Universitat Oberta de Catalunya

Universidad de los Andes
Colombia

RESEARCH ARTICLE

Open Access

The possibility of predicting learning performance using features of note taking activities and instructions in a blended learning environment

Minoru Nakayama[1]* , Kouichi Mutsuura[2] and Hiroh Yamamoto[3]

*Correspondence:
nakayama@ict.e.titech.ac.jp
[1]Information and Communications Engineering, Tokyo Institute of Technology, Ookayama, Meguro, 152-8552, Tokyo, Japan
Full list of author information is available at the end of the article

Abstract

A formative assessment was introduced to a blended learning course in order to predict participant's learning performance using measurements of their note taking activity and metrics of their attitudes. The lexical metrics were created by analyzing the contents of student's notes from every class, and by measuring their characteristics. In the results of two types of regression analysis of these measurements, features of note taking activities were a significant source of information for predicting the scores of final exams. In regards to temporal changes in prediction performance during the course, the progress of learning and the typical changes in features of note taking activity were discussed. During the analyses, the effectiveness of the lecturer's note taking instructions was also confirmed.

Keywords: Note taking, Participant's characteristics, Instruction, Prediction

Introduction

The online learning environment has recently become the preferred method of sharing learning content, which enables participant's learning activities to be analyzed, in order to maximize learning activity and optimize the overall system. For example, typical large scale systems, which are called massive open online courses (MOOCs), are used for delivering learning content (Seaton, Bergner, Chuang, Mitros, & Pritchard, 2014). In order to evaluate both participant's learning activity and learning systems, participant's access logs and learning performance were analyzed (Seaton, Nesterko, Mullaney, Reich, & Ho, 2014). When the current learning environment, which uses information communication technology, was developed, the effectiveness of the system was evaluated using participant's learning behavioral data (Nakayama, Kanazawa, & Yamamoto, 2007; Nakayama, Kanazawa, & Yamamoto, 2009). Using a conventional learning assessment approach, such as a summative evaluation, learning performance during the online course was discussed. Also, the cost benefit of an online learning environment has often been discussed (Bates, 2000). Another approach, known as an authentic assessment (Gulikers, Bastiaens, & Kirschner, 2004), focuses on the student's learning progress. Also, a conventional formative assessment can improve the learning progress of students

and the instructive activities of teachers (Bloom, Hastings, & Madaus, 1971). Since student's learning activity may elicit learning achievement, the formative assessment focuses on activities which improve student's learning progress in comparison with a summative assessment that is based on their final results. In addition to this, the formative assessment can be conducted using regular learning activities such as note taking and scores of various non-test work results. Though participants are evaluated using final test scores during most university courses, a formative assessment may reflect their overall learning performance. Therefore, a formative assessment can provide various kinds of information about participant's learning activity and performance. This approach has been applied to educational improvement of the learning progress (Bell et al., 2015; Bennett, 2015).

Using conventional approaches, the varying effectiveness of student's aptitudes (Cronbach & Snow 1977) and learning behavior, such as note taking activity (Kiewra, 1985, 1989; Kiewra, Benton, Kim, Risch, & Christensen, 1995; Kobayashi, 2005; Piolat, Olive, & Kellogg, 2005) was studied. The relationships between these factors and learning performance have also been widely analyzed and discussed (Cronbach & Snow, 1977; Nye, Crooks, Powley, & Tripp, 1984; Weener, 1974). These survey and analytical techniques were introduced to the study of the online learning environment, and the learning effectiveness of note taking, and some of the causal relationships between learning activities and note taking behavior were analyzed (Nakayama, Mutsuura, & Yamamoto, 2014a, 2015a, 2016b, 2017). Previous studies concerning note taking activity, which examined these overall activities, can be used as a summative assessment as student's achievements. This suggests that metrics of note taking activity can be used as one of the indices of the formative assessment when note contents for every session are assessed.

While some of these measurements affect student's learning, detailed factors and the timing of their effectiveness may be key issues. The analysis of participant's formative leaning process and the revision of metrics in response to student's behavior provide the means to resolve the problem.

This study extracts some of the features of student's attitudes and the contents of notes taken during learning activities, and tracks their contributions to learning performance as the course progress. During online learning courses, behavioral events are organized as learning activities which are supported by information communication technology. Therefore, the research questions this paper is concerned with are: examining the possibility of predicting scores of final exams as a measure of learning performance, using information collected during the learning progress, identifying the learning activity metrics necessary for the prediction, and evaluating the effectiveness of note taking instruction.

The following topics are also addressed in this paper:

- A procedure for the estimation of the final exam scores is proposed, using selected features of the contents of notes taken and participant's characteristics.
- The effectiveness of note taking instructions is evaluated, in order to examine the performance of the estimations.
- Possible time periods for the measurement of the progress of the course are discussed, regarding the performance of the predictions.

Nakayama *et al. International Journal of Educational Technology in Higher Education* (2017) 14:6

Page 3 of 14

Method

Surveys were conducted for two years during a credit course for Bachelor students at a Japanese university. The contents of notes taken by students were collected after every class, and other metrics were surveyed twice during the course. The metrics of the means of the first and second halves were then compared.

