:10.1145/2330601.2330664.

[7] Butler, D.L. and Winne, P.H. 2014. Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*. 65, 3 (2014), 245–281 doi:10.3102/00346543065003245.

[8] Dawson, S., Gaševi, D., Siemens, G. and Joksimovic, S. 2014. Current State and Future Trends : A Citation Network Analysis of the Learning Analytics Field. *International Conference on Learning Analytics and Knowledge* (2014), 232–240 doi:10.1145/2567574.2567585.

[9] Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H. and Wenderoth, M.P. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*. (May 2014), 1–6 doi:10.1073/pnas.1319030111.

[10] Gašević, D., Dawson, S., Rogers, T. and Gasevic, D. 2016. Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting learning success. *The Internet and Higher Education*. 28, (2016), 68–84.

[11] Gašević, D., Dawson, S. and Siemens, G. 2015. Let's not forget: Learning analytics are about learning. *TechTrends*. 59, 1 (2015), 64–71.

[12] Gibbs, G. and Simpson, C. 2004. Conditions under which assessment support students' learning. *Learning and Teaching in Higher Education*. 1, 1 (2004), 3–31.

[13] Hattie, J. and Jaeger, R. 1998. Assessment and Classroom Learning: a deductive approach. *Assessment in Education: Principles, Policy & Practice*. 5, 1 (Mar. 1998), 111–122 doi:10.1080/0969595980050107.

[14] Hattie, J. and Timperley, H. 2007. The Power of Feedback. *Review of Educational Research*. 77, 1 (Mar. 2007), 81–112 doi:10.3102/003465430298487.

[15] Jayaprakash, S.M., Moody, E.W., Lauría, E.J.M., Regan, J.R. and Baron, J.D. 2014. Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics*. 1, 1 (May 2014), 6–47.

[16] Jiang, Y., Paquette, L., Baker, R.S., Street, W., Clarke-midura, J. and Hill, O.M. 2015. Comparing Novice and Experienced Students within Virtual Performance Assessments. *International Conference on Educational Data Mining* (2015).

[17] Kovanović, V., Gašević, D., Joksimović, S., Hatala, M. and Adesope, O. 2015. Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*. 27, (Oct. 2015), 74–89 doi:10.1016/j.iheduc.2015.06.002.

[18] Lantz, B. 2013. *Machine Learning with R*. Packt Publishers.

[19] Romero, C., López, M.-I., Luna, J.-M. and Ventura, S. 2013. Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*. 68, (Oct. 2013), 458–472 doi:10.1016/j.compedu.2013.06.009.

[20] Romero, C., Ventura, S. and García, E. 2008. Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*. 51, 1 (Aug. 2008), 368–384 doi:10.1016/j.compedu.2007.05.016.

[21] Siemens, G. 2013. Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*. 57, 10 (Aug. 2013), 1380–1400 doi:10.1177/0002764213498851.

[22] Tanes, Z., Arnold, K.E., King, A.S. and Remnet, M.A. 2011. Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education*. 57, 4 (Dec. 2011), 2414–2422 doi:10.1016/j.compedu.2011.05.016.

[23] Wieman, C.E. 2014. Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences of the United States of America*. 111, 23 (Jun. 2014), 8319–20 doi:10.1073/pnas.1407304111.

[24] Wise, A.F. 2014. Designing Pedagogical Interventions to Support Student Use of Learning Analytics. *Proceedings of the International Conference on Learning Analytics and Knowledge* (2014).