

Analysis of Factors Influencing LMS Extracted Data using Learning Analysis on the Total Score of Learners

Jong-Teak Seo ¹, Young-gi Kim ², Ra-chel Ju ³

¹ Kookmin University, korea

² Kyunghee Cyber University , korea

³ Hanyang University, korea

¹ suho1793@hanmail.net, ² smileknock@gmail.com, ³ rachelju@naver.com

Abstract: The development of information and communication technology has led to change in today's modern society with various names, such as knowledge, information society, internationalization society, and especially, the development of distance education without time and space restrictions. With the development of distance education, problems of dropouts and incompleteness of learners who are constantly in trouble are emerging as and they must be solved. Therefore, in this study, log data (regularity of learning start interval, total number of learnings and total learning time) and personal background data (courses experience, education experience, gender and age) accumulated in the web through learning management system (LMS) in a distant education environment. The purpose of this study is to analyze the effect of age and educational experience on the total scores that determine completion (more than 60 points) and incompleteness (less than 60 points). The study was carried out with data from 1,130 learners of distance learning centers, which were conducted for a total of 16 weeks. Data was extracted and analyzed based on learning analysis, and the results were as follows: First, among the log data, it was found that the total learning time and number of learning had

significant effects on the total score. As the number of access to the LMS increased, the learning time and total score increased. Second, among personal background data, age was found to have a significant effect on the total score. It was concluded that the probability of completing the study, i.e., the probability of completing the study, increases as age increases, so the purpose of learning becomes more apparent. The data used in this study was used when a learner started signing up for LMS for learning and collected the consent in advance with the consent of the personal information agreement.

Keywords : Learning Analytics, Lifelong Distance Education, LMS (Learning Management System), Log data, Personal background data.

1. Introduction

Distance education has been recognized as an infinite educational potential because of the advantage of open learning suitable for the needs of learners anytime and anywhere without limitation of time and space. In addition, through distance education, it is possible to switch to a lifelong learning society, learners' accessibility becomes easier and teachers can provide up-to-date learning materials to learners without restrictions. In this phenomenon, self-directed learning, where the subject of learning becomes the learner himself, becomes possible through the meeting of information and communication technology and higher education [10][20].

Jong-Teak Seo

Kookmin University, korea
suho1793@hanmail.net

In a distance education environment, accurate information about a learner's learning status is an important factor in determining the success or failure of learning. There is an increasing need to provide learning status information to learners [1][6].

At this time, if the data stored in the learning management system is analyzed, the learner's learning type can be understood. By using this, it can be sufficient educational information data that can be used to predict the learner's learning status [1] [14]. In addition, since all records from the start of learning to the end are stored, a vast amount of data is created and numerous information data are accumulated. However, since most LMSs do not provide tools for instructors to track and analyze all learners' learning activities, learning analytics is a useful approach in predicting a learners' completion of data generated through LMSs. It can be such as [40]. Learning analytics can collect and analyze a learner's activity data generated during the learning process on the LMS to find the regularity in the data and predict future events, as well as analyze learner-related data generated on the LMS. Nowadays, LMS is receiving attention in terms of being able to evaluate and manage learning outcomes [35].

Prior studies using LMS data in relation to the Department of Learning Analysis included regularity of learning time intervals, total learning time, number of learning connections, attendance, degree of interaction between instructors and learners, assignment of tasks, participation in discussions, detailed predictors of major areas, etc. Academic achievement is predicted by analyzing various factors [1] [2] [12] [25] [27]. Among these, a learner's studies is an important information that can be observed and analyzed to understand a learner's online behavior so as to identify improvement points. It affects the learner's total score and helps their learning patterns to be regular. The number of times and total learning time invested for learning were selected as log data. In addition to log data, personal background data is stored in the LMS of the Distance Continuing Education Center. There are previous studies that studied the possibility of incompleteness based on age, gender, final education, occupation, grade, department, area, high school, admission selection, department satisfaction and major suitability corresponding to personal background data [3][17][38]. However, since the research mainly focused on the learner's personal environmental factors, attempts to analyze what the learner's

personal background cause may be are insufficient, which may be a direct cause of the completion. It simply remains in a negative view of limiting the completion to a learner's personal problem or to a maladjustment problem [15]. Therefore, this study compares and analyzes the actual log data (regularity of learning start interval, total number of learnings and total learning time) and personal background data (gender, age, educational experience and number of courses experienced) of learners of distance learning centers. Through this, we would like to verify how each factor affects the total score based on the learning analysis.