Blended learning courses

The subject was Information Networks. The courses were Bachelor level credit courses. The course was taught using the blended learning style, which consists of face-to-face sessions and an online learning environment outside of the classroom. The course consisted of weekly face-to-face sessions for 15 weeks (Nakayama, Mutsuura, & Yamamoto, 2011). Participants could take on-line tests (OT) after each session, using a function of the learning management system (LMS), and the assigned textbook. The function of on-line tests is as a review exercise of the content of each face-to-face session. Since students can repeat the tests until they are satisfied with their scores, most scores are perfect or nearly so. The LMS recorded the final scores of the on-line tests. Therefore, these scores were not included in the evaluation of participant's learning performance. A final exam (FE) was given to all participants, to measure their performance at the end of the course.

Note taking instructions

All participants were required to present their notebooks in order to track the progress of their learning. The contents of the notes were scanned and stored in a PC, and all texts were then converted into machine readable text. In addition to the recording of the contents of notes, the lecturer reviewed and assessed the contents of notes for 14 out of 15 weeks of the course. The contents were evaluated using a scale, and the sums of the scores were defined as individual note taking assessment score (NT-A). The overall survey conditions remained the same during the two years the course was studied.

This survey employed as an action research procedure. During the first year, the survey was conducted without any instructions having been given, nor suggestions to improve note taking skills made, and this condition is defined as the "without instruction" condition. Regarding our series of surveys (Nakayama, Mutsuura, & Yamamoto, 2011, 2014a, 2016b), note taking activity may be a key learning process, and can provide a possible means of improving learning performance. Therefore, instructions concerning note taking techniques and examples of good notes were shown twice during the course, at the beginning and mid-point of the course. This other condition is defined as the "with instruction" condition.

The valid number of participants is 32 for without instruction and 24 for with instruction.

Characteristics of participants

The individual characteristics of participants may affect their learning activity, as has been noted in previous studies (Nakayama, Mutsuura, & Yamamoto, 2014a). The following metrics were introduced to measure characteristics during the two surveys. The constructs are: Personality (Goldberg, 1999; International Personality Item Pool, 2004), Information Literacy (Fujii, 2007), Learning Experience (Nakayama, Yamamoto, & Santiago, 2007) and Note-taking skills (Nakayama, Mutsuura, & Yamamoto, 2014a). The

total number of variables is 13. The causal relationships between these characteristics and overall learning performance was confirmed using these metrics (Nakayama, Mutsuura, & Yamamoto, 2014a, 2016b).

Personality

The personalities of students were measured using an open resource such as the International Personality Item Pool (IPIP) inventory (International Personality Item Pool, 2004). This construct consists of 50 questions which measure the following five factor scores as follows: "Extroversion" (IPIP-1), "Agreeableness" (IPIP-2), "Conscientiousness" (IPIP-3), "Neuroticism" (IPIP-4) and "Openness to Experience" (IPIP-5).

Information literacy

Information literacy was measured using 32 question inventories which were developed and defined as 8 factor structures by Fujii (2007). The resulting two meta-factors were summarized as Operational Skills (IL-1), and Attitudes toward Information Literacy (IL-2) (Nakayama, Yamamoto, & Santiago, 2008).

Learning experience

Student's online learning experiences at the university were evaluated using three factor scores which were extracted using developed inventories. The inventories consist of 10 items which ask participants about their attitude toward learning. The items are then rated using a 5-point scale. Factor analysis of the surveyed data was conducted using Promax rotation. In the results, three specific factors were extracted (Nakayama, Yamamoto, & Santiago, 2007). The three factors are Factor 1 (LE-F1): Overall Evaluation of the e-learning experience, Factor 2 (LE-F2): Learning Habits, and Factor 3 (LE-F3): Learning Strategies (Nakayama, Yamamoto, & Santiago, 2007). The factor scores of student's learning experience may reflect their attitudes toward their learning behavior.

Note-taking skills

Though many universities provide students with self evaluation quizzes to measure their own abilities (Pennsylvania State University, Learning Centre, 2017), most Japanese universities do not provide these. A set of 17 question inventories was developed to measure student's note-taking skills. Also, three factor structures were extracted using factor analysis (Nakayama, Mutsuura, & Yamamoto, 2011, 2012). The three factors are as follows: NT-F1: Recognition of functions of note taking, NT-F2: Methodology of utilizing notes, and NT-F3: Presentation of notes.

Evaluations of note contents

Note taking assessment

The contents of all notes taken by participants in each of the sessions were evaluated by the lecturer (Nakayama, Mutsuura, & Yamamoto, 2014a, 2016b). The measurements were divided into two categories, "Good" and "Fair". The assessments of all sessions were summed up, and the mean scores were defined as note taking scores (NTS). As mentioned above, the machine readable texts of note contents were used in the following lexical analysis.

