Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback

Jalal Nouri
Department of Computer and Systems Sciences, Stockholm University
Stockholm, Sweden

Mohammed Saqr
School of Computing, University of Eastern Finland
Joensuu, Finland

Uno Fors
Department of Computer and Systems Sciences, Stockholm University
Stockholm, Sweden

ABSTRACT

Learning analytics (LA) is a relatively new research discipline that uses data to try to improve learning, optimizing the learning process and develop the environment in which learning occurs. One of the objectives of LA is to monitor students' activities and early predict performance to improve retention, offer personalized feedback and facilitate the provision of support to the students. Flipped classroom is one of the pedagogical methods that find strength in the combination of physical and digital environments – i.e. blended learning environments. Flipped classroom often make use of learning management systems in which video-recorded lectures and digital material is made available, which thus generates data about students' interactions with these materials. In this paper, we report on a study conducted with focus on a flipped learning course in research methodology. Based on data regarding how students interact with course material (video recorded lectures and reading material), how they interact with teachers and other peers in discussion forums, and how they perform on a digital assessment (digital quiz), we apply machine learning methods (i.e. Neural Networks, Naïve Bayes, Random Forest, kNN, and Logistic regression) in order to predict students' overall performance on the course. The final predictive model that we present in this paper could with fairly high accuracy predict low- and high achievers in the course based on activity and early assessment data. Using this approach, we are given opportunities to develop learning management systems that provide automatic data-driven formative feedback that can help students to self-regulate as well as inform teachers where and how to intervene and scaffold students.

Keywords: Learning analytics, feedback, assessment, machine learning, flipped classroom.

1. INTRODUCTION

Learning analytics (LA) is a relatively new research discipline that uses data to try to analyze learner behaviors and interactions, and the learning environments for the sake of improving learning, optimizing the learning process and the environment in which learning occurs (Siemens, 2013). LA often tries to make sense of the available massive data recorded by learning management systems, such as data resulting from clicking behavior, logins to learning management systems and access to educational resources etc. Investigators have used these data to create predictive models that uses computer algorithms to early detect students who are at risk of failing or underachieving (Leitner, Khalil, & Ebner, 2017; Saqr, Nouri, & Fors, 2018; Nouri et al., 2019).

Monitoring of students' activities and early prediction of performance to improve retention, offer personalized feedback and facilitate the provision of support to the students are the foremost objectives of learning analytics. Those insights hold the promise of the provision of practical solutions to high priority issues in education such as attrition, quality of learning experience, improve learning design, informing decision makers (Leitner et al., 2017; Saqr et al., 2017; Wong, 2017). Learning analytics may also reduce the gaps in personalized feedback in the resource-restricted educational landscape. Real life results of LA initiatives have proven to be fruit-bearing in a number of instances, such as Course Signals at the University of Purdue, E2Coach in the University of Michigan University and the usage of "Automated Wellness Engine" in the University of New England. Other trials in controlled studies where learning analytic informed intervention was compared to traditional methods provide more evidence on the worth of learning analytic; an example is the work in Marist College, USA, where the intervention led to significant improvements (Wong, 2017).

While early learning analytics results have showed promising results, they are still challenging to reproduce, generalize, or link to theory (Ga et al., 2016). These challenges call for a revision of the approach, an approach that increases the depth and width of data collection to data that are more representative of students' effort in learning (Chen, 2015; Dawson, Mirriahi, & Gasevic, 2015; Ga et al., 2016). Some data sources have received little attention in the field of learning analytics, most notably assessment data and video data. Assessment is a significant source of feedback for students and educators alike. Moreover, assessment has always been known to drive students' learning and define the efforts of students in learning (Cubric & Tosic, 2011). In contrast to traditional methods of assessment, e-assessments generates large amount of information about students and curricula (Cubric & Tosic, 2011; Timmis, Broadfoot, Sutherland, & Oldfield, 2016; Saqr, 2017). Assessment data is still under-explored and largely under-developed, probably because researchers treat assessment as an outcome they work to optimize, rather than an integral part of the analytics data cycle. This might be even more interesting when using formative assessment data. However, harnessing the power of e-assessment in analytics may probably help close the learning analytics data loop and add to the usefulness of learning analytics (Clow, 2012). Video data is another source of data that carries the potential of providing better insights and improvement the predictive models used nowadays. Video data are becoming increasingly common in blended learning such as flipped classroom scenarios (Nouri, 2016). In flipped classrooms, students rely on the video lectures as a major source of information. Previous research in the field of video analytics reported that frequent video views correlates with better levels of cognition and thinking as well as performance progress (Giannakos, Chorianopoulos, & Chrisochoides, 2015).

