A LECTURE SUPPORTING SYSTEM BASED ON REAL-TIME LEARNING ANALYTICS

Atsushi Shimada and Shin'ichi Konomi
Kyushu University, Japan

ABSTRACT

A new lecture supporting system based on real-time learning analytics is proposed. Our target is on-site classrooms where teachers give their lectures, and a lot of students listen to teachers' explanation, conduct exercises etc. We utilize not only an e-Learning system, but also an e-Book system to collect real-time learning activities during the lectures. The proposed system is useful for a teacher just before lecture starts and during the lecture. The system provides summary reports of the previews of given materials and quiz results. The teacher can check which pages were well previewed and which pages were not previewed by students using the preview achievement graph. Additionally, the teacher can check which quizzes were difficult for students, and the suggested pages that should be explained in the lecture to aid students. During the lecture, real-time analytics graphs are shown on the teacher's PC. The teacher can easily grasp students status whether or not students are following the teacher's explanation. Through a case study, we confirmed the effectiveness of the proposed system, in terms of high synchronization between a teacher and students, i.e., the teacher adjusted the speed of his lecture based on the real-time feedback, and many students followed the teacher's explanation.

KEYWORDS

Learning analytics, real-time feedback, on-site classrooms, lecture supporting system

1. INTRODUCTION

Learning analytics (Khalil 2016) has attracted many researchers recent years. One of the important issues in learning analytics is not only to perform analysis of data about learners, but also to realize feedbacks for optimizing the learning environment and learners themselves. There are roughly three types of feedback loops in terms of their frequency: yearly, weekly and real-time feedbacks. A typical example of a yearly (or term-by-term) feedback loop is the assessment and improvement of education. Students' grades, examination results, class questionnaires, and so on are typically analyzed and evaluated (Mouri 2016, Okubo 2016). The yearly feedback loop is designed so that the feedback results will be delivered in the next year (or term). In other words, students and teachers will not directly receive the feedback results acquired by analyzing their own learning logs. A weekly feedback loop can recommend related material based on student status determined using a prediction of academic performance through the analysis of learning logs such as attendance reports and quiz results (Shimada 2015, Shimada 2016). In contrast to a yearly feedback loop, the analysis results are directly fed back to the students and teachers who provide the learning logs. The main difference between weekly and real-time feedback loops is that the analysis results can be fed back to the on-site students and teachers even during a lecture. A teacher can check what students are doing, e.g., whether students are following the explanation, or whether they are doing something not related to the lecture. A teacher can flexibly control the speed of the lecture, and/or take more time for exercises rather than a nonstop talk, and so on.

There are several related work, which tackles real-time learning analytics. Minovic et al. proposed a visualization tool for teachers to track students learning progress in real-time, while in gameplay session (Minovic, 2013). Piech et al. collected tens of thousands of program codes, and applied a machine learning approach to identify "sink" states of students. A feedback is achieved for students just before they are about to enter such problematic "sink" states (Piech 2012). Fu et al. also proposed a real-time analysis of program codes (Fu 2017). They provides a learning dashboard to capture the behavior of students in the classroom and identify different difficulties faced by students. Although these studies realize real-time feedbacks, the target of the analytics and its feedback are activities in virtual learning environments.

Our study has focused on feedback, specifically, how to feedback efficient information to on-site classrooms even during lectures. The aim of this research is to realize real-time feedback, which has not often been discussed with respect to on-site educational environments. Our target is on-site classrooms where teachers give their lectures, and a lot of students listen to teachers' explanation, conduct exercises etc. In such a large classroom, it is not easy for teachers to grasp students’ situations and activities. We utilize not only an e-Learning system, but also an e-Book system to collect real-time learning activities during the lectures. We have developed two main feedback systems. One is useful for a teacher just before lecture starts. The system provides summary reports of the previews of given materials and quiz results. The teacher can check which pages were well previewed and which pages were not previewed by students using the preview achievement graph. Additionally, the teacher can check which quizzes were difficult for students, and the suggested pages that should be explained in the lecture to aid students. The other is real-time analytics graphs which are helpful for the teacher to control his/her lecture speed during the lecture. The system collects e-Book logs operated by students sequentially, and performs analytics in real time how many students are following the teacher’s explanation. In the rest of this paper, we introduce the details of our real-time feedback system and report experimental results.