Nakayama *et al. International Journal of Educational Technology in Higher Education* (2017) 14:6

Page 5 of 14

Lexical analysis for contents of notes taken

The machine readable texts from the notes students took were prepared, though figures and tables were excluded. During the face-to-face sessions, the lecturer wrote down and presented points texts using a document camera. These documents were also converted into text along with the participant's notes.

The texts of the participants and the lecturer were analyzed lexically using the Japanese morphological term analysis tool MeCab (2006). During processing, the nouns and the sequence of appearance of each individual noun in the texts and in each session were recorded. To measure the note taking activity of participants, some indices were defined as follows (Nakayama, Mutsuura, & Yamamoto, 2014b, 2016b).

- Word ratio (*WR*): the ratio between the number of terms written down and the number of terms given (the number of terms participants recorded vs. the number of terms the lecturer presented)
- Coverage (*CV*): the coverage ratio was calculated as a percentage of the number of terms recorded by participants.

The above indices represent written features which are based on noun term frequencies. To detect key phrases such as conceptual terminology, term connections or co-occurring terms were summarized. Concurrent term connections are called as 2-gram (Nakayama, Mutsuura, & Yamamoto, 2014b, 2016a, 2016b). For example, noun transitions of terms such as A-B and B-C were extracted from a text A-B-C. The term connections between these two nouns can be mathematically summarized as an adjacency matrix. Therefore, the adjacency matrix for each individual and each session can be used to represent the features of notes taken. In addition to this, the matrices of the two sets of texts can be compared using a distance metric such as the Levenshtein distance, which is a kind of edit distance. In comparing the contents of texts of the students and the lecturer, students sometimes omitted terms the lecturer presented, or some students recorded terms that were in addition to those the lecturer presented. The frequency of this kind of behavior can be evaluated using a distance metric which is based on a comparison of the two matrices of adjacency.

As a result, the two indices are defined as follows:

- Additional distance (*AD*) means the sum of the number of additional nodes or edges in a matrix.
- Insufficient distance (*ID*) means the sum of the number of reduced nodes or edges in a participant's matrix in comparison with the lecturer's matrix.

Both distances are influenced by the total number of terms in the lecturer's presentation during sessions, so that the relative distances are calculated using the overall distances of written contents presented by the lecturer in each session. As a result, four note taking indices were extracted from each session and each individual, then the ground averages across all sessions were calculated, and partial averages for the first and the second halves of the sessions were calculated, respectively.

Results

Note taking activities

Some simple statistics are summarized using an analysis of the contents of notes taken.

First, the numbers of terms the lecturer presented in each session is summarized in Fig. 1. This course includes mathematical equations and presentations in several sessions, so that the numbers of nouns detected in those sessions is not large. Also, the number of nouns presented is small when the lecturer gave explanations using descriptions from the text book during a few of the sessions at the end of course.

Second, the note taking activity of participants was measured using the four metrics mentioned above. These means are summarized across two conditions in Table 1. As the table shows, all means for the two conditions are comparable, except for AD (additional distance). When note taking instruction was given, the mean AD distance increased in accordance with the WD (word rate). The number of nouns in student's notes reflects the active promotion of recording nouns in their notes. However, for four metrics there are no significant differences between the two conditions. The effectiveness of note taking instruction was not observed in these metrics.

In a detailed analyses of note taking activity, the relationships between these metrics were measured using correlation analysis. The correlation coefficients between metrics without instruction are summarized as a triangular matrix in Table 2. A subjective assessment of the contents of notes was also conducted. In the table, the non-significant coefficients are indicated using parentheses. When instruction was not provided, the word rate (WD) strongly correlates with additional distance (AD), and also with insufficient distance (ID). Active writing promotes the replacement of terms the lecturer presented with the participant's own terms. In regards to the phenomenon, coverage (CV) negatively correlates with ID. That means that ID increases when CV decreases. In an overall assessment, NTS positively correlates with WD, ID and AD, since more note taking activity provides the lecturer with a better impression of the contents of student's notes. There are no significant relationships between final exams scores (FE) and activity however, as no contribution due to note taking was observed.