2. Theoretical background

2.1. Concept of learning analytics

Recently, learning analytics and educational data mining, etc. have been attracting attention as educational data mining field to improve the educational performance of learners or productivity of educational systems by using big data generated in the educational field [11]. It is expected that the use of big data in the education field will yield improvement effects or problem solving that the existing systems or platforms could not provide. In accordance with the proliferation of big data, the field of education analytics has been divided into learning and educational management analytics [18].

In the past, learning analysis in the traditional approach was intended to improve teaching and learning activities through learner evaluation or questionnaire. However, as a large amount of educational resources are concentrated in the LMS, it provides various functions, such as registration, progress and evaluation of learning courses required for learners, installation of learning contents, management of training courses and evaluation of education. It also provides numerous functions, such as lessons, announcements, interaction, lecture material provision, exams, assignments and discussions between instructors and learners. Participants data, such as the presence and absence of learners, and the duration, number of times, tests, assignments, etc., have been accumulated to measure and evaluate the learners' overall learning activities. Based on the accumulated learning data, various interactions occur and some activity data remains as log records in the learning management system. Learning analytics aims to improve teaching and

learning activities using these learning-related data [13][32]. It has long been done to evaluate and manage learners' performance by analyzing learning-related data. However, in the recent years, as the amount and types of learning data have increased, and analytical techniques for them have been developed, learning analytics is getting more attention [23].

Learners want to be provided with the quality of education provided through educational institutions and learning progress at a glance. For this reason, the distance lifelong education institution needs a system that can improve the learners' academic achievement through learning analysis and collects, analyzes and predicts data so that learners can successfully complete their studies [28]. Learning analytics takes a step further from educational data mining analysis methods whose main goal is efficiency in data analysis, and includes steps to control learning outcomes by applying teaching and learning prescriptions using data mining results [4][30]. Therefore, learning analytics is a research field that applies the predictive model found in the system of education to education beyond the simple technical meaning applied to education data. It provides a platform for learners to successfully complete learning or is useful as a new prediction tool. It can be used as in [16][19]. It also helps learners to predict difficult problems (tasks, discussions, post management, etc.) that students may experience during the learning process in advance. This is to ensure that they can finish their studies normally through proper prescription [31]. In the online learning environment, learning analysis is a prescription for learning, such as predicting learning outcomes and providing advice by analyzing data generated through various learning activities and digital footprints that learners left in the learning process via a computer. Learning analysis that presents a study becomes possible [19][26]. Therefore, by providing the results of the learning activities in the learner's online learning environment visually, it helps the learners to achieve self-directed learning, set goals themselves and manage the course and time to attain the results [34][39].

2.2. Learning analysis advanced research

In recent studies related to learning analytics, LMS log data was used to analyze the factors that affect learners' learning outcomes or incompleteness, or to focus on research that implements a system to analyze the learners' learning status.

In a study examining the relationship between log data of learners accumulated on LMS and learning outcomes, Kim (2011) analyzed the effects of regularity of learning time intervals, total learning time and number of learning accesses on academic achievement. As a result of analysis, it was found that the regularity of the learning time interval had a positive effect on academic achievement, but the total learning time and number of learning connections did not affect academic achievement. In other words, the analysis results suggested that having a lot of numerical learning time does not necessarily increase academic performance, and that frequent access to learning does not lead to good learning outcomes.

Han and Jeon (2015) analyzed the causal relationship between log data and academic achievement for business people taking e-learning. There was no statistically significant correlation between the indicators, such as the number of accesses to learning, time of study and number of days connected, and the indicators of academic achievement, such as the total score and completion rate. It was confirmed that learning performance did not increase simply because there were a large number of learning hours and number of learning connections. However, there was a statistically significant causal relationship between learning pattern and academic achievement.

