Yang, C. C. Y., Chen, I. Y. L., Akçapınar, G., Flanagan, B., & Ogata, H. (2021). Using a Summarized Lecture Material Recommendation System to Enhance Students' Preclass Preparation in a Flipped Classroom. *Educational Technology & Society*, *24* (2), 1–13.

Using a Summarized Lecture Material Recommendation System to Enhance Students' Preclass Preparation in a Flipped Classroom

Christopher C. Y. Yang[1*], Irene Y. L. Chen[2], Gökhan Akçapınar[3], Brendan Flanagan[4] and Hiroaki Ogata[4]

[1]Graduate School of Informatics, Kyoto University, Japan // [2]Department of Accounting, National Changhua University of Education, Taiwan // [3]Department of Computer Education and Instructional Technology, Hacettepe University, Turkey // [4]Academic Center for Computing and Media Studies, Kyoto University, Japan // yang.yuan.57e@st.kyoto-u.ac.jp // irene@cc.ncue.edu.tw // gokhanakcapinar@gmail.com // flanagan.brendanjohn.4n@kyoto-u.ac.jp // hiroaki.ogata@gmail.com

[*]Corresponding author

(Submitted May 27, 2020; Revised June 10, 2020; Accepted December 25, 2020)

ABSTRACT: Research has revealed the positive effects of flipped classroom approaches on students' learning engagement and performance compared with conventional lecture-based classrooms. However, because of a lack of out-of-class learning support, many students fail to comprehensively prepare the provided lecture materials before class. One promising solution to this problem is recommendation systems in the educational area, which have been instrumental in helping learners identify useful and relevant lecture materials that satisfy their learning needs. Thus, in this study, we propose a summarized lecture material recommendation system, which is integrated into an e-book reading system as an enhancement of the flipped classroom approach. This system helps students identify pages that contain essential knowledge that must be thoroughly studied before class. The proposed system was constructed on the basis of our previous work. In this study, a quasi-experiment was conducted in a graduate course that implemented the flipped classroom model: experimental group students learned with the proposed system, whereas the control group students had no access to the additional features. The findings of this study suggest that students who learn with the proposed recommendation system significantly outperform those who learn without the system in a flipped classroom in terms of their learning outcomes and engagement in preclass preparation.

Keywords: Recommendation systems, Flipped classrooms, Preclass preparation, Learning analytics, Learning outcomes

1. Introduction

Researchers have revealed the positive effects of flipped classrooms on students' motivation and performance compared with conventional lecture-based classrooms (Lai & Hwang, 2016). As defined by Bishop and Verleger (2013), the flipped classroom is a technology-supported pedagogical approach that combines asynchronous computer-assisted instruction outside the classroom and interactive in-class learning activities. Flipped classrooms are considered an effective model for engaging students in active learning and involving meaningful teaching and learning strategies during the in-class and out-of-class learning process (Forsey, Low, & Glance, 2013).

However, few studies have investigated how to improve students' out-of-class learning instruction in the flipped classroom model (Jensen et al., 2018). For flipped classroom approaches to be successful, educators must have confidence that their students learned the necessary information and skills from the lecture materials before class. Without appropriate guidance or support, most students tend to present a low level of learning and responsibility during the learning process in a flipped classroom (McLaughlin et al., 2013; Sun, Wu, & Lee, 2017). This lack of clarity regarding optimal class preparation instructions and methods may explain why educators are hesitant to adopt flipped teaching (Jensen et al., 2018). In addition, students' learning achievements in the course have also been reported to strongly correlate with their learning engagement (Yang et al., 2020; Kahu, 2013) and frequently used to examine the effects of their learning approaches on students' learning in flipped classrooms (e.g., Jovanović et al., 2017; Lai & Hwang, 2016). These reports emphasize the importance of considering students' behavioral engagements as measurements when evaluating the effects of educational systems on students' learning in flipped classrooms.

Shimada et al. (2017) revealed that learners are not always willing to prepare for classes, particularly when the volume of the lecture material is too large for them to adequately prepare before class. In their work, it is reported that students have a marked preference for studying summarized lecture material instead of the original material or additional content from external learning environments when aiming for thorough preparation before class. Accordingly, they recommended that students be provided with summarized lecture material instead of requiring them to study the full content; such a measure can increase students' preclass browsing time and overall learning performance.

Learning has shifted from the conventional classroom to the e-learning environment, which uses electronic media, principally through the Internet (George & Lal, 2019). With increasing research attention on recommendation approaches in e-book-based learning resource as well as the rapidly growing amount of online lecture materials, recommendation systems (RSs) in technology-enhanced learning (TEL) have become valuable in supporting learners in identifying useful and relevant lecture materials that satisfy their learning needs (Tarus, Niu, & Mustafa, 2018). The aim of developing an educational RS is to reduce information overload by retrieving the most relevant information and services from a large amount of data (Lu et al., 2015). Thus, current intelligent agents and RSs have been widely studied and accepted as solutions for overcoming information retrieval challenges by learners because of information overload (Montaner, López, & De La Rosa, 2003). These systems could therefore address some of the concerns in flipped classroom models, such as students not preparing before class (Louhab et al., 2020).

However, most existing and proposed educational RSs focus on recommending lecture materials from an external learning environment (Bauman & Tuzhilin, 2018; Ghauth & Abdullah, 2010). Thus, comprehensively satisfying students' learning needs, particularly for students' preclass preparation in flipped classrooms, remains challenging. Therefore, investigations regarding recommendations within the lecture material itself is quite rare, and further clarification is needed.

To address the aforementioned concerns regarding flipped classrooms and improve students' preclass preparation, in this study, based on our previous work (Yang et al., 2019), we propose the summarized lecture material recommendation system that utilizes several text processing and image processing techniques to generate recommendations within the lecture material itself. The proposed system is considered an enhancement of the flipped classroom model as it aims to retrieve the most important information from an expected large amount of contents from the online learning materials and thereby providing suitable guidance to support students' preclass preparation in flipped classrooms.