Fig. 1 The number of words the lecturer presented in each session
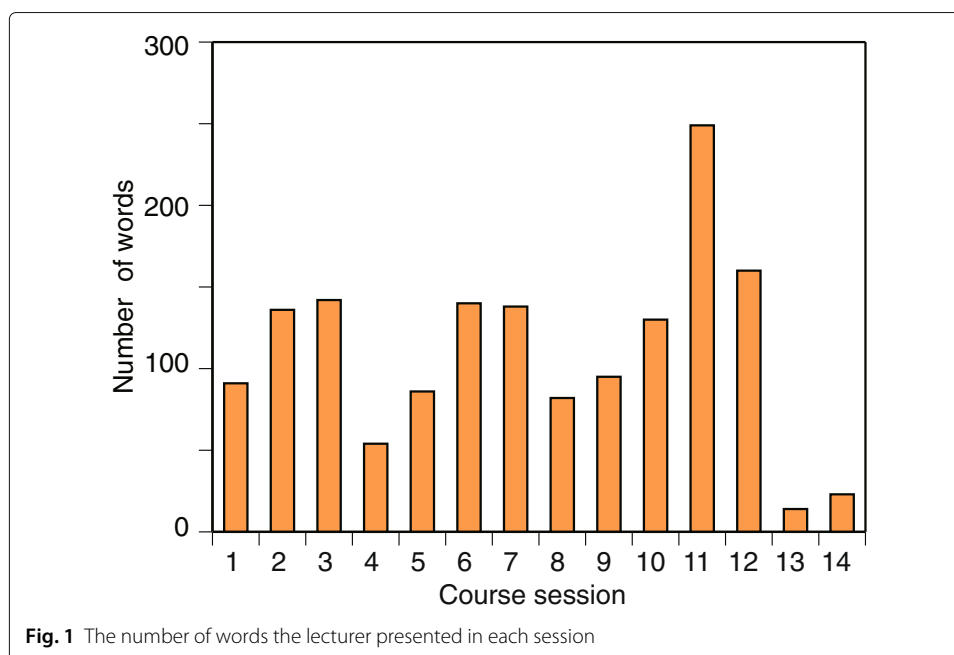
Table 1 Relative means of features of note taking activity

	Without Inst. (N=32)		With Inst. (N=24)	
	Mean	STD	Mean	STD
WD: word rate	1.51	0.40	1.66	0.46
CV: coverage	0.77	0.15	0.72	0.14
ID: insufficient distance	0.86	0.03	0.83	0.05
AD: additional distance	1.42	0.57	2.24	0.98

When note taking instruction was given, some relationships changed, as Table 3 shows. The word rate (WD) also correlates with coverage (CV), in addition to correlating with additional distance (AD). In this condition, CV correlates with AD. The results suggest that more note taking activity causes an increase in both CV and AD, so that participants record nouns the lecturer presented as well as nouns they created themselves. As a result, the lecturer's assessment of the contents of notes (NTS) is affected by the volume of nouns recorder, which is based on rates of both WD and AD. Also, all note taking metrics correlate with the final exams scores (FE). The activities of participants contribute to these relationships significantly, and the effectiveness of note taking instruction during the course was examined. The features of note taking activity for the two conditions were comparable, as Table 1 shows. The individual features may depend on the learning progress of the student. A detailed analysis will be presented in the following Performance in regards to the progress of the course.

Multiple regression analysis

Some of the effectiveness of note taking activity and instructions which promote note taking were confirmed during the respective correlation analysis in the previous section. The contribution of learning performance was not specified. To evaluate the impact of note taking instruction on participant's learning performance, the relationships between all variables mentioned above, and scores of final exams (FE) were examined. A set of variables consisted of 13 variables for participant characteristic variables and four note taking activity variables. A possible relationship using the set of variables ($\mathbf{x}$) can be noted as follows.

$$\mathbf{x} \in \{WR, CV, AD, ID, IPIP_{1-5}, IL_{1-2}, LE_{1-3}, NT - F_{1-3}\}$$
$$G(\mathbf{x}) = \ <\mathbf{w}, \mathbf{x}> \ + b$$

Performance comparisons between the first and second halves of the courses

First, a simple multiple regression analysis using liner models for $G(\mathbf{x})$ was introduced, in order to determine key variables and their contributions to the scores of final exams.

Table 2 Correlation coefficients across features of NT (without Inst.; N=32)

	WD	CV	ID	AD
WD: word rate	–			
CV: coverage	(–.09)	–		
ID: insufficient distance	0.47	–.53	–	
AD: additional distance	0.94	(–.05)	0.54	–
NTS: note taking scores	0.76	(–.10)	0.48	0.79
FE: scores of final exams	(–.10)	(0.05)	(0.02)	(–.02)

Nakayama *et al. International Journal of Educational Technology in Higher Education* (2017) 14:6

Page 8 of 14

Table 3 Correlation coefficients across features of NT (with Inst.; N=24)

	WD	CV	ID	AD
WD: word rate	-			
CV: coverage	0.49	-		
ID: insufficient distance	(-.24)	-.75	-	
AD: additional distance	0.96	0.49	(-.27)	-
NTS: note taking scores	0.70	(0.36)	(-.24)	0.56
FE: scores of final exams	0.49	0.72	-.53	0.48

During the analysis, a step wise method of selection was introduced to choose the variables which were significant. This procedure can optimize the model in order to calculate the rate of contribution, which is expressed as an R-squares.

The results are summarized in Table 4. After taking into consideration the stimulation of note taking instruction and the differences in the number of terms the lecturer presented in Fig. 1, four note taking metrics were summarized as three conditions: means of the 1st and the 2nd halves of the course, and the overall course.

In the results, the effective variables and their contributions (partial R-squares) are summarized and compared between two conditions, such as data with instruction and data without instruction. In a comparison of an index of overall performance of the regressions between the two conditions, the R-squares for the condition with instructions are much higher than for the other condition. In regards to these results, the contributions of final exam scores can be explained using individual sets of variables during the course with instruction. Also, performance is better when variables from the first half of the course are employed. The contributions using multiple regression models are higher than are the ones for single regression models. Therefore, every index can contribute to the relationships.

