Teacher Interventions to Enhance the Quality of Student Comments and their Effect on Prediction Performance

Shaymaa E. Sorour*†
* Dept. of Advanced Information Technology
Kyushu University, Japan
shaymaa@ma.ait.kyushu-u.ac.jp
Dept. of Educational Technology
† Faculty of Specific Education
Kafrelsheikh University, Egypt
shaymaa.srour@spe.kfs.edu.eg

Shaymaa Abd El Rahman
Dept. of Educational Technology
Faculty of Specific Education
Kafrelsheikh University, Egypt
shaymaa.abdelrahman@spe.kfs.edu.eg

Tsunenori Mine
Dept. of Advanced Information Technology
Kyushu University, Japan
mine@ait.kyushu-u.ac.jp

***Abstract*—Today, the use of learning analytics is becoming more crucial in the learning environment for the purpose of understanding and optimizing students' learning situations. The purpose of this paper is to examine the impacts of Teacher Interventions (TIs) on students' attitudes and achievements involved with the lesson by analyzing their freestyle comment data after every lesson. The current study proposes a new method for building an accessible prediction model, which represents students' activities, situations and viewpoints; the method classifies words in the student comments into six attribute types and indicates the most important types that affect the prediction results. Further, the prediction results are compared with the topic-based statistical method that uses Latent Dirichlet Allocation and Support Vector Machine models. The results proved that there were positive correlations between TIs and the quality of writing comments that affect on improving the prediction results.**

Keywords*—Quality education, Teacher Interventions, Comment Data, Attribute-base Method, Topic-based Method.

I. Introduction

Nowadays, Higher Education Institutes (HEIs) play a crucial role in improving the quality and the efficacy of education. Thus, it is important to ensure the quality of the educational processes and identify the methods by which they can be validated and improved in order to provide quality education to students [1], [2].

Quality education is one of the key responsibilities of any University/HEI to its stakeholders denoting not only the requirement for production of high level of knowledge but also the need for efficient provision of education, so that students achieve their learning objectives without any problem and improve their performance [1], [3]. One way of enhancing the quality of educational processes is to improve the decision-making procedures on the various processes by providing the administration of an educational institute with useful knowledge, which is currently unknown to the decision makers [2].

Understanding students' situations in the classroom is a challenging task. The instructor has faced some difficulties in observing all students and to grasp their learning situations. The current study uses comment data mining, as a type of text mining [4], to enhance the quality of educational processes in the classroom. Comment data have significant features of teaching and learning in university settings. They foster students' self-expression, construction of identity, understanding, and knowledge-building. Comment data also help teachers to modify the way of giving lessons, improve the quality of making communication with students and allow teachers at the same time to develop monitoring of assessment tasks. Furthermore, they enable student interactions, especially for the students with an introvert character, and help them express their views or asking questions [5], [6]. Although comment data have some benefits for students and teachers, it is not an easy task to encourage students to write good comments that express their behaviors, situations and attitudes involved with each lesson.

In this study, the capability of comment data mining is demonstrated to provide vital information of students' situations and attitudes and improve the quality of learning process and prediction results. The current study has two main objectives: (1) to address the necessity of TIs in collecting students' comments with high quality, (2) to build an accessible prediction model based on students' learning attributes that reflect their attitudes, understanding, learning situations, and so forth.

For the first goal, the TIs are characterized by a series of supports or scaffolds, whereby students are guided through the learning process. We emphasis on the progress with small steps based on our previous experiences to check students' achievements of active and successful participation with their peers. For the second, the extracted words from all student comments are classified into six attributes: lecture's objective, understanding, cooperation, lecture subject, students' attitude and others; the words related to the attributes are selected from student comments and an attribute is assigned to each selected word; using this alignment, each student comment is transformed into an attribute vector. The attribute vectors are used to build a prediction model of final student grade. We call this method Attribute-based Method (AM). Using the attributes will help the teacher to understand student learning situations more deeply in the learning environment and to provide vital feedback to each grade group. Moreover, the prediction results are compared with those of an existing statistical method using Latent Dirichlet Allocation (LDA) and Support Vector Machine (SVM) models; we call this method Topic-based Method (TM). The main point of this method is to extract topics from student comments and build strong relationships between input data (topics) and a final student grade, which is the output data. The merit of TM is to extract latent topics from student comments and build a prediction model of final student grades automatically. On the other hand, it is difficult to interpret the results and give the meaning appropriate to the topics.