2. IMPLEMENTATION

2.1 Learning Logs

The e-Book logs were collected via an e-Book system called “BookRoll”. Figure 1 shows samples of e-Book logs. There are many types of operations recorded in the logs, for example, OPEN means that the student opened the e-book file and NEXT means that the student clicked the next button to move to the subsequent page. The browsing duration for each page can be calculated by subtracting the subsequent timestamps. Learning logs on the e-Leaning system such as attendances and quiz scores are collected from tables in the Moodle database. The system analyzes the quiz scores and class attendances by integrating related tables.

User	Material	Operation	PageNo	Date	Time
X	00000000NLAT	OPEN	0	2014/10/15	9:01:09
X	00000000NLAT	CLOSE	1	2014/10/15	9:01:13
Y	00000000P82P	PREV	25	2014/10/29	10:05:35
Y	00000000P855	NEXT	2	2014/11/19	8:52:47
Z	00000000P84Z	NEXT	9	2014/11/12	9:31:30
...	...	...	...	...	...

Figure 1. Samples e-Book logs

2.2 On-site Lecture Supporting System

We present the example case study shown in Figure 2, which was actually applied to a lecture in our university. The time line is divided into two parts: before starting a class and during a class. During the previous lecture, a teacher gave students some preview materials, that were automatically generated by summarization technique (Shimada 2017). Students previewed the given materials before the class, and the operation logs recorded during the previews were collected by the system. Before the class started, students answered the quizzes and the results were collected on the server.

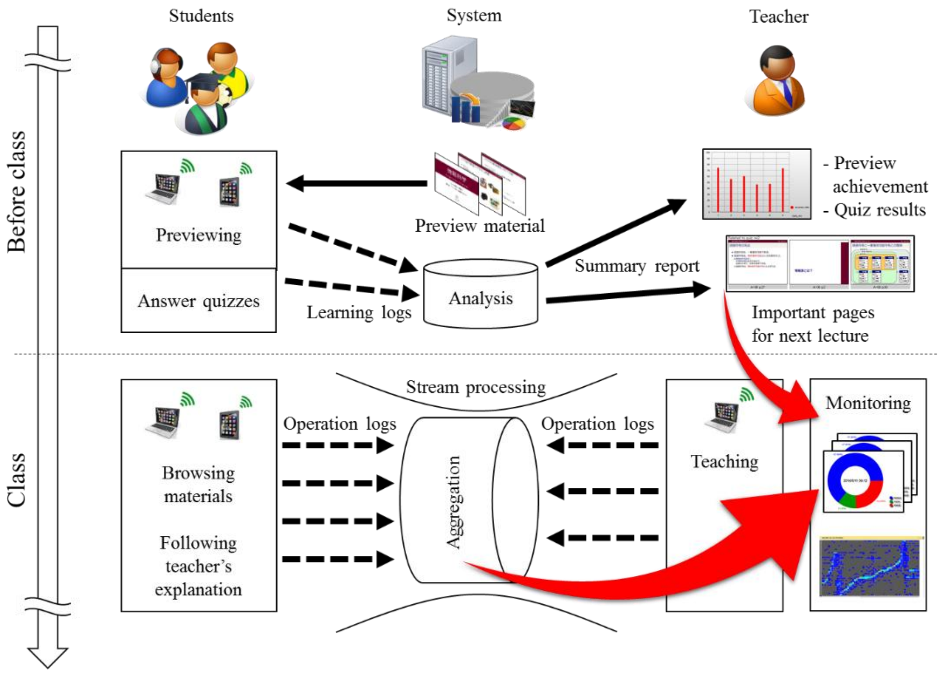

Figure 2. A case study. Students previewed materials before their class and answered the quizzes. Learning logs (e-book operation logs and quiz results) were collected in the system and summary reports were sent to the teacher. The teacher checked the preview achievements (which page was browsed the most, which page was not read, etc.) and quiz results before starting their lecture. Additionally, they were told some important pages that should be explained well in the next lecture. During the lecture, the operation logs of the teacher and students were collected and analyzed. The results were immediately visualized on the teacher's PC. The teacher could check the students' activities (e.g., how many students were following the lecture) and adjust the their speed accordingly

Just before the lecture started, our system analyzed the learning logs to make a summary report containing previews of the achievement and quiz results (details are given in the section 2.6. Additionally, the system provided information regarding important pages that should be explained well in the lecture. For example, the teacher should focus on pages that are related to the quizzes, especially those that have led to lower quiz scores. Our system analyzed the relationship between quiz statements and their related pages in the lecture material in advance. Section 2.3 explains how we automatically discover important pages.