The lists of variables indicate that some variables of student's characteristics were selected from the course without instruction, while note taking features were included when instruction was given. In particular, four metrics of note taking activity in the first session, and NT3 (presentation of notes) were selected for regression analyses of the 1st half of the course, and the 1st and 2nd halves of the course.

Table 4 Selected variables and R-squares as a contribution to multi regression analysis

Data set	Instruction	Selected variables (partial R^2)	R^2
Overall	N/A	IPIP1(0.15), LE1(0.10)	0.25
	done	mCV(0.52), IPIP5(0.10), LE1(0.04)	0.66
1st half	N/A	IPIP2(0.15), LE1(0.10), fAD(0.09)	0.34
	done	fCV(0.51), fID(0.13), NT3(0.10), fAD(0.04), fWD(0.03)	0.81
2nd half	N/A	IPIP2(0.15), LE1(0.10)	0.25
	done	sCV(0.49), IPIP5(0.09)	0.58
1st & 2nd	N/A	IPIP2(0.15), LE1(0.10), fAD(0.09)	0.34
halves	done	fCV(0.51), fID(0.13), NT3(0.10), fAD(0.04), fWD(0.33), sID(0.04)	0.85

IPIP1 extroversion, *IPIP2* agreeableness, *LE1* Overall evaluation of the e-learning experience
NT1 function of note taking, *NT3* presentation of notes, *mCV* mean CV, *mID* mean ID, *mAD* mean AD
fCV mean CV for 1st half, *fWD* mean WD for 1st half, *fID* mean ID for 1st half, *fAD* mean AD for 1st half
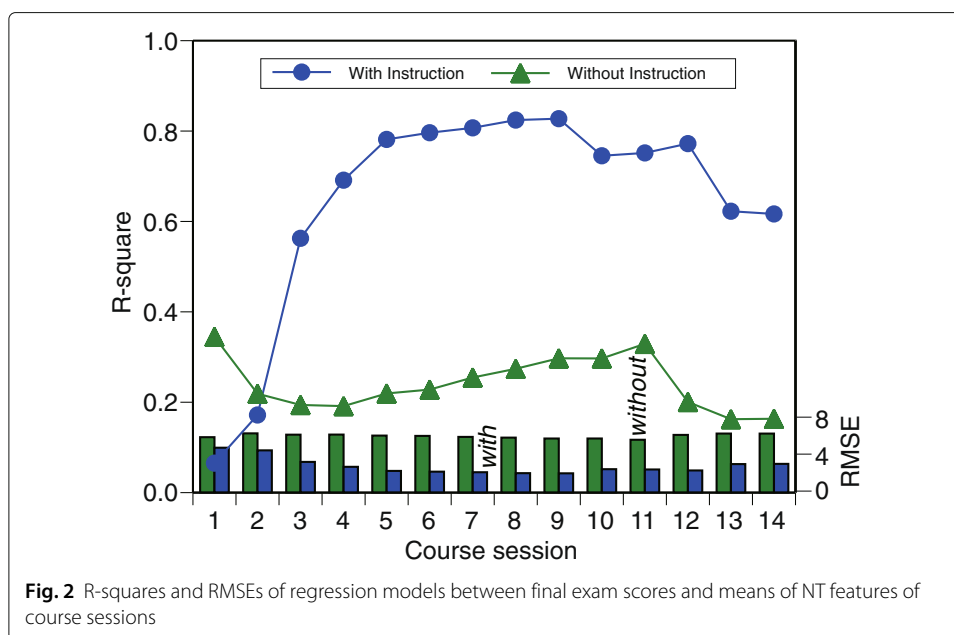*sCV* mean CV for 2nd half, *sID* mean ID for 2nd half, *sAD* mean AD for 2nd half

Performance in regards to the progress of the course

Since participants learn content and make improvements to their notes during each session, learning performance may be based on cumulative learning behavior. For example, the learning performance of the i th session may reflect overall learning activities from the first to i th session. As mentioned in the previous section, the four note taking metrics and one of the note taking skills (NT-F3) contributed to final exams scores (FE) such as $\mathbf{x} \in \{WR, CV, AD, ID, NT - F_3\}$, and temporal prediction performance in regards to the evaluation of cumulative indices.

The contribution ratios of the variables were calculated as R-squares, and the accuracy of prediction was indicated as root mean square errors (RMSE). The temporal changes are summarized in Fig. 2. The contribution ratios stay around 0.2–0.4 without instruction. When instruction was made, the ratios increased with the number of sessions. During sessions 5–12, the ratios were distributed around 0.8, which was quite a high rate. In the last two sessions, the ratios decreased, though they were above 0.6. In addition to these improvements, after the fourth session RMSEs with instructions tended to be smaller than those for without instruction. These results suggest that the effectiveness of note taking activity appears at an early stage in the course, and can be maintained until the end of the course.