In this paper, we report on a study performed with focus on a flipped learning course in research methodology. Based on data regarding how students interact with course material (video lectures and reading material), how they interact with teachers and other peers in discussion forums, and how they perform on a digital assessment (digital quiz), we apply machine learning methods in order to predict students' performance on the course.

2. METHOD

Study context

Data was collected in a course on research methodology at Stockholm University that was given in autumn 2017. The course used flipped classroom methodology as basis meaning in this case that a large portion of the lectures were made available in digital format in an online environment. The course focused on in this study prepares students for the bachelor thesis with respect to scientific methodology and communication. The learning objectives are on the one hand to facilitate students understanding of the fundamentals of re- search strategies, data-collection methods, and analysis methods, and on the other hand to familiarize students with application of qualitative and quantitative methods of analysis. Put differently, the course aimed at equipping students with conceptual knowledge (an understanding of scientific methods), and procedural knowledge (application of analysis methods and scientific writing). The course was divided into three parts with three different examination tasks that was the basis of the final grade on the course. The first part concerned gaining a theoretical understanding of the fundamentals of research strategies, data-collection methods, and analysis methods. The pedagogical structure for this part comprised of independent reading of course literature. Students reading of the course literature was supported by three longer video lectures (in average 60 min each), one traditional campus lecture (teacher presenting and summarizing the fundamentals of research strategies), and one interactive flipped classroom lecture.

The second part was a practical qualitative analysis project that students conducted in groups of two. The task of this project was to use a qualitative analysis method to analyse qualitative interview data and communicate the results in a report following scientific standards of qualitative data presentation. During this project, the students were supported by five digital lectures (in average 35 min each), three flipped lectures on campus, and digital supervision through the learning management system. In the three flipped lectures on campus students worked with their projects and were scaffolded by several teachers that answered questions and provided feedback. When the teachers identified common misunderstandings, or needs among the students, they provided elaborated explanations to the whole class. The examination of the second part comprised of a written group report.

The third part of the course was similar to the second part, comprising of a project with a focus on using quantitative methods to analyse a questionnaire and communicate the results according to scientific standards of quantitative result presentation. During this project, the students were supported by seven video lectures (in average 30 min each), three flipped lectures in class with teachers scaffolding practical work, and digital supervision in the learning management system. The videos covered the theoretical fundamentals of descriptive and inferential statistics as well as how different statistical tests can be performed and interpreted in SPSS. The examination of the third part comprised of a written group report. All video lectures made available to the students during the course were produced by teachers and researchers in a professional video studio at Stockholm University. The video lectures were specifically tailored for the course.

Data collection and data processing

We informed the participating students in the beginning of the course about that we were collecting digital data about their learning processes and interactions with learning material in the online environment used during the course

(Moodle learning management system) and asked for consent. The data collected and analyzed is based only on the students (n=251) that gave consent for this research, making sure ethical requirements are met.