In a study that analyzed LMS log data of online course students in general universities, Kim (2013) predicted the regularity of learning start intervals, total learning time, number of learning connections and academic achievement to analyze the relationship between learner time management strategy and academic performance. A model was then developed. In particular, the regularity of learning time intervals was found to have a positive effect on academic achievement. Subsequently, through the academic performance prediction model, the time to predict students' academic performance significantly before learning was completed was verified. As a result, it was found that, after the learning period of at least the second quarter elapsed, it was possible to predict academic performance with only some time management strategy data of learners.

In a study on the behavioral variables of learners affecting academic achievement, Hwang (2016) based on the learner's web log data, time management (total learning time, number of learning access and regularity of learning time interval) behavior data. A

study was conducted to analyze the effect on achievement. As a result, it was confirmed that the regularity of the number of learning connections and learning time interval had a significant effect on academic achievement. Moreover, the effect of the learning motivation on the time management behavior and academic achievement had a moderating effect only in relation to the regularity of the learning time interval.

You (2014) used LMS log data of liberal arts e-learning course students operated by a general university to determine how much attendance, total learning time and number of accesses to drop out of scheduled learning for each parking period. Logistic regression analysis was performed. As a result of the study, attendance and total learning time were significantly analyzed for dropout prediction, but the total number of connections was not significant. It was confirmed that regular access to regular learning is an important factor to prevent learning persistence and dropout.

2.3 Learning analytics analysis procedure

The analysis process of learning analytics starts from the data collection stage. At this stage, activity data related to personal learning activities and learning elements in the LMS is collected. This includes log data, etc. In the data collection stage, it is important to solve the learner's personal information problem. This is because personal information leakage may occur during data extraction and processing. Next, after processing, such as grouping, classification and analysis of related data, the results are provided to the learner through a dashboard or other visualized graphic means. The analyzed data can also be provided to the instructor to be used for research on improving teaching methods related to learning methods and outcomes, or used to take appropriate prescriptions and measures for the unexpired risk group of learners [29].

2.3.1. Data collection and processing

In the data collection and processing stage, it can be divided into static data that does not change well in a short time and dynamic data generated during the learning process. Static data has a refined form, i.e., it is systematically arranged according to the scale for each clearly distinguished variable, and it is easy to combine between different databases through a relational database design. Static data includes the characteristics of learners managed by the learning

management system (LMS), demographic information (age and gender) and results of prior learning (number of courses experienced and educational experience). Dynamic data is data generated during the course of learning, and includes learner log data (regularity of learning start interval, total learning time and number of learning) and learning performance data (attendance, assignment, test and total score). Raw data of dynamic data is usually unstructured 'big data', which has no fixed form or cannot be operated on. Therefore, it cannot be analyzed in full without pre-processing and filtering. The data collection and processing stage is a process of filtering only the main data required through the characteristics of the data and converting it into an analytical form [21][37].

For example, if the total learning time is used as a predictor to predict the learner's total score and self-reported questionnaire data is collected, the learner is asked how much time is being devoted to learning the subject. Then, they will answer. After the arithmetic process, the learner's answer is converted into a standardized indicator of the total learning time. On the other hand, if the total learning time is calculated based on the dynamic data called log data, there is no standardized data called 'total learning time' in the log data. Therefore, the calculation for this can vary sufficiently depending on how to define the variable they have. Therefore, the data collection and processing step is a process of filtering only the main data required through the characteristics of the data and converting it into an analyzable form. This pre-treatment process can be divided into four main steps.

First, in the data cleaning step, the missing data, noise and outlier data are selectively removed or adjusted, and the data consistency problem is solved.

Second, in the data reduction step, there are dimensional reduction, number reduction and data compression methods, and even if less data is obtained than the raw data, integrity with the raw data must be maintained.

Third, in the data transformation stage, a strategy is needed to transform or integrate data into a form that is suitable for data mining.

Fourth, in the data integration stage, data is integrated from various databases, and it is necessary to reduce unnecessary data and eliminate data that does not match duplication [18][26].