In contrast to previous works on flipped classroom approaches and educational RSs, internal content (e-book pages) is the target in this proposed system. Instead of the entire digital lecture material, summarized lecture material is generated by collecting the recommended pages from the original lecture material. Moreover, a supervised machine learning approach is applied that extracts more features from e-book text and image content compared with the approach presented in Shimada et al. (2017). To elucidate the effects of the proposed system on students' learning outcomes and engagement in preclass preparation, we investigated the following research questions:

RQ1: Do students who learn with the proposed system achieve better learning outcomes in comparison with students who learn without the system in a flipped classroom?
RQ2: Do students who learn with the proposed system exhibit greater engagement in preclass preparation in comparison with students who learn without the system in a flipped classroom?

2. Literature review

2.1. Flipped classrooms

Researchers have continually investigated how to maximize students' academic performance in higher education. The flipped classroom is gaining popularity as a means to support student learning in such contexts (Boevé et al., 2017). The organization of flipped classroom lectures requires students to prepare for the in-class meetings, often facilitated through additional online video lectures, and demands involvement during lectures by means of problem-solving and peer instruction (Abeysekera & Dawson, 2015).

The flipped classroom uses a student-centered approach, focusing on student learning and placing the responsibility for learning more on students than on teachers while encouraging them to experiment (Sams, 2011). However, researchers have argued that without appropriate guidance or support, most students tend to exhibit a low level of learning and responsibility during the learning process in a flipped classroom (McLaughlin et al., 2013; Sun, Wu, & Lee, 2017). Thus, to adequately support students' preclass preparation in flipped classroom models, Diwanji, Hinkelmann, and Witschel (2018) proposed an application based on artificial intelligence (AI) and data analysis techniques to intrinsically motivate the students' preclass preparation in a flipped classroom. Louhab et al. (2020) proposed the smart adaptive management for flipped learning plugin that can be combined with a Learning Management System to provide students with adapted lecture materials according to their knowledge level and skills when learning in a flipped classroom. Their findings indicated that the students who used the plugin obtained more favorable final examination results than students who did not. This also exhibits the positive effects on students' knowledge levels when integrating the advanced educational technologies into flipped classroom settings.

The aforementioned studies have demonstrated the considerable advantages of implementing flipped classrooms as a new teaching strategy. However, few studies have focused on applying educational RSs that can provide suitable guidance to support students' preclass preparation in flipped classrooms. Hence, investigating the effects of educational RSs on students' learning outcomes and their engagement in preclass preparation in a flipped classroom (i.e., reading time) is paramount.

2.2. Educational recommendation systems

Using a different educational setting and novel methods, Bousbahi and Chorfi (2015) proposed an architecture for personalizing a massive open online course by using a case-based reasoning RS. Benhamdi, Babouri, and Chiky (2017) developed a new multi-personalized recommender system that considers a student's level of study, preference, and memory capacity when providing recommendations. Klašnja-Milićević et al. (2011) proposed an online tutoring system, Protus, in which learners are clustered on the basis of their learning styles. Furthermore, the learners' sequential patterns were mined. Finally, recommendations were generated by applying the collaborative filtering approach.

AI planning and case-based planning can also be applied to achieve a more personalized e-learning solution for students (Garrido, Morales, & Serina, 2016). Obeid et al. (2018) generated personalized recommendations by creating ontologies for students, employees, and higher educational institutions. Machine learning algorithms were then applied for clustering similar graduate students. Hussain et al. (2019) developed a system using machine learning algorithms such as support vector machine, logistic regression, naive Bayes classifier, and artificial neural networks. The objective was to predict students' difficulty when undertaking courses online and provide feedback to teachers on those who are using the system.

Using the features of educational RSs, the aforementioned systems were helpful in supporting students' preclass preparation because they could retrieve essential information from the overwhelming amount of online resources in different educational settings. However, most existing educational RSs were not designed to provide suitable guidance to support students' preclass preparation in flipped classrooms. To achieve such a goal, the proposed system takes into the advantages of AI techniques to generate recommendations within the online lecture material itself. Specifically, the proposed system predicts e-book pages that contain essential knowledge from online lecture materials as the learning guidance for students to perform a thorough preclass preparation in flipped classrooms.

3. A summarized lecture material recommendation system

3.1. System overview

The proposed system aims to rank e-book pages according to their corresponding importance using a machine learning approach. The proposed system selects the e-book pages that are top-ranked for containing essential information that must be learned in advance to form the summarized lecture material. Figure 1 illustrates the procedures of the proposed system. First, we collected different types of data from the lecture material uploaded to an e-book system as the system's input data. Second, we extracted text features and image features from the collected input data. Third, we preprocessed the extracted features. Fourth, we trained a machine learning model to retrieve class probabilities from the training process and rank the input e-book pages accordingly. Last, after selecting several top-ranked e-book pages in a preference range from the ranking, the summarized lecture material recommendation is generated as the output.

Because of the characteristics of machine learning-based systems, the proposed system becomes more predictive as it receives new materials from the same instructor and the same course. Notably, the proposed system only uses new lecture materials from the same instructor and course for model training. This is inspired by the notion that learning and educational technologies should be adopted and applied in course-specific contexts rather than relying on a generalized model that applies one trained prediction model in all the educational contexts (Gašević et al., 2016).