During the lecture, a teacher explained the contents of the materials and students browsed the materials in their lap-top PCs. In our university, students are asked to open and browse the same page as the teacher, and to put highlights or notes on the important points. During the lecture, learning logs were sequentially collected and stored. The analysis results were immediately visualized on the web interface, and updated each minute. Therefore, the teacher could check the latest student activities. The visualization included real-time information regarding how many students were following the lecture, how many students were browsing previous pages, etc. The web interface is described in section 2.6. The teacher adaptively controlled the speed of the lecture according to the students. For example, if many students were not following the lecture and were still on previous page, the teacher slowed down the lecture.

2.3 Important Page Mining

There are strong relationships between lecture materials and quizzes, because quizzes are often generated using the contents of the lecture materials. Related pages are important to understanding the contents of the materials. However, lecture materials and quizzes are stored separately, or are very weakly connected in the system using, for example, subject names. We can manually assess the relationship between a quiz item and its related pages, but this is not easy nor realistic when the number of quiz items and/or the number of pages increase. Furthermore, if the lecture material is updated, i.e., the page numbering changes, the teacher must

ISBN: 978-989-8533-68-5

update the correspondence. Therefore, we developed a method that automatically determines the correspondences.

Our strategy assumes that a related page contains the same keyword as the quiz statement. Each quiz statement, QS, is divided into morphemes. Then, we extract the nouns $n\ (1, \cdots, n, \cdots, N)$. For each noun n , a normalized histogram h_n is created. Each bin $b_{u,n}$ of the histogram h_n represents how many times page u contains noun n. Note that the bins are normalized after counting the number of times noun n appears in all the pages. To acquire the final result, we sum the frequencies of all nouns. We define the normalized value r_u as the related score of page u.

Although the mining method finds pages that are highly related to a given quiz statement, it does not consider the relationships among pages. Therefore, we also apply a ranking method that assigns a ranking score to each page. The idea was inspired by VisualRank (Jing 2008). A ranking vector R is iteratively updated using

$$R = \alpha(S \times R) + (1 - \alpha)B$$

where S is the column normalized similarity matrix, and $S_{u,v}$ measures the page similarity between pages u and v. In this study, we simply evaluate the similarity using the L2 norm between two feature vectors represented by bag of words (Zhang 2010). B is a bias vector. We use the relate score r_u as an element of B. R is repeatedly updated until it converges. α, $(0 \leq \alpha \leq 1)$ controls the balance between the similarity matrix and the bias vector. According to the literature (Jing 2008), $\alpha > 0.8$ is often used in practice. After the ranking vector R converges, pages that are related to important pages have larger ranking scores. We select the top N ranked pages as important.

2.4 Preview Achievement

By analyzing e-book operation logs, we can know how long students spend previewing each page of a given material. The previewing time period for each page can be easily acquired by subtracting two successive time stamps from the operation logs. Note that we ignored durations less than 3 seconds and more than 600 seconds to discard skipped and abandoned pages. Figure 4 shows an example of a visualized result of preview achievement. A teacher can check the preview status of given materials in advance before beginning his/her lecture.

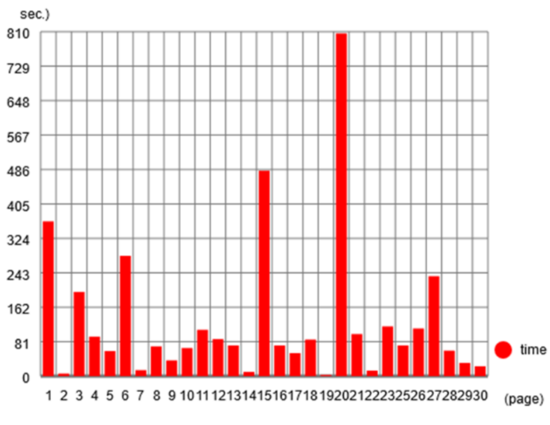

Figure 3. Example preview achievement. The horizontal axis is the page number and the vertical axis is the time spent previewing by students