A. *The contributions of this research are as follows:*

- The factors or attributes from student comments are identified that affect students' achievement and help to improve their performance.
- AM and TM are proposed to predict final student grades.
- Understandable prediction models are built by classifying comment data into the main six types of attributes that reflect students' attitudes and situations in each lesson.

The rest of the paper is organized as follows: Section II gives an overview of some related work. Section III describes the problem formulation. Section IV displays the prediction model of final student grade . Sections V and VI display and discuss some of the highlighted experimental results. Finally, Section VII concludes the paper and notes some open issues for future research.

II. Related Work

Predicting students' performance is one of the oldest and most useful applications of data mining and its goal is to estimate the unknown values of students' performance, such as knowledge, score or mark from other information, and aspects or behavior of students [7]. However, it is a difficult problem to solve due to the large number of factors or their characteristics, such as demographic, cultural, interactions between student and faculty, etc. [8].

Recently, there has been substantial work on constructing student models that can accurately predict student performance (e.g. [7]–[12]). From an application perspective, typical data mining algorithms usually work as black boxes, and grasping the factors that affect students' academic achievement is a critical input to understanding and improving the educational landscape. Moreover, several studies analyzed student attitudes by creating the criteria and questionnaires (e.g. [13]–[15]) to investigate how the student's attitude affected their achievements in the class. Moreover, Sen et al. [16] identified the factors that lead students to success or failure in placement tests. They developed models to predict secondary education placement test results, and used sensitivity analysis on those prediction models to identify the most important predictors. Minami et al. [17] investigated correlations of student achievement and their learning characteristics by analyzing the texts of their answers to a questionnaire. The results showed there were different words usage distribution between low, middle and high achievement students. Milena and Rubiano [18] applied data mining techniques (J48 and PART) to students' academic and demographic data to predict their performance and find the most adequate to extract classification rules. The results indicated that the sociodemographic variables influenced academic performance prediction, specially marital status and social stratum. In addition, the extracted rules could provide the relationships between attributes to predict poor academic performance and reduce the possibility for academic desertion. Ashenafi et al. [19] presented a linear regression model that utilizes data generated by the activities of students in two courses to predict their final exam scores. They implemented the web-based peer-assessment system and used the data from the system to build the model. They employed 14 features that capture various activities of students. The Root Mean Squared Error (RMSE) of the prediction models is used to evaluate their performance. The final model recorded an RMSE of 2.93 for one course and 3.44 for another one in predicting students' grades. Petkovic1 et al. [20] focused on assessment and prediction of Software Engineering (SE) teamwork in terms of ability of student teams to apply best SE processes and develop SE products. They used the quantitative measures of team activity from multiple sources, such as: the statistics of student time use, software engineering tool use, and instructor observations. They presented a machine learning framework which applied the Random Forest (RF) algorithm to the team activity measurements and team outcomes focusing on predicting teams. The results showed that RF was able to predict SE process and SE product team performance in intuitively explainable manner.

On the other hand, comment data mining have recently been used to understand the student behavior and situation. Goda et al. [21], [22] proposed the PCN method to estimate student learning situations on the basis of freestyle comment. The PCN method categorizes the comments into three items: P (Previous activity), C (Current activity), and N (Next activity). Item P indicates the learning activity before the class time. Item C shows the understanding and achievements of class subjects during the class time, and item N expresses the learning activity plan until the next class. They conducted a study on using the PCN scores to determine the validity of assessment based on student comments and showed there were strong correlations between the PCN scores and the accuracy of predicting final student grades.