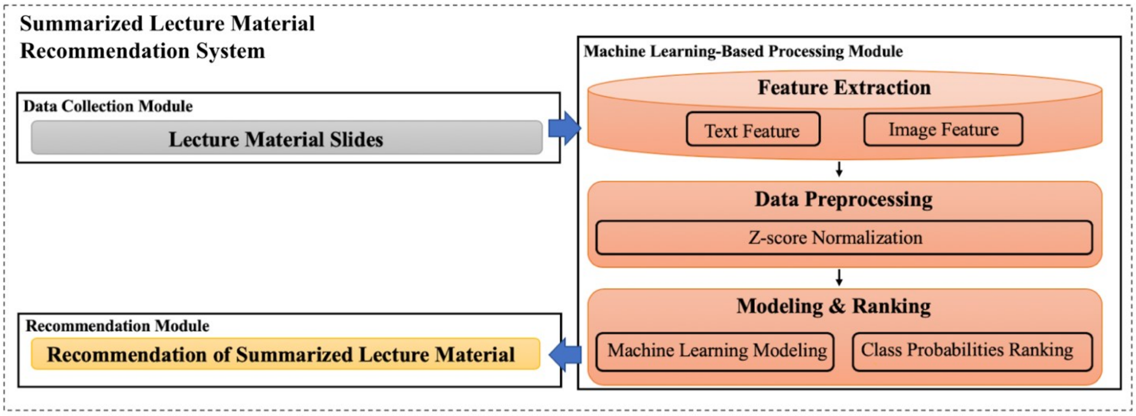

Figure 1. Procedures of the proposed system

3.2. Data collection, feature extraction, and data preprocessing

In this study, we used the BookRoll system (Ogata et al., 2015), which is an e-book reading system that allows users to browse the uploaded content anytime and anywhere. Figure 2 illustrates an example of BookRoll's user interface. Several functions such as page-turning, highlighting, bookmark making, note-taking, page jumping, and internal content searching are available. Students' reading behaviors while using BookRoll are stored in its database. During the data collection process, by using BookRoll, text content, and image content from the original digital lecture material were collected for further feature extraction and data preprocessing. Note that the types of BookRoll learning behavioral logs were described by the developers in their 2015 study (Ogata et al., 2015); therefore, they are not detailed in the present study.

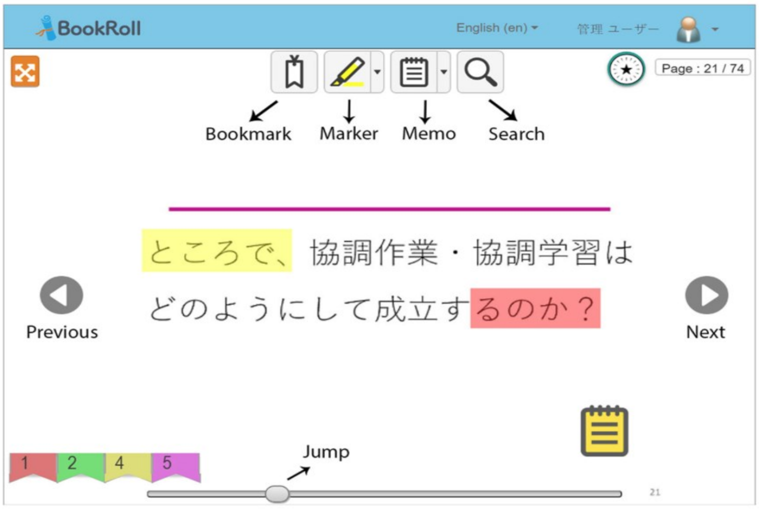

Figure 2. Screenshot of the BookRoll system's user interface

The characteristics of the essential document or e-book pages and the associated content features during the feature extraction process have been mentioned and applied in several studies (Neto, Freitas, & Kaestner, 2002; Shimada et al., 2017). Therefore, we extracted two categories of features from the collected data on the basis of the characteristics of essential pages, as described in Table 1, to train our summarized lecture material generator. In the present study, an essential document page or e-book page should have the following characteristics:

- Sufficient content to be worth browsing
- Unique visual content
- Keywords that appear frequently on the page

- Keywords that rarely appear throughout the whole lecture material
- Similarity to the title of the lecture material
- Similarity to other pages under the same lecture title

We applied text processing techniques including N-gram tokenizer, cosine similarity, and lemmatization to extract text features from the collected e-book pages. We used cosine similarity measurement to calculate text similarity between paired e-book pages and similarity to the title and keywords of the lecture, which represents the values for Page–Page cohesion, similarity to the title, and similarity to keywords, respectively. We counted the total number of characters and punctuations on a page from the generated corpus to represent the values of TotalChar and Punctuation. We applied the vector space modeling technique term frequency–inverse document frequency (TFIDF) weighting to calculate the weight of each term on the page and calculated the average TFIDF value for each page, representing the value of AvgTFIDF.

For image feature extraction, we applied image processing algorithms including the background subtraction method and the interframe difference method, which have both been used for the extraction of visual information from the lecture materials, to extract values representing Background subtraction and Background subtraction + Inter-frame difference. The details of the extracted image features were described in a previous article (Shimada et al., 2017).

Next, to obtain an improved distribution of the dataset that contains all the feature values of each e-book page, we preprocessed each extracted feature by using the z-score normalization method to redefine and transform the range of feature values into a smaller and specific range.

Table 1. Description of the extracted features

Feature	Description of the feature
TotalChar	Total characters on the page
AvgTFIDF	Average TFIDF after preprocessing by the bigram tokenizer
Similarity to title	Cosine similarity to the title of the content
Similarity to keywords	Cosine similarity to the keywords of the content
Page–Page cohesion	Sum of cosine similarities to the remaining pages
Punctuation	Total instances of punctuations on the page
Background subtraction	Total amount of foreground pixels on the page
Background subtraction + Inter-frame difference	Absolute foreground pixel differences in comparison with the previous page and the next page (higher value is selected)

3.3. Model generation and recommendation of summarized lecture material

In this study, our problem had a binary classification; however, we considered class probabilities as an output instead of class labels. This provides the flexibility to rank e-book pages on the basis of their possibilities to be positively classified. Thus, the instructor can determine which volume of the summarized lecture material becomes the recommendation for preclass preparation. In the proposed system, we requested that instructors label "recommended pages" in lecture material to represent the gold standard. These labeled pages are recommended reading before class because they contain essential information. After receiving the labeled lecture material, we trained a supervised machine learning model to generate the class probabilities of each page. Next, these e-book pages were ranked by class probabilities, and the top-ranked pages were then selected to generate the summarized lecture material recommendation.