Using a Moodle plugin, we performed database queries collecting the following data (features):

- Results on digital quiz (total score and score on individual quiz questions)
- Interactions with available video lectures and reading material (captured in clicks)
- Interactions with teachers in digital forums (no. of posts)
- Interactions with other students in digital forums (no. of posts)

In total, we collected data for 55 features that were used for the final predictive models. We also collected data regarding the students' final grades on the course. This variable, which was used as a target variable, was transformed into a categorical variable representing low- and high achievers (grade A-B for high achievers, and C-F for low achievers). Data were inspected and explored for corrupted, missing, or incomplete records. Collected data were combined with performance data in a single table. IDs and identifying information were removed to completely anonymize the data.

Data analysis

We performed descriptive statistical analysis using SPSS. Spearman correlation test was performed to investigate the correlation between performance on digital quiz and course performance. We then performed predictive analytics using several machine learning models (Naïve Bayes, Logistic regression, kNN, Neural Network, and Random Forest) in order to investigate if student's interactions (and absence of interactions) with available online material as well as their performance on the digital quiz could predict how well they performed on course level in terms of low- and high performance. 55 features were used in these models with performance on course level measured in low- and high achievers as target variable. Features were ranked using information gain ratio. To prevent overfitting 10-fold cross-validation was performed.

3. RESULTS

Table 1 below presents descriptive statistics concerning students' performance on course. As can be seen in the table, the total number of participating students on the course were n=251, of which approximately 90% passed the course, 80% passed the digital examination in the first attempt. Among the students that passed the course, approximately 63% were low-achieving and 37% high-achieving on course level, and approximately the same proportion of low- and high achievers in terms of performance on the digital examination.

Table 1. Descriptive statistics

	n	%	M	SD
Students on course	251			
Passing students on digital quiz	197	79.49		
Failing students on digital quiz	54	21.51		
Passing students on course	225	89.64		
Failing students on course	26	10.35		
Average student grade on digital quiz	225		9.21	2.77
Average student grade on course	225		3.14	0.93
Low & High achievers on quiz	223			
Low	148	66.37	7.62	1.86
High	75	33.63	12.33	1.13
Low & High achievers on course	225			
Low	141	62.67	2.55	0.61
High	84	37.33	4.13	0.34

When performing a correlation analysis, we could not see any significant correlation between performance on the digital quiz and the final grade ($r=0.09$, $p>0.05$).

Predicting performance on course based on online activities and digital quiz

We performed predictive analytics using several machine learning models (Naïve Bayes, Logistic regression, kNN, Neural Network, and Random Forest) in order to investigate if student's interactions (and absence of interactions) with available online material as well as their performance on the digital quiz could predict how well they performed on course level in terms of low- and high performance. Number of clicks with all available digital learning resources was used as features, as well as score on individual quiz questions and total score of quiz. In table 2 we see how well the different models performed. The kNN model was the best performing model with highest accuracy (81%) and AUC levels.

Table 2. Prediction accuracy and ROC

Model	Accuracy	AUC
kNN	0.81	0.73
Logistic regression	0.77	0.68
Naïve Bayes	0.73	0.59
Random Forest	0.69	0.67
Neural Network	0.65	0.64

In table 3 below we see a confusion matrix for the kNN model. As can be seen, the kNN model proved to predict the low achievers with a fairly high precision, 78% of the actual low achievers were correctly predicted, and 70 % of the high achievers.

Table 3. Prediction of low and high achievers using kNN

	True High	True Low	Class precision
Predicted High	58	32	64%
Predicted Low	24	111	79%
Class recall	70%	78%	

These results tell us that performance on course level can be predicted by a fairly high accuracy based on students' online activity data and their performance on a digital quiz. Furthermore, we can also see what features (online activities/resources) that had most weight in term of gain ratio (see table 4). In this case, the digital quiz stood out among the features, followed by a number of central video lectures covering the basics of research methodology. Such information can be used to provide formative feedback (and automatic recommendations) to students and teachers and be the basis of early-warning systems. The information produced can also be used to inform the development of the assessment instrument (digital quiz) as well as course design.