2.3.2. Data analysis and prediction

For successful analysis and prediction, it must be possible to sufficiently reflect the characteristics of learners who fail or succeed in learning, along with selecting an effective algorithm. In the analysis and prediction phase, the processed data are analyzed in earnest. It is divided into summary, description and prediction, and the summary shows the data visualized in an easy-to-understand and clear manner mainly using charts, scatter plots, graphs and 3D shapes [21][33].

The specific analysis step refers to a step of analyzing a pattern of learner statistical data using static and dynamic data. It is the process of analyzing meaningful results using the refined data or predicting future events.

Prediction has the advantage of analyzing the learning patterns of learners with or without the possibility of unfinished risks and suggesting appropriate measures at individual time points. For example, a warning message can be sent to a learner who is close to the unfinished risk level and can guide the learning activity to enter the normal or safety level. So far, the most reliable method of judging whether or not a student has completed by the total score was a method of judging by taking a test at the start of learning. However, since this method has limitations that are static and cannot reflect the learning status, it can be said that the prediction method is more effective in order to judge based on the real-time learning status and level [21].

3. Research method

3.1. Research problem

The total score that determines whether the LMS log data (regularity of learning start interval, total learning time and total number of learning) and personal background data (number of subjects, gender, age and educational experience) in distance learning education are completed (60 points or more: completed, less than 60 points: not completed). How does it affect?

1.1 How does the regularity of learning start interval affect the total score?

1.2 How does the total learning time affect the total score?

1.3 How does the total number of lessons affect the total score?

1.4 How does the number of course experiences affect the total score?

1.5 How does gender affect the total score?

1.6 How does age affect the total score?

1.7 How does education experience affect the total score?

3.2. Research subject

The study subjects were learners who had conducted remote classes at A Remote Lifelong Education Center for 16 weeks, with 771 male students and 359 female students, totaling to 1,130 students. Table 1 shows the total learner ratios.

Table 1: Subject Gender Ratio

Male		Female		Total	
Number of people	ratio(%)	Number of people	ratio(%)	Number of people	ratio(%)
771	68.2	359	31.8	1,130	100

3.3. Definition of Terms

The data used in this study is accurate data based on facts because the data collected through LMS is based on data entered by the learner during membership registration and learning activity data generated during the learning process. The log data generated during the course of learning is the total number of learning sessions, which shows the regularity of the learning start interval and total number of times the learner has learned, and the standard deviation at the time each learner starts learning through the 16-week learning process. It is composed by re-processing the raw data extracted from the LMS database as an Excel file to match the research purpose by summing up all the times of accessing the learning window and viewing the learning page as the total learning time.

3.3.1. Log data

In this study, log data refers to 'regularity of the learning start interval', 'total learning time' and 'total learning frequency' calculated based on all records made when the learner accesses the LMS.

1. Regularity of learning start interval

Regularity of the learning start interval is a variable representing the regular learning pattern of learners and was classified as an index to measure 'regular class participation'[2].

This refers to how regularly the learner accesses the learning window and the standard deviation of the values obtained in minutes by the difference between the start times of each learning. That is, an interval ($\Delta t_1, 2$) between the start point A and end point B at which a learner of the subject has undergone the learning process, and between the first access point (t_1) and second access point (t_2) for the subject. Is referred to as the 'login interval', n and the previous connection interval can be expressed by $\Delta t_{n-1}, n$. In this way, the standard deviation value of the access interval was calculated based on the deviation between the learner's access interval values and average value.

2. Total learning time

The total learning time was calculated by summing the time that the learner accesses the learning page and accesses all the learning pages of the subject from the initial learning to the later reviewed learning time. In this study, the total learning time was calculated in terms of minutes.

3. Total number of learning

During the learning period, the number of reading the learning pages was limited. The total was calculated by adding all the records that the learner accessed in the corresponding parking learning window for learning within a predetermined learning period. Even if it was accessed multiple times in a parking lot, all of the values were added and included in the total number of lessons.

3.3.2 Personal background data

This refers to the gender, age and educational experience entered by the learners when they sign up for membership, and the number of courses experienced by learners automatically managed by the LMS.