3.4. Preliminary evaluation for selecting the best-performing model and constructing the proposed system

To evaluate the predictive performance of the machine learning models, we conducted a preliminary evaluation using the extracted text and image features of lecture material containing 91 pages. The instructor labeled 45 out of 91 pages as recommended reading for preclass preparation. The remaining 46 pages were then considered "not recommended" pages, indicating optional reading before class. In the preliminary evaluation, we trained and compared eight well-known machine learning classification models from the Python (Pedregosa et al., 2011) libraries that involve different algorithms for class probability prediction during the model training process. We used all the extracted text features and image features for the model training, and we ranked the input e-book pages by the predicted class probabilities accordingly. After the ranking process, we selected the top-ranked e-book pages to form the recommendation for learners' preparation before class. All the prediction models were trained with default parameters.

The models' performance was evaluated using precision, recall, and area under the curve (AUC) of receiver operating characteristic (ROC) metrics. The precision and recall metrics were calculated according to the confusion matrix, and the AUC was calculated from the ROC curve. The AUC value ranges between 0 and 1, and an AUC value near 0.5 indicates that the classification process is similar to a random guess. The values of precision and recall are always the same in each model because the number of pages that the proposed system predicted as the recommended pages are always equivalent to the number the instructor labeled as recommended pages. Hence, the performance was validated using threefold cross-validation, meaning that our dataset was randomly divided into three disjoint subsets of equal size in a stratified manner (maintaining the original class distribution) and each model was executed three times. In each repetition, one of the three subsets was used as the testing set, and the other two were combined to form the training set. The preliminary results are presented in Table 2. The results revealed that the multilayer perceptron classification model (indicated in boldface in Table 2) was the optimal model with the greatest potential for use in the development of our proposed system. These results provided the data necessary to implement our experiment, which applied the proposed system on a graduate course that employed the flipped classroom model.

Table 2. Preliminary evaluation results of model performance

Model	Precision	Recall	AUC
DTNB	0.422	0.422	0.429
JRip	0.489	0.489	0.494
Random Forest	0.556	0.556	0.560
J48	0.533	0.533	0.538
BayesNet	0.422	0.422	0.429
Gaussian Naive Bayes	0.467	0.467	0.472
Logistic Regression	0.622	0.622	0.626
Multilayer Perceptron	**0.667**	**0.667**	**0.670**

3.5. Validation of the system's performance in comparison with human decisions

After the preliminary evaluation for finding the ideal model, we asked instructors to compare the recommendation provided by the system to the material they identified as essential reading before class. Tests were conducted on the six different lecture materials that were provided to students for preclass preparation in flipped classroom settings during the experiment. The results of the predictive performance of the system, as measured by precision, recall, and AUC, are shown in Table 3. The volume of the lecture materials (in pages) is also presented. The proposed system exhibited decent predictive performance, with scores higher than 60% for precision, recall, and AUC for each week.

Table 3. Predictive performance of the proposed system in comparison with human decisions

Lecture Material	Precision	Recall	AUC	Volume (# of Page)
Lecture material 1	0.742	0.737	0.739	30
Lecture material 2	0.744	0.741	0.742	27
Lecture material 3	0.816	0.806	0.812	27
Lecture material 4	0.733	0.733	0.733	31
Lecture material 5	0.647	0.645	0.646	30
Lecture material 6	0.62	0.613	0.616	32

4. Experiment

4.1. Participants and study contexts

A quasi-experiment was conducted in a flipped classroom graduate course, namely creative learning, in a university in northern Taiwan. A total of 30 graduate students from the department of computer science participated in the experiment. Students enrolled in the course were required to use BookRoll to study the given scientific articles during out-of-class learning time. During each class, students presented and discussed the content of the scientific articles with their classmates in addition to receiving some instruction from their teacher. Among the participants, 15 students were assigned as the experimental group and the remaining 15 students were assigned as the control group. Students in the experimental group were allowed to use BookRoll freely for their in-class and out-of-class learning with additional requirements to the uses of summarized lecture material recommendations in addition to the original lecture material for preclass preparation. By contrast, students in the control group used BookRoll for their learning but without receiving any of the recommendations from our proposed system throughout the experiment.

4.2. Measurements

A pretest and posttest were issued by the course instructor. The pretest aimed to evaluate the students' prior knowledge of the course. It consisted of 10 open-ended questions pertaining to data science and statistical analysis (e.g., what is text mining, what is t-test, how do you define low-skill readers), with a perfect score of 100. The posttest aimed to assess the students' learning outcomes in the course. It also consisted of 10 open-ended questions pertaining to data science and statistical analysis (e.g., what is the key feature of the text mining method, what does the t-test show and how to interpret the results, what is the main reason why low-skill readers use text-marking strategies), with a perfect score of 100. In addition, two experts from the department of computer science at the participating university were invited to ensure that the pretest and posttest were sufficient to evaluate the students' prior knowledge and learning outcomes in the course. Pearson's correlation between the two tests was 0.026, indicating a low correlation between the two tests. Moreover, the Kuder–Richardson Formula 20 of the posttest was 0.59, indicating acceptable internal consistency (Cortina, 1993).

The following three indicators were monitored throughout the experiment to measure students' learning behaviors (redesigned as "engagement" or "engagements" in the following sections) in preclass preparation when learning in the flipped classroom. It is noted that considering the cases that participants just open the lecture materials and quickly surf through the contents or do not pay attention to the contents, we pruned the outlier BookRoll activities and sessions. In particular, activities that were taken with too much (more than 10 minutes) or too few (less than three seconds) time were eliminated from the analysis of learning behaviors in the following sections. Moreover, the outlier sessions, in particular, overly long sessions (include more than 50 occurrences of activities) and overly short sessions (include less than three occurrences of activities) were eliminated from the analysis as well:

- Time (hour): Time spent in preclass preparation. The students' preclass time spent on learning materials has been indicated as a major factor that needs to be taken into account when utilizing flipped classroom models (Persky & Hogg, 2017). With its potential in indicating the students' learning engagements in flipped classrooms, in this experiment, students' preclass time spent on BookRoll is measured by summing up the reading time spent on each reading activity during preclass preparation. As shown in Figure 3, the timestamp of each reading activity can be recorded by the BookRoll system. For example, the subtracted value of the two timestamps of reading activities A and B indicates the time spent on reading activity A, etc.
- Activity: Number of BookRoll activities completed during preclass preparation. The students' preclass activities have been evident with the relevance to their learning outcomes in the course when learning in a flipped classroom where students with more preclass activities in a flipped classroom obtained higher course performance (Jovanović et al., 2017). With its potential in indicating the students' learning engagements in flipped classrooms, in this experiment, students' preclass reading activity on BookRoll is measured by summing up the occurrence of reading activities (as shown in Figure 3). For example, a total of four rows of BookRoll activities shown in Figure 3 represents the total of four occurrences of preclass reading activity. Among them, only three activities will be taken into account for the analysis of BookRoll learning behaviors as the time taken for the third activity is more than 10 minutes.
- Session: Number of reading sessions during preclass preparation. The students' preclass learning sessions have been reported to positively correlate with their learning outcomes in the course when learning in a flipped classroom where students with more preclass learning sessions obtained higher course performance (Jovanović et al., 2017). With its potential in indicating the students' learning engagements in flipped classrooms, in this experiment, students' preclass reading session on BookRoll is measured by summing up the occurrence of the reading activity OPEN as shown in Figure 3. For example, a total of one row of BookRoll activity OPEN, shown in Figure 3, represents the total of one preclass reading session. In this case, after pruning the outlier activities, the occurrence of activities in this session is three; therefore, this reading session will be taken into account for the analysis of BookRoll learning behaviors. On the other hand, if the occurrence of BookRoll activities in the session is more than 50 or less than three, this session will not be taken into account for the analysis.

An example of the learning log collected in BookRoll is illustrated in Figure 3.

User ID	Content ID	Operation Name	Operation Date
15910	ed645f3821e	OPEN	2019-12-20 10:03:52
15910	ed645f3821e	NEXT	2019-12-20 10:05:17
15910	ed645f3821e	ADD MEMO	2019-12-20 10:07:03
15910	ed645f3821e	NEXT	2019-12-20 11:27:14

Figure 3. Example of the BookRoll learning behavioral log collected in this study

4.3. Experimental procedures

Figure 4 illustrates the flow chart of the experiment. The duration of the experiment was eight weeks. The instructor released the original version of the lecture material to the BookRoll system one week before the class to enable students to prepare at their own pace and when and where they chose. In the first week of the experiment, the course syllabus and usage of the BookRoll system were introduced to all the students who were participating in the experiment. Moreover, the experimental (n = 15) and control groups (n = 15) completed the pretest examination to evaluate their prior knowledge of the course. Between the second and the seventh week of the experiment, the two groups prepared for class in different manners. Students in the experimental group were allowed to use BookRoll freely for their preclass preparation. Moreover, they were explicitly required to use the proposed system to view the summarized lecture material recommendations in addition to the original lecture material for preclass preparation. By contrast, students in the control group prepared without receiving any of the recommendations from our proposed system. Notably, both groups learned in a flipped classroom using BookRoll; the only difference between the two groups was the use of the proposed system to support preclass preparation. The length of the summarized lecture material that was provided to students in the experimental group depended on the instructor (in case of this study, 50% of the lecture materials were identified as essential reading before the class every week). Figure 5 illustrates the different BookRoll user interfaces for the control and experimental groups. For the experimental group, recommendations for summarized lecture materials were provided on the lower left side of the page (bookmarks), and students could browse the content by clicking the bookmarks. In the last week of the experiment, a posttest was conducted to evaluate both groups' learning outcomes in the course.