Possibility of prediction of final exams scores

In regards to the results of multi regression analysis, the scores of final exams reflect metrics of note taking activity with instruction. The results of analysis show some significant relationships between the scores of final exams (FE) and the four metrics of note taking activities and the one factor score for note taking skills (NT-F3). The accuracy of prediction for FE scores was not evaluated precisely. Here, prediction accuracy means precision of estimation of scores of a novel participant using his or her scores of metrics of note taking behavior as an index of the capability to generalize.

Fig. 2 R-squares and RMSEs of regression models between final exam scores and means of NT features of course sessions

To estimate scores accurately, a support vector regression (SVR) technique was introduced as a more robust model. A Gaussian kernel was employed as the function (G) using the constant b. Prediction accuracy was evaluated using a leave-one-out procedure which estimated the scores of each individual using a model trained with the rest of the data. With this, the model training process and prediction process can be evaluated independently of each other.

Prediction performance using SVR

The computed function G($\mathbf{x}$) can provide estimates of the scores of final exams ($\tilde{FE}$) after optimization training using data of participant that does not target any one specific individual. The actual calculation was conducted using a LIBSVM package (Chang & Lin, 2011). To evaluate prediction performance, R-squares and prediction error RMSEs were calculated. The performance was tested under several conditions using selected variables. The results of some trials suggest that the four features of note taking activity contribute to performance prediction.

The results of several feature sets are summarized in Table 5. In comparing the performance of estimations between with and without note taking instruction, the predictions were more accurate when note taking instruction was given. The results of R-squares show that some note taking instruction metrics can contribute to the prediction of FE ($R^2 > 0.6$), though the contribution of these metrics without instruction is low. Both RMSEs for the two sets of data are comparable with those which use multiple regression analysis, as Fig. 2 shows.

Prediction performance of course progress

The possibility of predicting final learning performance provides useful information to improve student's learning activities during the course. In particular, generalization and the period of evaluation are important issues. Prediction performance using SVR and the cumulative information about participant's learning features is calculated in this section, and in subsection "Performance in regards to the progress of the course" as well.

Here, prediction was also conducted using cumulative values of four note taking features and the note taking factor score (NT-F3) and comparing them with subsection "Performance in regards to the progress of the course". Performance is displayed in Fig. 3 using the same format as in Fig. 2. R-squares remain around 0.6 after the 4th session,

Table 5 R-squares and RMSEs between final exam scores and predictions across sets of selected feature variables

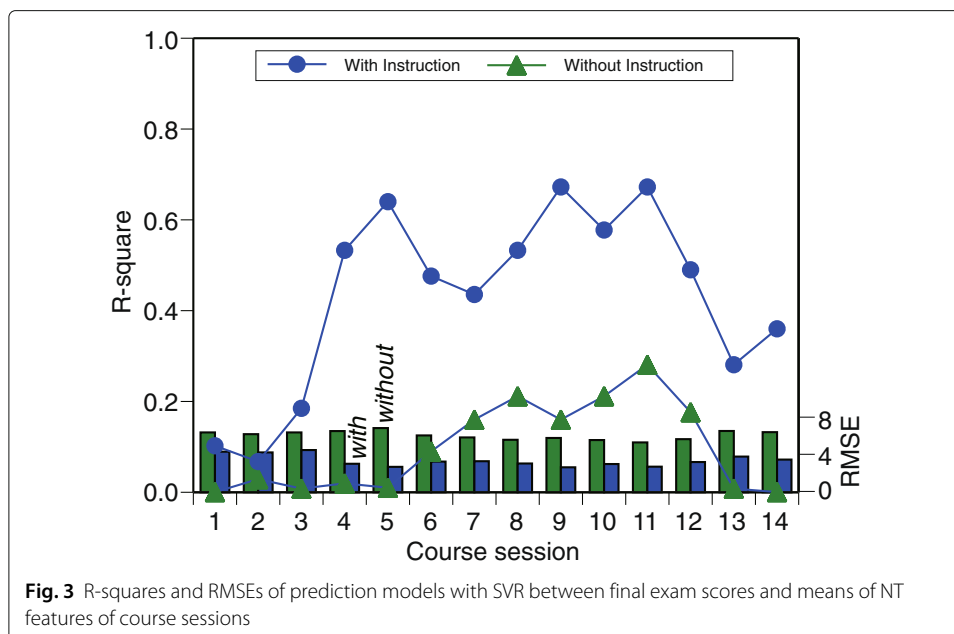| Feature set	R-squares		RMSE	
	Without Inst.	With Inst.	Without Inst.	With Inst.
NT-f + 13 variables[a]	0.06	0.07	6.6	4.4
NT-f + 2 variables[b]	0.17	0.64	5.9	2.8
NT-f + 2 variables[c]	0.18	0.72	5.9	2.5
NT-f (the first half means)	0.04	0.36	6.7	3.6
NT-f (the second half means)	0.02	0.02	6.9	4.9
3 variables[d]	0.04	0.08	6.6	5.1

NT-f: Means of features of NT(WD, CV, ID and AD)
[a]Four features of overall means in NT-f and other 13 features of characteristics
[b]Selected features of NT-f (Mean WD and AD in the first half sessions, and Mean ID in the second half sessions)
[c]Four features of NT-f in the first half sessions, NT-F3, and Mean ID in the second half sessions
[d]IPIP2, LE-1, and Mean AD in the first sessions

Fig. 3 R-squares and RMSEs of prediction models with SVR between final exam scores and means of NT features of course sessions

except for the last two sessions when instruction was provided. When instruction was not provided, R-squares remained around 0.2. In order to evaluate prediction accuracy, a leave-one-out procedure was introduced. Though overall performance is lower than for multiple regression analysis, the accuracy of with instruction shows the possibility of estimating individual final exams scores as the course progress.