4. Course Experience

The number of course experience courses refers to

the history of distance learning at this institute at the time of analysis among the selected course learners. The number of courses taken by students is displayed as a constant. When collecting data used in the study, data was extracted by excluding duplicate students during the measurement period on the raw data.

5. Education background

The educational background represents the final educational background of the learner's current state. This is the information entered by learners when they first signed up for membership on the website. Then the types were collected in three; high school graduate, college graduate and 4-year university graduate. The three variables used in the analysis were analyzed by dummy variables

6. Gender

Gender was used for the actual input value, and was used for analysis by a dummy variable.

7. Age

It was analyzed using the actual age.

8. Total score

The total score represents the final score calculated based on the 100 points at the time when all the learning processes were completed according to the evaluation criteria. The final grade consists of midterm (30%), final exam (30%), attendance (15%) and assignment (25%). The total score was extracted by selecting one subject with the highest number of participants and using only 1,130 data excluding missing values.

3.4. Data analysis results

After collecting LMS raw data in MS-excel format, the data collected for this study were reprocessed to meet the research objectives and variables. Data was then analyzed using the IBM Statistical Package for the Social Science (SPSS) 21 program.

3.4.1. Descriptive Statistics

The average and standard deviation, minimum and maximum values, skewness and kurtosis of the total scores of the dependent variables were obtained through general trends for the analysis variables and

are presented in table 2. It can be seen that all the variables corresponding to the measured variables form a normal distribution with a skewness of ± 2 or less and a kurtosis of ± 7 or less.

Table 2 : Descriptive Statistics

Variable	Average	Standard Deviation	Minimum value	Maximum value	Skewness	Kurtosis
Regularity of learning start interval	4.037	.782	.731	8.660	1.111	3.162
Total learning time	1,389.404	406.451	3.4	3,744.0	1.699	5.118
Total number of learning	47.009	14.247	7.000	129.000	.708	1.489
Course Experience	10.050	7.6497	1.0	44.0	1.611	2.437
Gender	.682	.466	0.0	1.0	-.784	-1.388
Age	27.804	5.986	20.0	53.0	1.275	1.485
Educational background	1.453	.641	1.0	3.0	1.105	.081
Total score	74.387	16.380	0.0	98.0	-1.928	4.684

3.4.2. Correlation analysis

In this study, the correlation between each variable was analyzed. The regularity of the learning start interval showed a negative (-) correlation with the total learning time and number of learning, age, educational experience and total score. However, it did not show any correlation with the number of subjects and gender. The total score was significant for all the variables except for the number of courses. Negative (-) correlation with regularity at the beginning of the learning interval was found to be positive (+) between the remaining variables [Table 3].

Table 3 : Correlation analysis

Variable	1	2	3	4	5	6	7	8
1. Regularity of learning start interval	1							
2. Total learning time	-.228**	1						
3. Total number of learning	-.500**	.302**	1					
4. Course Experience	.016	-.003	-.072*	1				
5. gender	-.001	.037	.059*	.087**	1			
6. age	-.075*	.159**	.095**	.095**	.342**	1		
7. Educational background	-.070*	.051	.087**	.045	.156**	.216**	1	
8. Total score	-.190**	.240**	.325**	.021	.084**	.168**	.108**	1

* p < .05, ** p < .01

3.4.3. Regression

Table 4 shows the results of the multiple regression analysis. The final model with all independent variables had an F value of 26.429 and significant probability of .000, which showed a significant result

in the statistical regression model at a significance level of .05. It was concluded that 15.9% of the total score (adj. $R^2 = 15.3$) was accounted for by the independent variables.