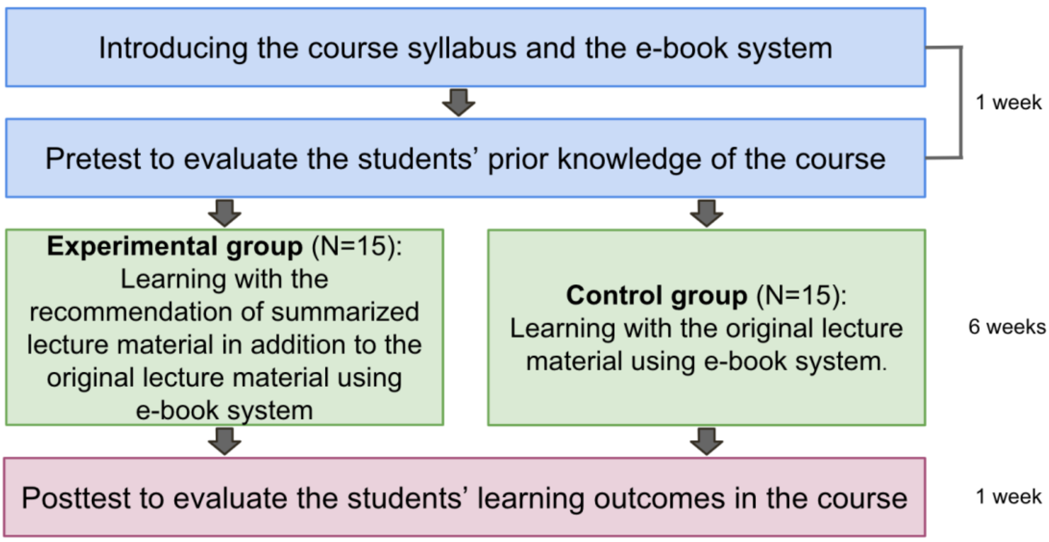

Figure 4. Experimental procedures

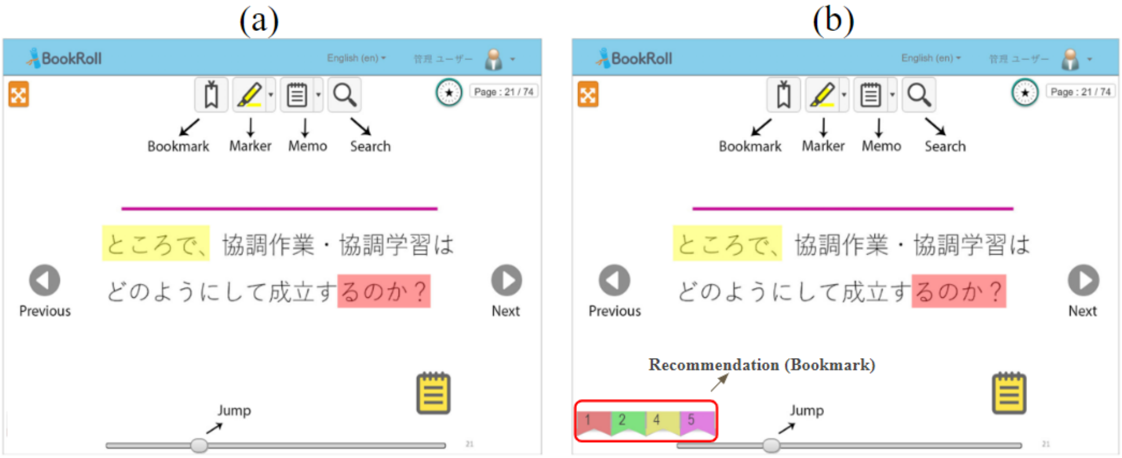

Figure 5. Different BookRoll user interfaces for (a) control group students that contained no recommendations and (b) experimental group students that contained recommendations on the lower left side using BookRoll's bookmark feature

5. Results

5.1. Analysis of learning outcomes

To investigate the effects of the proposed system on the learning outcomes of the students, the IBM SPSS one-way analysis of covariance (ANCOVA) was employed, with the students' pretest scores serving as the covariate. Levene's test of determining homogeneity of variance was not violated (F = 0.68, p > .05), suggesting that a common regression coefficient was appropriate for the implementation of the one-way ANCOVA. Moreover, the Shapiro–Wilk test was employed to examine the normality of the data. The results (control group: 0.93 [p > .05]; experimental group 0.95 [p > .05]) indicated that both groups' data were normally distributed in terms of posttest scores.

Table 4 presents the one-way ANCOVA results of the two groups' posttest scores. A significant difference was observed between the two groups of students in terms of their posttest scores (F = 36.58, p < .001, η^2 = 0.58), with an adjusted mean and adjusted standard deviation of 85.83 and 2.69, respectively, for the experimental group and 62.2 and 2.69, respectively, for the control group. The posttest scores of the experimental group were significantly higher than those of the control group. The effect size (η^2) was 0.58, indicating a medium to large effect size (Cohen, 1988). Overall, the results demonstrated that the proposed system exerted positive impacts on students' learning outcomes in flipped classrooms.

Table 4. One-way ANCOVA results of the posttest scores of the two groups

Group	N	Mean	SD	Adjusted mean	Adjusted SD	F	η^2
Experimental group	15	86.07	10.26	85.83	2.69	36.58***	0.58
Control group	15	62.2	9.89	62.44	2.69		

Note. $^{***}p$ < .001.

5.2. Analysis of engagement in preclass preparation

Next, the effects of the proposed system on the two groups' engagement in preclass preparation were investigated. The Mann–Whitney U test was employed because the values of the three indicators did not meet the assumptions of data normality and homogeneity of variance, which are required for a parametric test.

Table 5 reveals the Mann–Whitney U test results of the three indicators of the students' engagement in preclass preparation. First, a significant difference was observed in terms of time spent on preclass reading by the median, 25th percentile, and 75th percentile of the experimental and control groups: 18.81 versus 4.12, 7.58 versus 2.5, and 30.65 versus 5.21, respectively (U = 25, p < .001). The preclass reading time of the experimental group students was significantly higher than that of the control group students. This indicates that students who used the proposed system had a greater engagement in terms of the time spent on preclass reading than did students who learned without the proposed system.

Table 5. Mann–Whitney U test results of the indicators of engagement in preclass preparation between two groups

Indicator	Group	*N*	Median	25th / 75 percentiles	*U*
Time (hour)	Experimental group	15	18.81	7.58 / 30.65	25^{***}
	Control group	15	4.12	2.5 / 5.21	
Activity	Experimental group	15	651	343 / 945	49^{**}
	Control group	15	209	157 / 325	
Session	Experimental group	15	35	30 / 66	59^{*}
	Control group	15	23	17 / 31	

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

A significant difference was observed in terms of the number of preclass learning activities engaged in by the median, 25th percentile, and 75th percentile of the experimental and control groups: 651 versus 209, 343 versus 157, and 945 versus 325, respectively ($U = 49$, $p < .01$). The number of preclass activities completed by students in the experimental group was significantly higher than that of students in the control group. This indicates that students who used the proposed system had greater engagement in terms of the number of preclass activities than did students who learned without the proposed system.

Moreover, a significant difference was evident in the number of preclass learning sessions engaged in by the median, 25th percentile, and 75th percentile of the experimental and control groups: 35 versus 23, 30 versus 17, and 66 versus 31, respectively ($U = 59$, $p < .05$). The number of preclass reading sessions of the experimental group was significantly higher than that of the control group. This indicates that students who used the proposed system had greater engagement in terms of the number of preclass reading sessions than did students who learned without the proposed system.

6. Discussion

6.1. Educational impacts of the proposed system on students' learning outcomes in flipped classrooms

The success of implementing flipped classroom approaches has been positively correlated with learners' learning achievements in the course based on the findings reported in Jovanović et al. (2017) and Lai and Hwang (2016). The experimental results in this study have revealed that the students who used the proposed system obtained significantly higher learning outcomes compared to the students who did not. These findings provide evidence that students who learn with the proposed system can achieve better learning outcomes than students who learn without it when learning with an e-book system in flipped classrooms. Specifically, it is implied that the students' learning outcomes in the course can be improved by helping them identify pages that contain essential knowledge that must be thoroughly studied before the class when learning with an e-book system in flipped classrooms. The results were consistent with previous findings reported in Shimada et al. (2017), where the students who learned with the summarized lecture material recommendations before class obtained more favorable learning outcomes than did students who did not have access to the recommendations. The results were also consistent with the findings of Louhab et al. (2020), where students who received guidance for preclass preparation through an RS exhibited higher knowledge levels than did students who did not use such a system. From the perspective of educational RSs, the results were consistent with those of previous studies (Bauman & Tuzhilin, 2018; Ghauth & Abdullah, 2010; Hsu, Hwang, & Chang, 2013), which have provided evidence that students who use an educational RS to support their learning tend to exhibit stronger academic performance than students who do not.