Discussion

As mentioned in the introduction, student's characteristics, including note taking scores (NTS), have been confirmed to have an affect on the scores of final exams. Also, the effectiveness of note taking instruction was introduced, though the detailed relationships between these variables were not specified. This paper tries to examine the relationships mathematically in order to improve the learning process.

Fundamentally, the relationships between the scores of final exams and features of note taking activities were analyzed. Though the metrics for note taking activity for courses with and without instruction are comparable in Table 1, the correlation relationships between the scores of final exams and features of note taking activities have changed due to the note taking instruction that was given. All correlation coefficients between variables were significant during the course with instruction, while the coefficients were not significant for the course without instruction. In comparing correlation relationships across metrics of note taking activity between Tables 2 and 3, the contents of notes partici-pants took may have changed. Generally, students replaced the terms presented with their own words in their notes. When note taking instruction was given, students recorded their own words in addition to recording the words presented. Therefore, the metric of additional words written down increased during the course with instruction, as is shown in Table 1.

In the next step, the process of increasing the effectiveness of note taking metrics was measured as a formative assessment, using two approaches. First, the relationship between the scores of final exams and variables of student's individual behaviors were

evaluated using multiple regression analysis and a step wise method of selecting variables. In the results, the overall contribution of most note taking activity metrics which were selected was high when instruction was given. However, the contribution of some variables of student's characteristics which were selected is small for the no instruction condition. The contribution of a set of metrics of note taking activity increased with the number of sessions as the course progressed. The contribution, when measured as an R-square, remained at a high level between the 4th and 12th sessions of the course, as is shown in Fig. 2. However, the contribution decreased in the last two sessions. In regards to the change in the number of terms the lecturer presented in Fig. 1, the number of terms in the last two session was the smallest of all sessions during the course. The lecturer explained the contents using mainly a textbook, and so the number of terms presented was small. Therefore, this information may influence the metrics of cumulative values.

When the relationships were validated using multiple regression analysis, the possibility of prediction of final exam scores during the progress of the course was confirmed. To do this, SVR was introduced as a robust prediction procedure to test note taking activity metrics. Also, the contribution to predicting final exam scores remained at around 0.6 between the 4th and 12th sessions of the course, as is shown in Fig. 3. These results suggest that it is possible to improve participant's final exam scores during the course. Since the prediction function is based on metrics of note taking activity, it may be possible to provide each participant with appropriate instruction regarding their individual note taking abilities.

However, the results are based on the case of one course at a single Japanese university, and the number of participants was not large. The validity of this approach should be investigated with care, and the validation for these points will be the subject of our further study.

Conclusion

The possibility of developing a diagnostic procedure to improve learning performance using participant's characteristics and features of note taking activity during a blended learning was confirmed. The relationships between these indices and the procedure used to predict the scores of final exams were examined. The following results were obtained.

1. Some metrics of note taking activity were defined using lexical analysis of the contents of notes taken by students. The statistics of two groups of students in courses with and without note taking instruction were compared. The effectiveness of note taking with instruction was observed in correlational relationships between these metrics.

2. The relationships between final exam scores and the metrics of note taking activity were analyzed, and the contribution of these metrics during the course with instruction was confirmed. Additionally, the relationships were established after only several sessions of the course, using formative analysis.

3. The possibility of predicting the scores of the final exams was confirmed using support vector regression (SVR) functions. Also, the contribution of the metrics of note taking activity was confirmed. The sessions available for predicting scores were examined using formative analysis.

These techniques can be applied to improving participant's learning activity through the use of better note taking methods. A detailed procedure and confirmation of the effectiveness of this using various other courses will be a subject of our further study.

Authors' contributions

MN and HY conceived and designed the survey experiments; KM performed the experiments; MN analyzed the data; MN wrote the paper and KM and HY reviewed it. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Information and Communications Engineering, Tokyo Institute of Technology, Ookayama, Meguro, 152-8552, Tokyo, Japan. [2]Faculty of Economics, Shinshu University, Asahi, 390-8621, Matsumoto, Japan. [3]Shinshu University, Asahi, 390-8621, Matsumoto, Japan.

References

Bates, A.W.T. (2000). *Managing Technological Change: Strategies for College and University Leaders*. San Francisco, CA, USA: Jessey-Bass Publishers.

