Educational Technology Approach toward Learning Analytics: Relationship between Student Online Behavior and Learning Performance in Higher Education

Taeho Yu
Purdue University
100 N. University Street, Rm3297
West Lafayette, IN 47906
765-494-6864
yu134@purdue.edu

Il-Hyun Jo
Ewha Womans University
Institute for Teaching & Learning,
11-1, Daehyeon-dong, Seodaemun-gu,
Seoul, Korea, 120-750
+82-2-3277-6689
ijo@ewha.ac.kr

ABSTRACT
The aim of this study is to suggest more meaningful components for learning analytics in order to help learners improving their learning achievement continuously through an educational technology approach. Multiple linear regression analysis is conducted to determine which factors influence student's academic achievement. 84 undergraduate students in a women's university in South Korea participated in this study. The six-predictor model was able to account for 33.5% of the variance in final grade, $F_{(6, 77)} = 6.457$, $p < .001$, $R^2 = .335$. Total studying time in LMS, interaction with peers, regularity of learning interval in LMS, and number of downloads were determined to be significant factors for students' academic achievement in online learning environment. These four controllable variables not only predict learning outcomes significantly but also can be changed if learners put more effort to improve their academic performance. The results provide a rationale for the treatment for student time management effort.

Categories and Subject Descriptors
K.3.1 Computer uses in education

General Terms
Measurement, Design, Human Factors.

Keywords
Learning analytics, Educational Technology, Higher education, E-learning.

1. INTRODUCTION
Learning analytics based on educational big-data mining is an emerging trend in the field of higher education. Whenever students utilize the Internet, computers, or Learning Management System (LMS), many log files are recorded. We can understand the current status of students' learning and even predict their possible learning achievement in a course by analyzing those log data that they leave in the database. This systematic process, which includes data collection-analysis, achievement prediction, and intervention treatment is called learning analytics. For instance, Purdue's Course Signals Model includes four components, which are as follows: a) academic performance in course, b) interaction with Learning Management System (LMS), c) prior academic data such as high school GPA and SAT score, and d) student characteristics such as age or residency [2].

However, three categories except b) are all about exogenous variables which are powerful predictor variables, but are not controllable by educators during instructions. Prior academic data and student characteristics cannot be changed by students' efforts [1] although they are a significant factor for learning outcomes.

From the perspective of constructivism which is one of the fundamental theories for educational technology, learning is a self-developing process by creating or reorganizing a concept or cognitive structure with learner's experience or beliefs [3]. For this reason, it is important to provide precautionary interventions for individual learners to support their learning process. Therefore, this study determines which controllable components need to be included in learning analytics to help learners improving their learning achievement continuously through an educational technology approach.

LAK '14, Mar 24-28 2014, Indianapolis, IN, USA
ACM 978-1-4503-2664-3/14/03.
http://dx.doi.org/10.1145/2567574.2567594

2. RESEARCH QUESTIONS
Do the six suggested independent variables (IV1: Total login frequency in LMS, IV2: Total studying time in LMS, IV3: regularity of learning interval in LMS, IV4: Number of downloads, IV5: Interactions with peers, and IV6: Interactions with instructor) predict learners' academic achievement?

3. METHODS
3.1 Research Context
The participants in this study were 84 undergraduate students in a face-to-face course entitled 'Understanding of Science of Public Administration'. This course had the following features: a) a three credit core course for undergraduate students in the department of Science of Public Administration, b) the instructor taught this course during spring semester 2013 for 16 weeks, c) 20% of the final grade was assigned for online discussion participation in LMS, and d) students use LMS to download course materials,

including syllabus or assigned readings. All participants are female students since this is a women's university.

3.2 Data Collection

Web-log data were collected from the Moodle-based Learning Management System (LMS), and the independent variables for this study, as shown in Table 1, were computed by automatic data collection module embedded in the LMS. Final grades were collected as a dependent variable in this study. Data matching process between independent variables and final grades was executed automatically in the database system.

Table 1. Data collecting methods for independent variables

Number	Independent variables	Data collecting methods
IV1	Total login frequency in LMS	Adding up the number of individual student's login time into the LMS
IV2	Total studying time in LMS	Calculating the total amount of time spent between login and logout
IV3	Regularity of learning interval in LMS	Calculating the standard deviation of average login time into the LMS
IV4	Number of downloads	Adding up the numbers of course materials downloaded
IV5	Interactions with peers	Counting the total number of student's postings responding to peers
IV6	Interactions with instructor	Counting the total number of student's postings responding to instructor

4. RESULTS

Multiple linear regression analysis was conducted to develop a model for predicting students' academic achievement based on their total studying time in LMS, total login frequency in LMS, (ir)regularity of learning interval in LMS, interactions with instructor, interactions with peers, and interactions with content. Results are shown in Table 2.

The six-predictor model was able to account for 33.5% of the variance in final grade, $F(6, 77) = 6.457$, $p < .05$, $R2 = .335$. Total studying time in LMS and interaction with peers had a significant ($p < .05$) correlation with final grade. In addition, (ir)regularity of learning interval in LMS and number of downloads had significant ($p <. 10$) partial effects in the full model. However, total login frequency in LMS (ß=-.131, t=-.980, $p > .05$) and interaction with instructor (ß=-.035, t=-.344, $p > .05$) did not predict final grade.

5. CONCLUSION

The purpose of this study is to suggest more meaningful components for learning analytics to help learners improving their learning achievement continuously in terms of educational technology approach. The regression model with only controllable variables that can be affected by learners' effort was able to account for 33.5% of the variance in students' academic achievement. The main focus of learning analytics tends to focus on the prediction of the future learning outcome by adding geographical, demographical, or characteristic factors such as high school GPA, SAT score, age, or residency.

Table 2. Results of multiple linear regression analysis

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	77.824	5.344		14.562	.000
IV1	-.011	.011	-.131	-.980	.330
IV2	.000	.000	.238	2.069	.042
IV3	-.227	.126	-.245	-1.801	.076
IV4	.099	.050	.208	1.981	.051
IV5	.174	.065	.278	2.689	.009
IV6	-.262	.761	-.035	-.344	.732
R^2 (adj. R^2) =.335(.283), F=6.457, p = .000					

a. Dependent Variable: Final grade

However, these factors are not controllable because they were fixed in the past and given to the instructional setting. Therefore, based on the results of this study, we tested six controllable variables for our learning analytics model and confirmed that four of them were significantly correlated with learning final grade. Moreover, these four variables not only predict learning outcomes significantly but also can be changed if learners put more effort to improve their academic performance. The advantage of learning analytics using big-data mining is to predict students' future performance. Yet educators should pay more attention to improving the process of learners' achievement rather than predicting achievement. However, this study was conducted with a face-to-face course in a women's university in South Korea. In addition, attendance rate, assignments and assessment composite, and discussion composite were not added as independent variables for this study because of the research context. Thus, more research should be implemented using various course subjects, different learning environments, and diverse participants with different school settings, ages, sex, nationalities, and level of student-instructor interactions.

6. REFERENCES

[1] Arnold, K. E. 2010. Signals: Applying Academic Analytics. *EDUCAUSE Quarterly, 33*(1).

[2] Arnold, K. E., and Pistilli, M. D. 2012. Course Signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge.* LAK '12. ACM, New York, NY, 267-270.

[3] Jo, I., and Kim, J. H. 2013. Investigation of Statistically Significant Period for Achievement Prediction Model in e-Learning. *The Journal of Educational Information and Media, 29*(2), 285-306.