6.2. Educational impacts of the proposed system on students' preclass engagements in flipped classrooms

The importance of measuring students' behavior changes and engagements when evaluating the effects of newly proposed systems or methods on students' learning in flipped classroom settings has been emphasized based on the findings reported in Kahu (2013), Jovanović et al. (2017), and Lai and Hwang (2016). The experimental results in this study have revealed that the students who used the proposed system exhibited significantly greater engagement in preclass preparation compared to the students who did not. These findings provide evidence that students who learn with the proposed system exhibit greater engagement in preclass preparation than students who learn without it when learning with an e-book system in flipped classrooms. This also implies that by helping the students identify pages that contain essential knowledge that must be thoroughly studied before the class, the students tend to exhibit greater engagements in preclass preparation when learning with an e-book system in flipped classrooms. Notably, the current results are inconsistent with those of Shimada et al. (2017), where the students who learned with the summarized lecture material recommendations actually spent less time preparing for the lecture before class than did students who did not have access to the system. One reason for the disparity in the results could be the objective and design of the proposed system. In this study, the proposed system aimed to provide suitable guidance by recommending summarized lecture material in addition to the original lecture material to increase the students' engagements in preclass preparation, whereas Shimada et al. (2017) aimed to reduce the volume of the lecture material with the ultimate goal of minimizing the students' browsing time during preclass preparation; consequently, students were only allowed to browse the summarized lecture material. The divergent results of these two studies emphasize the importance of clarifying the objective of such systems (increasing or decreasing students' browsing time before class). The results obtained from the comparison results between the experimental group and the control group in terms of preclass reading time, number of learning activities, and number of learning sessions provide evidence that students who learned with the proposed system exhibited greater engagement in preclass preparation than did students who learned without it in a flipped classroom. The proposed system is expected to provide an extra option for instructors to automatically generate summarized lecture materials as the learning guidance for the learners' preclass preparation in flipped classrooms without expending too much time and effort. This also exhibits potential in enabling teachers to support students in increasing their behavioral engagements in preclass preparation in a more efficient manner.

6.3. Limitations and suggestions for future studies

This study had some limitations. Although the students' learning outcomes were normally distributed, only 30 students participated in the experiment. Therefore, future studies should employ a larger sample to obtain stronger results with greater external validity. Moreover, we investigated the effects of the proposed system using eight features from both text and image content in the e-book lecture material for the machine learning model training task. However, we did not establish whether other types of educational data can be used for similar recommendation tasks. Consequently, future studies could include other features from different educational data sources. Furthermore, although the proposed system was evident with its effects on students' learning outcomes in the course as well as their engagements in preclass preparation according to the experimental results, it is difficult to assert that the proposed system will not limit the students' strategies in preclass preparation when learning with flipped classroom settings. Therefore, future studies are suggested to investigate such an issue. Lastly, since we focused on the uses of the proposed system but did not include the investigation of the influence of participants' learning style or personality traits on their learning outcomes and engagement in preclass preparation, the analysis results might contain some bias. Consequently, it would be needed to extend our investigation with multi-modal approaches where participants' learning style or personality traits are combined with the analysis of data obtained from other sources such as self-reports or interviews with instructors, to better interpret the findings of this study, particularly in terms of the participants' engagements of learning.

7. Conclusions

Online learning continues to grow, in part, because of the reduced cost, increased flexibility regarding class schedule, and improved mobility when taking a class (Allen & Seaman, 2014). With a dramatically increasing volume of digital lecture material, methods to support students when they study the materials are more crucial than ever. These methods can streamline students' preclass preparation and thereby maximize their engagement in a flipped classroom. To support students in achieving better learning outcomes and exhibiting greater engagement in preclass preparation in flipped classrooms, on the basis of our previous work, a summarized lecture material recommendation system was proposed. The proposed system was integrated into an e-book system and implemented in a flipped classroom model, constituting a different learning approach than that employed in typical flipped classrooms.

In this study, a quasi-experiment was conducted in a graduate course in northern Taiwan that implemented the flipped classroom model to evaluate the effects of the proposed system on students' learning. The experimental group learned with the proposed system, whereas the control group had no access to the additional features. The findings of this study suggest that the proposed system can enhance the flipped classroom and mitigate, to some extent, some concerns with the flipped classroom approach, such as difficulty providing appropriate guidance to students for preclass preparation. The experimental students outperformed the control students in terms of their learning outcomes in the course and engagement in preclass preparation including reading time, number of learning activities, and number of learning sessions at statistically significant levels. This implies that students who learn with the proposed system can not only achieve better learning outcomes but also exhibit greater engagement in preclass preparation than can students who learn without the proposed system in a flipped classroom. Furthermore, the proposed system can be regarded as an effective option for instructors to automatically generate summarized lecture materials in addition to the original lecture materials without expending too much time and effort. This enables teachers to support students in enhanced flipped classrooms by increasing the students' engagements in preclass preparation and thereby improving their learning outcomes in the adopted flipped classroom approach.

As a follow-up to the current study, we are now developing a knowledge-based RS that can be adapted and personalized by considering the students' learning behavior, assessment data, and the knowledge structure of the lecture materials. We aim to develop and propose in the future a bigger framework for educational RSs for students' preclass preparation and postclass reflection by integrating the two RSs into the course.

Acknowledgement

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (B)20H01722, JSPS Grant-in-Aid for Scientific Research (S)16H06304 and NEDO Special Innovation Program on AI and Big Data 18102059-0.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development*, *34*(1), 1-14.

Allen, I. E., & Seaman, J. (2014). *Grade change: Tracking online education in the United States*. Babson Park, MA: Babson Survey Research Group and Quahog Research Group.