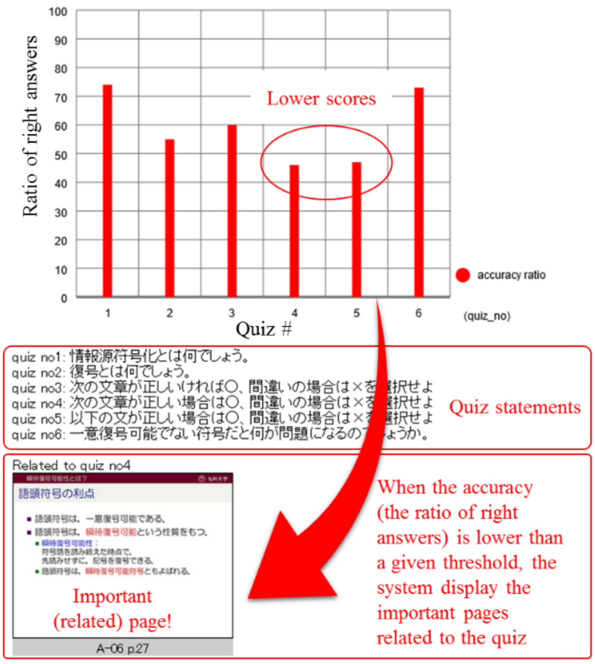

Figure 4. Example quiz results and related information

2.5 Quiz Results

The quiz results and questions are collected from the e-Learning system, and the scores are aggregated in the class. We set a threshold for the ratio of correct answers (in our implementation, we set the threshold to 50), and if the accuracy is lower than the threshold, important pages, which are automatically mined in advance, are displayed below the summary graph. See Figure 3 for an example of the web page. A summary graph of the quiz results is followed by the quiz statements, and related page information if necessary.

2.6 Visualizer on Web Pages

The proposed visualizer of the analysis results was implemented as a web system. A teacher can easily access the web page from a PC. Before the lecture starts, a teacher can access the web pages that provide summary reports of the previews of given materials and quiz results, as shown in Figure 4 and Figure 3. The teacher can check which pages were well previewed and which pages were not previewed by students using the preview achievement graph. Additionally, the teacher can check which quizzes were difficult for students, and the suggested pages that should be explained in the lecture to aid students.

During the lecture, the teacher can access to two kinds of real-time analytics graphs. One is the real-time heat map shown in Figure 5. The horizontal and vertical axes represent the time of day and the page number, respectively. In other words, a vertical line corresponds to the distribution of the number of students who are browsing each page. The vertical lines are updated each minute, i.e., a new line is added per minute. Each cell represents the number of students. The page being explained by the teacher is highlighted by red colored rectangles. If a brighter color (red, orange, yellow, or green) is used on the page being explained by the teacher, most of students are following the teacher's explanation. Students are asked to try to be on the same page as the teacher, and to add highlights or notes if necessary. Therefore, when the distribution of the students is skewed to the below direction, some students are still browsing previous pages. In such a case, the teacher should slow down the lecture so that students can catch up.

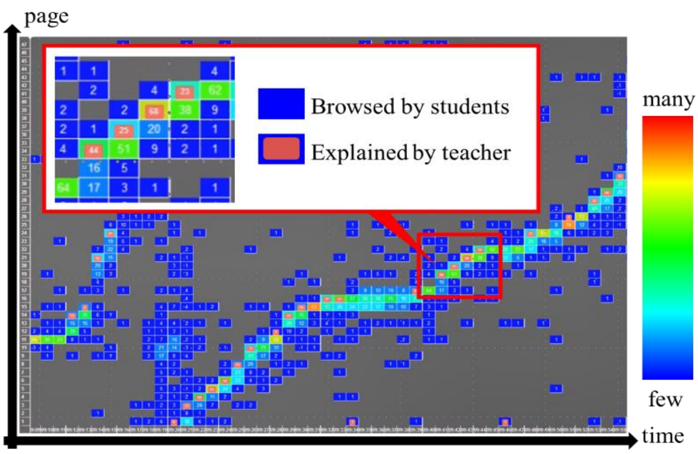

Figure 5. Real-time heat map of browsed pages. The horizontal axis is the time of day, and the vertical axis is the page number. A column corresponds to the distribution of the number of students browsing each page. The page explained by the teacher is highlighted by a red colored rectangle. The heat map is automatically updated minute-by-minute

The other real-time analytics graph is the circular chart (left part of Figure 6), which is a summarized version of above heat map. A teacher could take some time to check and understand the situation from the heat map. To just visualize a summary of the heat map, the second visual focuses on the ratio of three types of students: browsing previous pages (blue), browsing the same page as the teacher (green); and browsing the next pages (red). This chart is also updated each minute to display the latest status of students. The visualizer also provides a breakdown of the three types based on whether students previewed the page in advance (light color) or not (dark color). For example, if many students are still browsing previous pages, most of them had previewed the pages in advance, and the pages are suggested as important ones that are related to a difficult quiz, the teacher should wait for students and explain the material slowly and carefully. Another example is that a teacher should proceed with the lecture when many students are browsing the next pages, and most of them have previewed the materials in advance. In such a situation, students may have got bored during a teacher's long explanation, or some students may have finished a given exercise.