Table 4. Weights of 7 features with highest information gain

Features	Weights
Grade on digital quiz	1.0
What is science? Video Lecture 1	0.27
What is science? Video Lecture 4	0.25
What is science? Video Lecture 2	0.25
What is science? Video Lecture 3	0.24
Quantitative methods. Video lecture 1	0.22
Qualitative methods. Video lecture 1	0.21

4. CONCLUSIONS

Flipped classroom is one of the pedagogical methods that find strength in the combination of physical and digital environments – i.e. blended learning environments. Flipped classroom often make use of learning management systems in which video lectures and digital material is made available, which thus generates data about students' interactions with these materials. In this paper, we took as departure point to study if and how data generated by students in a flipped classroom environment, data such as their clicks on video resources, interactions with peers and teachers, as well as assessment data, could be used to predict overall performance in a course. Using a total of 55 features in our predictive modelling, and especially when using the kNN machine learning algorithm, we were able to fairly accurately predict both low- and high achievers on the course. The analysis also resulted in the identification of features with the highest information gain.

Such an information, in combination with the predictive model, can be used as the basis of early-warning systems and as a basis for automated formative feedback mechanisms, i.e. creating awareness about students' performance and recommending resources/activities, that support students to self-regulate towards increased performance. Furthermore, the information provided by the predictive analytics can be part of teachers monitoring practices and be used by teachers for interventions and provision of scaffolding in a data-driven manner.

5. REFERENCES

Chen, B. (2015). From theory use to theory building in learning analytics: A commentary on "Learning Analytics to Support Teachers during Synchronous CSCL". *Journal of Learning Analytics*, *2*(2), 163–168. https://doi.org/10.18608/jla.2015.22.12

Clow, D. (2012). The learning analytics cycle. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, 134. https://doi.org/10.1145/2330601.2330636

Cubric, M., & Tosic, M. (2011). Towards automatic generation of e-assessment using semantic web technologies. *International Journal of E-Assessment*.

Dawson, S., Mirriahi, N., & Gasevic, D. (2015). Importance of Theory in Learning Analytics in Formal and Workplace Settings. *Journal of Learning Analytics*, *2*(2), 1–4. https://doi.org/10.18608/jla.2015.22.1

Ga, D., Dawson, S., Rogers, T., Gasevic, D., Ga??evi??, D., Dawson, S., … Gasevic, D. (2016). Internet and Higher Education Learning analytics should not promote one size fits all : The effects of instructional conditions in predicting academic success. *Internet and Higher Education*, *28*, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002

Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making Sense of Video Analytics : Lessons Learned from Clickstream Interactions , Attitudes , and Learning Outcome in a Video-Assisted Course, *16*(1), 260–283.

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning Analytics in Higher Education—A Literature Review. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications* (Vol. 94, pp. 1–23). Cham: Springer International Publishing.

Nouri, J. (2016). The flipped classroom: for active, effective and increased learning–especially for low achievers. *International Journal of Educational Technology in Higher Education*, *13*(1), 33.

Nouri, J., Ebner, M., Ifenthaler, D., Saqr, M., Malmberg, J., Khalil, M., Viberg, O., Bruun, J., Conde-Gonzalez, M., Papamitsiou, Z. & Berthelsen, U. (2019). Efforts in Europe for Data-Driven Improvement of Education – A review of learning analytics research in six countries. International Journal of Learning Analytics and Artificial Intelligence for Education. 1(1).

Saqr, M. (2017). Assessment analytics: The missing step. *International Journal of Health Sciences*.

Saqr, M., Nouri, J. Fors, U. (2018). Time to focus on the temporal dimension of learning. A learning analytics study of the temporal patterns of students' interactions and self-regulation. International Journal of Technology Enhanced Learning.

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

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, *42*(3), 454–476.

Wong, B. T. M. (2017). Learning analytics in higher education: an analysis of case studies. *Asian Association of Open Universities Journal*, *12*(1), 21–40.