Table 3. Correlation analysis

Dependent variable	Independent variable	Denormalization factor B	Standard error	Standardization factor β	T	P
	(a constant)	43.060	4.700		10.013	.000
Log data	Regularity of learning start interval	-.332	.707	-.016	-.469	.639
	Total learning time	.004	.001	.101	3.513	.000
	Total number of learning	.221	.031	.257	7.186	.000
Totalscore	Course Experience	-.071	.057	-.037	-1.258	.290
	Gender dummy	-1.378	.927	-.043	-1.486	.138
	Age	.344	.075	.131	4.581	.000
	Educational background dummy 1	2.654	.961	.080	2.762	.006
	Educational background dummy 2	-.568	1.666	-.010	-.341	.733
$R^2(\text{adj. } R^2) = .159(.153)$ F = 26.429, p = .000						

*p < .05n = 1,13(High school =0)

4. Discussion

Through this study, we tried to find out how log data and personal background data stored on the web affects each the total score. Through this, a study was conducted to find a way to predict incompleteness in advance, inform learners at risk of unfinished risks, as well as to inform risks of unfinished risks by using SMS, as well as various warning messages, such as telephone and e-mail, and induce them to actively study.

5. Conclusion

This study is based on objective data because research was conducted based on actual LMS data of learners using distance learning centers. The results of the study were summarized as follows.

First, it was found that the total learning time and number of learning had a significant effect on the total score. In a previous study, Kim (2003) stated that the longer the time required for learning access, the higher the total score. Kang, Kim and Park (2009) showed that the attendance rate, number of discussions, classroom access time and bulletin board posting times significantly influence the prediction of the total score. In addition, Rau and Durand (2000) suggested

that the learning time accumulated in LMS is the main variable for predicting the total score. In Kwon (2009) study, the total score was higher for the adult learners of the remote graduate school. Thus, a higher the number of responses and the number of discussions indicating participation in learning. In other words, the study suggested that it is difficult to expect good learning results without active participation of learners. Jeon and Han (2015) showed that the number of LMS accesses and time of learning in the university online learning process are variables that can significantly predict the total score.

Second, among personal background data, age was found to have a significant effect on the total score. Jeong (2004) stated that the higher the age of the learner's demographic characteristics (gender, age, grade, education, occupation and residential area), the higher the sense of purpose for study. Therefore, the higher the probability of completion. Lim (2007) said that the higher the age among the learner-related variables (gender, age, education, rating, completion rate, job and residence), the lower the probability of completion. Kim (2012) added admission, age, gender and job. It was analyzed that only age among classification, residence and final education had a significant relationship with incompleteness. It can be estimated that the younger the age, the more economic burden of tuition provision, thus leading to incompleteness. In particular, in the 20s, when work and study were performed in parallel, there was a clear tendency to give priority to work, and there was a high probability of not completing the study due to lack of learning time. In addition, people in their 20s are not only likely to choose a variety of degree acquisition institutions other than distance learning institutions, but also can be considered to have a high rate of completion due to relatively poor time and economic conditions. However, due to personal circumstances, people in their 30s and 40s have the highest uncompleted rate, and those in their 40s with the lowest uncompleted rate can be expected to be more serious in their studies with stable status and time in the workplace. For this reason, it can be seen that a high rate of unfinished students was in their 20s with relatively poor temporal and economic conditions [9] [22].

Next, I would like to make the following suggestions for further research. First, in this study, a study was conducted in advance to predict and classify learners who may be at risk of not completing their studies. However, research on learning support methods, such

as dashboards, in which feedback from teachers can be effectively delivered to learners has not been conducted. Second, through guidance on effective and efficient learning activities, including the initial time management method, it is possible to provide information in various forms by automatically analyzing data generated by learners so that learners can actively check their learning. Research on specific teaching design strategies is needed[41].

References

- [1] J. H. Kim, "Impact of Learner's Time Management Strategies on Achievement: A Learning Analytics Approach", The Graduate School of Ewha Womans University, Seoul Korea, 2013.
- [2] Y. M. Kim, "Impact of Regularity of Learning Interval, Total Learning Hours and Number of Access to E-learning in a Corporate E-learning Environment on Academic Achievement", The Graduate School of Ewha Womans University, Seoul Korea, 2011.
- [3] S. Y. Kim, "An Analysis of College Student Dropouts' Mobility Paths and Structure", The Journal of Educational Studies, Vol. 43, No. 3, pp. 131-163, 2012.
- [4] K. G. Kim, "Learner Activity Modeling Based on Teaching and Learning Activities Data", Software and Data Eng, Vol. 5, No. 9, pp. 411-418, 2016.
- [5] S. T. Kim, "A correlation between Learning Behavior and Achievement Level of Learners in e-Learning," Korea University of Technology and Education, Chungcheongnam-do, Korea, 2003.
- [6] E. H. Kim, M. K. Park, Coreference resolution of Korean anaphoric zero objects: Towards a supervised machine learning approach. International Journal of Computer Science and Information Technology for Education. Vol. 1. No. 1. Dec. 2016. GVPress. pp:1-6.
- [7] M. S. Kang, J. I. Kim, and I. W. Park, "The Examination of the Variables related to the Students' e-learning Participation that Have an Effect on Learning Achievement in e-learning