The right part of Figure 6 is a time-series of the circular chart. The teacher can see the recent trends of each status. As described above, the real-time analytics graphs provide an opportunity to flexibly adjust the lecture progress based on the statuses of students.

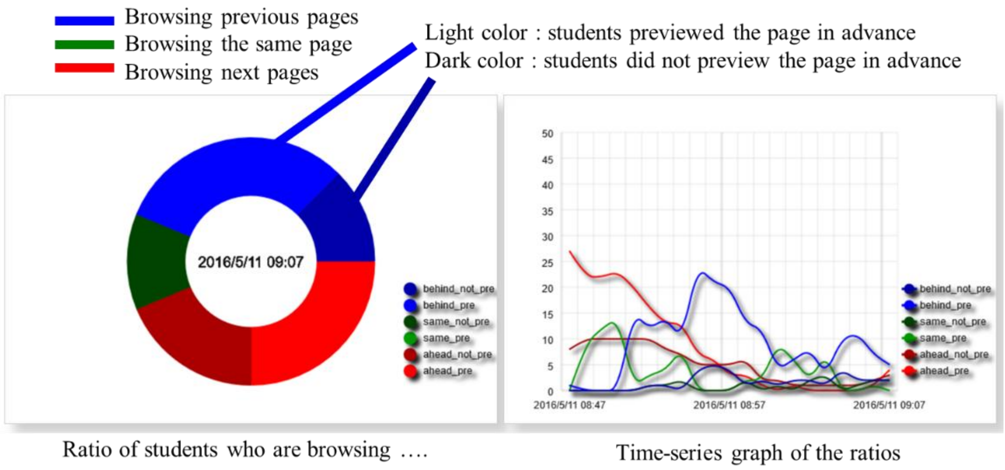

Figure 6. Real-time circular chart of student status. There are three kinds of status: browsing previous pages (blue), browsing the same page as the teacher (green), and browsing next pages (red). In each status, the system uses different brightness of colors to distinguish whether students previewed the page in advance (i.e., previewed the page before the lecture started)

3. EXPERIMENTAL RESULTS

3.1 Settings

We investigated the effectiveness of the proposed system in two classes at Kyushu University, Japan. One was a control group (N = 58) without the system, and the other was an experimental group (N = 157) with the system. The contents of two lectures are completely the same. Students chose one of them according to their schedules accordingly. Therefore, the number of students are not balanced between two classes. The class was designed to provide an introduction to information and communication technology in a number of disciplines. First-year students, including both arts and science students, attended the class, which commenced in October 2016. All of the students brought their own laptops to the class.

The lecture was given by the same teacher using the same materials. The teacher used two materials: material 1 consists of 37 pages, and material 2 consists of 47 pages. The teacher firstly began with the material 1 followed by material 2, and asked students to follow the page of materials with putting bookmarks, highlights and notes. Operation logs were sequentially collected to the server, and real-time analysis was performed. The results were fed back to the teacher minute-by-minute in the class of experimental group only. More details are summarized in Table 1. We conducted a pre-test to check the basic knowledge about information science. There was not significant difference between two groups.

Table 1. Detailed information of each group. n.s.: not significant

	control	experimental	p-value
# of students	58	157	
pre-test average	6.85±2.28	6.99±2.38	n.s.
e-Book logs	16335	39722	
logs / students	281.6±123.3	253.0±129.1	n.s.

3.2 Synchronization

When the teacher gave his lecture to students in the experimental group, he monitored the display on which real-time analysis results were drawn. He controlled the speed of lecture to make students catch up the lecture as much as possible. We evaluated the synchronization of classroom; how many students opened the pages which were explained by the teacher. We counted up the number minute-by-minute with setting allowable delay, which is a short period to accept the delay of e-Book operations.

Table 2 shows the ratio of synchronization of each group. For example, if we set the allowable delay to be 3 minutes (i.e., if students opened the same page with the teacher within 3-minute delay), the synchronization ratio of the experimental group was 0.7661. The score was significantly different from the score of the control group. In other allowable delay settings, the synchronization ratios of the experimental groups were higher than those of the control group. We consider that such high synchronization was realized by the lecture speed control through the real-time feedback of classroom activities.