Bauman, K., & Tuzhilin, A. (2018). Recommending remedial learning materials to students by filling their knowledge gaps. *MIS Quarterly*, *42*(1), 313-332.

Benhamdi, S., Babouri, A., & Chiky, R. (2017). Personalized recommender system for e-Learning environment. *Education and Information Technologies*, *22*(4), 1455-1477.

Bishop, J. L., & Verleger, M. A. (2013). The Flipped classroom: A Survey of the research. In *ASEE national conference proceedings, Atlanta, GA* (Vol. 30, No. 9, pp. 1-18).

Boevé, A. J., Meijer, R. R., Bosker, R. J., Vugteveen, J., Hoekstra, R., & Albers, C. J. (2017). Implementing the flipped classroom: An Exploration of study behaviour and student performance. *Higher Education*, *74*(6), 1015-1032.

Bousbahi, F., & Chorfi, H. (2015). MOOC-Rec: A Case based recommender system for MOOCs. *Procedia-Social and Behavioral Sciences*, *195*, 1813-1822.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cortina, J. M. (1993). What is coefficient alpha? An Examination of theory and applications. *Journal of applied psychology*, *78*(1), 98-104. doi:10.1037/0021-9010.78.1.98

Diwanji, P., Hinkelmann, K., & Witschel, H. F. (2018). Enhance classroom preparation for flipped classroom using AI and analytics. In *Proceedings of the 20th International Conference on Enterprise Information Systems (ICEIS 2018)* (pp. 477-483). doi:10.5220/0006807604770483

Forsey, M., Low, M., & Glance, D. (2013). Flipping the sociology classroom: Towards a practice of online pedagogy. *Journal of Sociology*, *49*(4), 471-485.

Garrido, A., Morales, L., & Serina, I. (2016). On the use of case-based planning for e-learning personalization. *Expert Systems with Applications*, *60*, 1-15.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The Effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68-84.

George, G., & Lal, A. M. (2019). Review of ontology-based recommender systems in e-learning. *Computers & Education*, *142*, 103642. doi:10.1016/j.compedu.2019.103642

Ghauth, K. I., & Abdullah, N. A. (2010). Learning materials recommendation using good learners' ratings and content-based filtering. *Educational technology research and development*, *58*(6), 711-727.

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A Personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, *63*, 327-336.

Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, *52*(1), 381-407.

Jensen, J. L., Holt, E. A., Sowards, J. B., Ogden, T. H., & West, R. E. (2018). Investigating strategies for pre-class content learning in a flipped classroom. *Journal of Science Education and Technology*, *27*(6), 523-535.

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*(4), 74-85.

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, *38*(5), 758-773.

Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers & Education*, *56*(3), 885-899.

Lai, C. L., & Hwang, G. J. (2016). A Self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, *100*, 126-140.

Louhab, F. E., Bahnasse, A., Bensalah, F., Khiat, A., Khiat, Y., & Talea, M. (2020). Novel approach for adaptive flipped classroom based on learning management system. *Education and Information Technologies, 25*(2), 755-773. doi:10.1007/s10639-019-09994-0

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A Survey. *Decision Support Systems*, *74*, 12-32.

McLaughlin, J. E., Griffin, L. M., Esserman, D. A., Davidson, C. A., Glatt, D. M., Roth, M. T., Gharkholonarehe, N., & Mumper, R. J. (2013). Pharmacy student engagement, performance, and perception in a flipped satellite classroom. *American Journal of Pharmaceutical Education*, *77*(9). doi:10.5688/ajpe779196

Montaner, M., López, B., & De La Rosa, J. L. (2003). A Taxonomy of recommender agents on the internet. *Artificial intelligence review*, *19*(4), 285-330.

Neto, J. L., Freitas, A. A., & Kaestner, C. A. (2002). Automatic text summarization using a machine learning approach. In *Brazilian symposium on artificial intelligence* (pp. 205-215). Berlin, Heidelberg: Springer. doi:10.1007/3-540-36127-8_20

Obeid, C., Lahoud, I., El Khoury, H., & Champin, P. A. (2018). Ontology-based recommender system in higher education. In *Companion Proceedings of 2018 Web Conference Companion (WWW 2018)* (*pp.* 1031-1034). Lyon, France: ACM. doi:10.1145/3184558.3191533

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K., & Yamada, M. (2015). E-Book-based learning analytics in university education. In *Proceeding of 23rd International Conference on Computer in Education (ICCE 2015)* (pp. 401-406). Hangzhou, China: APSCE.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vandeplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, *12*, 2825-2830.

Persky, A. M., & Hogg, A. (2017). Influence of reading material characteristics on study time for pre-class quizzes in a flipped classroom. *American journal of pharmaceutical education*, *81*(6), 103. doi:10.5688/ajpe816103

Sams, A. (2011). The Flipped class: Shedding light on the confusion, critique, and hype. *The Daily Rift*, 1.

Shimada, A., Okubo, F., Yin, C., & Ogata, H. (2017). Automatic summarization of lecture slides for enhanced student previewtechnical report and user study. *IEEE Transactions on Learning Technologies*, *11*(2), 165-178.

Sun, J. C. Y., Wu, Y. T., & Lee, W. I. (2017). The Effect of the flipped classroom approach to OpenCourseWare instruction on students' self-regulation. *British Journal of Educational Technology*, *48*(3), 713-729.

Tarus, J. K., Niu, Z., & Mustafa, G. (2018). Knowledge-based recommendation: A Review of ontology-based recommender systems for e-learning. *Artificial intelligence review*, *50*(1), 21-48.

Yang, C. C., Chen, I. Y., Huang, A. Y., Lin, Q. R., & Ogata, H. (2020). Can self-regulated learning intervention improve student reading performance in flipped classrooms? *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *10*(4), 1-13.

Yang, C.C., Akçapinar, G., Flanagan, B, & Ogata, H. (2019). Developing e-book page ranking model for pre-class reading recommendation. In *Proceeding of 27th International Conference on Computer in Education (ICCE 2019)* (pp. 360-362). Kenting, Taiwan: APSCE.