- Environment of Cyber University”, Korean Society For Internet Information, Vol. 10, No. 5, pp. 135-143, 2009.
- [8] S. Y. Kwon, “The Analysis of differences of learners' participation, procrastination, learning time and achievement by adult learners' adherence of learning time schedule in e-Learning environments”, Journal of Learner-Centered Curriculum and Instruction, Vol. 9, No. 3, pp. 1-86, 2009.
- [9] Y. R. Kim, “Investigation of the relationship between university students characteristics and dropout factors”, Graduate School of Education, Kyungpook National University, Kyungpook Korea, 2012.
- [10] J. Y. Shin, O. R. Jeong and D. S. Cho, “The Analysis of individual Learning Status on Web-based Instruction”, The Korean Association of Computer Education, Vol. 6, No.3, pp. 107-120, 2003.
- [11] J. H. Shin, J. W. Choi and W. Koh, “A study on the Use of Learning Analytics in Higher Education: Focusing on the perspective of professors”, Journal of Educational Technology, Vol. 31, No. 2, pp. 223-252, 2015.
- [12] J. W. You, “Dropout Prediction Modeling and Investigating the Feasibility of Early Detection in e-Learning Courses”, The Korean Association of Computer Education, Vol. 17, No.1, pp. 1-12, 2014.
- [13] J. Lee, “LCMS(Learning Content Management System) based e-Learning Development and Application”, Korean Association for Educational Information and Media, Vol. 8, No. 2, pp. 93-113, 2002.
- [14] H. Y. Lee, “Development of prediction models based on the clustered online learners' behavioral patterns in university e-Learning environment”, Graduate School of Ewha Womans University, Seoul Korea, 2016.
- [15] Y. W. Lim, “A substantial study on the Relationship between students' variables and dropout in Cyber university”, Journal & Artical management System, Vol. 11, No. 2, pp. 205-219, 2007.
- [16] M. L. Ahn, Y. Y. Choi, Y. H. Bae, Y. M. Ko, and M. H. Kim, “A Literature Review on Learning Analytics: Exploratory study of empirical researches utilizing log data in Korea”, Journal of Educational Technology. Vol. 32, No. 2, pp. 253-291, 2016.
- [17] J. Y. Chung, M. S. Sun, and M. J. Jeong, “An Analysis of Institutional Factors Affecting on College Dropout Rates” Education Research Institut, Vol. 16, No. 4, pp. 57-76, 2015.
- [18] Y. S. Cho, and R. J. Abel, “Prospects for the Application of Learning Analytics” RM 2013-15, KERIS.
- [19] I. H. Cho, “Learning analysis and learning design, development of the convergence horizon”, Journal of Educational Technology, Vol. 2015, No. 2, pp. 422-434, 2015.
- [20] B. M. Chang. Cross-Cultural Distance Learning Project for Promoting English Language Proficiency and Cultural Awareness of University Students in Korea. International Journal of Computer Science and Information Technology for Education. Vol. 1. No. 1. Dec. 2016. GVPress.
- [21] I. H. Cho, Y. J. Park, and J. H. Kim, Understanding Learning Analytics, Seoul : Young sa Park, 2019, pp. 76-92.
- [22] Y. J. Joo, W. J. Shim, and S. M. Kim, “A Study on the Factors Affecting the Drop-out in Corporate Cyber Learning”, The Journal of Educational Information and Media, Vol. 14, No. 4, pp. 5-25, 2008.
- [23] Y. H. Jeong, “Education Analytics”, MEDIA & EDUCATION, Vol. 5, No. 1, pp. 44-49, 2015.
- [24] B. Y. Jeong, “Analyses of Learning Achievement and Satisfaction on Demographic Characteristics in Cyber Universities : A Case Study”, The Journal of Educational Information and Media, Vol. 10, No. 3, pp. 127-150, 2004.
- [25] D. H. Han, and B. J. Jeon, “Analysis of Learning Type Factors that Affect e-learning Performance : Centering on the Comparison Analysis of