Table 2. Synchronization ratio of each group in three length of allowable delay. *: $p<0.05$, **: $p<0.01$

	control	experimental	p-value
1 min.	0.4275	0.5174	0.0403 *
3 min.	0.6598	0.7661	0.0033 **
5 min.	0.7508	0.8599	0.0014 **

3.3 Effectiveness of Important Page Suggestion

The analyses of preview status and quiz scores were performed just before the lecture started. The system reported that most students wrongly answered two of eleven questions. The pages related to the quizzes (actually, the page #10 of material 1 and the page #27 of material 2) were shown on the display, and the teacher confirmed them. The teacher spent a little bit longer time for the explanation of these pages. In fact, the page #10 was opened by the teacher for 3 minutes in the experimental group, meanwhile one minute for the control group.

We analyzed the number of bookmarks, highlights and notes on the above two pages, where the teacher emphasized the explanation. About 61% of students used the functions in the experiment group. On the other hand, about 53% of the students in the control group used the functions.

In addition, we analyzed utilization ratios of three functions through the materials, and compared the ratios between two groups. Table 3 shows that more students in the experimental group used the functions compared with the students in the control group. We guess that students in the experimental group had enough time to put bookmark, highlight and/or note because the teacher emphasized the explanation about important pages with adjusting the speed of his lecture based on the real-time situation of the classroom.

Table 3. Utilization ratios of three functions during the lecture

function	control	experimental
bookmarks	0.828	0.904
highlight	0.759	0.834
note	0.293	0.471

4. CONCLUSION

We proposed a lecture supporting system based on real-time learning analytics, which is available in on-site classrooms. Our system provided summary reports of previewing and quiz scores just before a lecture had started. The report was helpful for a teacher to check which pages were well previewed and which pages were not previewed by students using the preview achievement graph. Additionally, the teacher can check which quizzes were difficult for students. Our system automatically suggested related pages that should be explained in the lecture to aid students. Furthermore, real-time analytics graphs were helpful for the teacher to control his/her lecture speed during the lecture. We conducted a case study of real-time learning analytics in our university. Through the experiments, we found out the following things. The proposed real-time learning analytics system supported the on-site lecture in terms of following aspects.

- The teacher could adjust the speed of his lecture based on the real-time feedback system.
- The teacher emphasized the important points which were mistakenly understood by the students.

As the results, the following effects were confirmed.

- The students could catch up the lecture with following pages explained by the teacher.
- Many students put bookmarks, highlights and memos on important pages.

In our future work, we will continue to use the proposed real-time learning analytics system for further evaluation. Besides, we are going to develop other report graphs which support teachers' decision in classrooms.

ACKNOWLEDGEMENT

This work was supported by JST PRESTO Grant Number JPMJPR1505, Japan.

REFERENCES

M. Khalil and M. Ebner. 2016. What is learning analytics about? A survey of different methods used in 2013-2015. *In Proceedings of Smart Learning Conference*, pp. 294-304

K. Mouri. et al, 2016. Bayesian network for predicting students' final grade using e-book logs in university education. *In IEEE International Conference on Advanced Learning Technologies(ICALT2016)*, pp. 85-89.

F. Okubo. et al, 2016. Learning activity features of high performance students. *In Cross-LAK2016.*

A. Shimada. et al, 2015. Automatic summarization of lecture slides for enhanced student preview. *In The 23rd International Conference on Computers in Education.*

A. Shimada. et al, 2016. Automatic generation of personalized review materials based on across-learning-system analysis. *In Cross-LAK2016.*

M. Minovic and M. Milovanovic, 2013. Real-time learning analytics in educational games. *the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, pp.245-251.

C. Piech. et al, 2012. Modeling how students learn to program. *In Proceedings of the 43rd ACM Technical Symposium on Computer Science Education, SIGCSE'12*, pp.153-160.

X. Fu. et al, 2017. Real-time Learning Analytics for C Programming Language Courses. *The 7th International Conference on Learning Analytics & Knowledge Understanding*

A. Shimada. et al, 2017. Automatic Summarization of Lecture Slides for Enhanced Student Preview -Technical Report and User Study-. *IEEE Transactions on Learning Technologies*, Vol.PP, No.99

Y. Jing and S. Baluja, 2008. Visualrank: Applying pagerank to large-scale image search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, pp.1877-1890.

Y. Zhang. et al, 2010. Understanding bag-of-words model: a statistical framework. *Int. J. Machine Learning and Cybernetics*, pp.43-52.