Bell, C., Jones, N., Lewis, J., Qi, Y., Kirui, D., Stickler, L., …Liu, S. (2015). *Understanding Consequential Assessment Systems of Teaching: Year 2 Final Report to Los Angles United School District*. Princeton, NJ, USA: Educational Testing Service.

Bennett, R.E. (2015). The changing nature of educational assessment. *Review of Research in Education*, *39*, 370–407.

Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York, USA: McGraw-Hill Inc.

Cronbach, L.J., & Snow, R.E. (1977). *Aptitudes and Instructional Methods*. New York, USA: Irvington Publishers, Inc.

Chang, C.C., & Lin, C.J. (2011). Libsvm: a library for support vector machines. *ACM Transaction Intelligent Systems and Technology*, *2*, 27–12727. http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Fujii, Y. (2007). Development of a scale to evaluate the information literacy level of young people –comparison of junior high school students in Japan and Northern Europe. *Japan Journal of Educational Technology*, *30*(4), 387–395.

Gulikers, J.T.M., Bastiaens, T.J., & Kirschner, P.A. (2004). A five-dimensional framework for authentic assessment. *Educaitonal Technology Research & Development*, *52*(3), 67–86.

Goldberg, L.R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, *7*, 7–28.

International Personality Item Pool: (2004). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences. http://ipip.ori.org. Accessed 28th, Oct. 2004.

Kiewra, K.A. (1985). Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology*, *10*, 378–386.

Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, *1*(2), 147–172.

Kiewra, K.A., Benton, S.L., Kim, S.I., Risch, N., & Christensen, M. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology*, *20*, 172–187.

Kobayashi, K. (2005). What limits the encoding effect of note-taking? a meta-analytic examination. *Contemporary Educational Psychology*, *30*, 242–262.

MeCab: (2006). Yet Another Part-of-Speech and Morphological Analyzer. http://mecab.sourceforge.net. Accessed 27 Dec 2010.

Nakayama, M., Kanazawa, H., & Yamamoto, H. (2009). Detecting incomplete learners in a blended learning environment among Japanese university students. *International Journal of Emerging Technology in Learning*, *4*(1), 47–51.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2011). Evaluation of student's notes in a blended learning course. *International Journal of New Computer Architectures and their Applications*, *1*(4), 1080–1089.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2012). Causal analysis of student's characteristics of note-taking activities and learning performance during a fully online course. In *Proceedings of 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communication* (pp. 1924–1929). Liverpool, UK.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2014a). Impact of lernear's characteristics and learning behaviour on learning performance during a fully online course. *The Electronic Journal of e-learning*, *12*(4), 394–408.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2014b). A note taking evaluation index using term networks in a blended learning environment. In *Proceedings of Eighth International Conference on Complex, Intelligent and Software Intensive Systems* (pp. 486–490). Birmingham, UK.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2015a). Relationship between the final test scores and indices of note-taking activity. *Japan Journal of Educational Technology*, *39(Suppl.)*, 53–56.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2015b). The prediction of learning performance using features of note taking activities. In *Proceedings of 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)* (pp. 325–330). Brugge.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2016a). Note-taking evaluation using network illustrations based on term co-occurence in a blended learning environment. *International Journal of Distance Education Technologies*, *14*, 77–91.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2016b). Lexical analysis of student's learning activities during the giving of instructions for note-taking in a blended learning environment. *International Journal of Information and Educational Technology*, *6*, 1–6.

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2017). Effectiveness of student's note-taking activities and characteristics of their learning performance in two types of online learning. *International Journal of Distance Education Technologies*, *15*.

Nakayama, M., Yamamoto, H., & Santiago, R. (2007). The impact of learner characteristics on learning performance in hybrid courses among Japanese students. *The Electronic Journal of e-Learning*, *5*(3), 195–206.

Nakayama, M., Yamamoto, H., & Santiago, R. (2008). Impact of information literacy and learner characteristics on learning behavior of Japanese students in online courses. *International Journal of Case Method Research &, Application*, *XX*(4), 403–415.

Nye, P.A., Crooks, T.J., Powley, M., & Tripp, G. (1984). Student note-taking related to university examination performance. *Higher Education*, *13*, 85–97.

Pennsylvania State University (2017). Learning Centre: Lecture Note Taking. http://www.ulc.psu.edu/studyskills/note_taking.html. Accessed 30 Aug 2010.

Piolat, A., Olive, T., & Kellogg, R.T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology*, *19*, 291–312.

Seaton, D.T., Nesterko, S., Mullaney, T., Reich, J., & Ho, A. (2014). Characterizing video use in the catalogue of MITx MOOCs. *eLearning Papers*, *37*, 33–41.

Seaton, D.T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D.E. (2014). Who does what in a massive open online course? *Communication of the ACM*, *57*(4), 58–65.

Ueno, M. (2007). Online outlier detection for e-learning time data. *IEICE Transaction*, *J90-D*, 40–51.

Weener, P. (1974). Note taking and student verbalization as instrumental learning activities. *Instructional Science*, *3*, 51–74.