- Whole Learners Log and Excellent Learners”, *Journal of the Korean Data Analysis Society*, Vol. 17, No. 2, pp. 897-912, 2015.
- [26] J. S. Han, Y. M. Kim and M. K. Kang, "Suggestions to Support 'Flipped Learning' on Stage Based on the Survey of Primary and Secondary Teachers' Recognition. *Asia-Pacific Journal of Educational Management Research*. Vol. 1. No. 1. Dec. 2016. GVP Press. pp:145-154.
- [27] S. R. Hwang, "Impact of Learner's Learning Behavior on Achievement: The Moderating Effect of Learning Motivation", *The Graduate School of Education Ewha Womans University*, Seoul Korea, 2016.
- [28] Dietz-Uhler, B., & Hurn, J. E. Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning*, Vol. 12, No. 1, 17-26. 2013.
- [29] Dyckhoff, A. L., Zielke, D., Bultmann, M., Chatti, M. A., & Schroeder, U, "Design and implementation of a learning analytics toolkit for teachers", *Educational Technology & Society*, Vol. 15, No.3, pp.58-76, 2012.
- [30] Elias, T, *Learning Analytics: Definitions, Processes, and Potential*. Creative Commons, 2011.
- [31] Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. *The 2011 horizon report*. Austin, Texas: The New Media Consortium, 2011.
- [32] Hall, R, "Schedules of Practical Work For the Analysis of Case Studies of Learning and Development", *The Journal of the learning science*, Vol.10, No.1-2, pp.203-222, 2001.
- [33] Liao, S. H., Chu, P. H., & Hsiao, P. Y, "Data mining techniques and applications—A decade review from 2000 to 2011", *Expert Systems with Applications*, Vol. 39, No. 12, pp.11303-11311, 2012.
- [34] Long, P, & Siemens, G, *Penetrating the fog: Analytics in learning and education*. EDUCAUSE Review, Vol. 46, No.5, pp.30-32, 2011.
- [35] Macfadyen, L. P., & Dawson, S, "Numbers are not enough. why e-learning analytics failed to inform an institutional strategic plan", *Educational Technology & Society*, Vol. 15, No. 3, pp. 149-163, 2012.
- [36] Rau, W., & Durand, A, "The Academic Ethic and College Grades: Does Hard Work Help Students to 'Make the Grade'?", *Sociology of Education*, Vol. 73, No.1, pp.19-38, 2000.
- [37] Shum, S, B, *LEARNING ANALYTICS*, UNESCO Institute for Information Technologies in Education, Policy Brief. November 2012.
- [38] Tinto, V, *Leaving college: rethinking the causes and cures of student ttrition*. Chicago: University of Chicago Press, 1987.
- [39] Wagner, E., & Ice, P, "Data changes everything: Delivering on the promise of learning analytics in higher education", *EDUCAUSE Review*, Vol. 47, No. 4, pp. 32-36. 2012.
- [40] Zorrilla, M. E., Menasalvas, E., Marin, D., Mora, E., & Segovia, J, *Web usage mining project for improving web-based learning sites*. In *Web Mining Workshop Cataluna*(pp.1-22), 2015.
- [41] Seo, J.T., Kim, Y.G. & Ju, R. (2020). The Effect of LMS Data on Total Score of Learner's in Lifelong Distance Education Center: A Learning Analytical Approach. *Journal of Education and Social Science (JESS)*, HolyKnight, vol. 1, 15-22. doi: 10.46410/jess.2020.1.1.03.