For the purpose of predicting student grades from comment data, Luo et al. [23] applied Word2Vec [24] and Artificial Neural Network (ANN) to student comments. The average prediction F-measure result of final student grades was 63.8%. In addition, Sorour et al. [5], [25] proposed methods that analyzed student comments using Latent Semantic Analysis (LSA) and probabilistic Latent Semantic Analysis (pLSA). They compared the two models: SVM and ANN. The results indicated that the SVM model outperformed the ANN model. Their methods achieved an average F-measure results for LSA and pLSA, 52.27%, and 70.4%, respectively. Sorour et al. [26] conducted another experiment based on the LDA model and generated prediction models with SVM. The LDA model had significant results compared to the pLSA and the LSA models. It achieved an average F-measure result 76.9% in predicting final student grades. Although the prediction models proposed have achieved good prediction accuracy, they have difficulty to interpret their provided results and judgment criteria.

From the previous studies, we can conclude that researchers need to elaborate and go further into such factors that influence student learning and help to build an interpretable prediction model of student performance. In this study, we propose new methods to collect and analyze student comments and to build a new understandable prediction model.

III. Problem Formulation

A. Target of Comment Data Collection

The targets for collecting comments in this study were 150 undergraduate students at Faculty of Science and Arts from Saudi Arabia, who took the course covered computer applications and their advantages in education. After finishing this course, students have learned how to use the computer technical competencies and integrate them into the classroom teaching.

Teachers asked the students to write and submit their comments after each lesson. Student comments were collected for 12 lessons (one a week). Comment data can potentially eliminate barriers between students and their teacher, and provide opportunities for students to think back on their learning behaviors in connection with the subject. Thus, collecting comment data becomes an essential accessory to support learning in the classroom.

B. Teacher Interventions for Collecting Comment Data

Collecting comment data with high quality involved with the lesson and reflecting students' attitudes, tendencies and activities is a hard task. To this end, a teacher should give some pointers that lead students to describe their comments well. Moreover, the teacher should continuously tell students the importance of writing comments in order to help the students think back and improve their learning.

In this study comment data were collected using the same PCN method [21]. The teacher asked students to fill in three simple items about their learning status: P-, C- and N. Moreover, the teacher presented some instructions interventions to students through the three intervals: Lessons 1 to 4, 5 to 8, and 9 to 12, as shown in Table I. In order to extract the necessity assignments from all students and provide general instructions to the students based on several cases. We describe the content and problems arisen in each interval as follows:

1) **The first interval:** The teacher (1) displayed the advantages of writing comment data (e.g., they help to improve your performance, they give your viewpoints about the lesson, they show how much efforts you did), (2) presented instructions to follow (e.g., please describe your actual actions, feel free to display your progress to achieve your goals. For example, you can write clearly about understanding the lesson, the problems you faced and the attempts you did), and (3) showed some examples of comment data.
The problem: The teacher faced many difficulties to obtain comments that present students' situations; most of them only described the lesson objective and sentences related to the lesson subject by copying them from the lesson materials. Others did not have the confidence to write their situations and did not appreciate the importance of writing their comments. So, the instructor had to guide students in describing their comments. Giving some pointers helped students to describe their learning situations, behaviors, and attitudes in their comments accurately.

2) **The second interval:** The teacher employed the PCN method with the format consisting of some questions. In this research, we chose only C-comments that included four questions as follows:

(a) What did you learn from this lecture?
(b) Do you have anything you did not understand from this lecture? any question?
(c) What are the good points and bad points for this lecture?
(d) Did you teach something to your friends, get answers or discuss with them?

The questions were chosen based on our previous experience of collecting comment data. We decided to focus on the important questions regarding students' attitudes, activities and make an effort connected with the lesson subject. In addition, we used the same items of the PCN method to enable teachers to acquire a temporal learning status to each student and examine students' behavior in each lesson (e.g. review or prepare the lesson, study the current lesson).

In question (a) students wrote about the objective of the lesson and the lecture subject. In question (b) students described whether they understood the lesson or not and what parts they didn't understand and why. In question (c) students illustrated the good points and bad points of the lecture; they wrote about their attitudes, especially if they were willing to improve their level and overcome problems. Also, some of them presented the bad and good points of the lecturer (e.g. she explained so fast, she explained in an easy and interesting way, or her voice was small). In question (d) students displayed the cooperation with their friends by discussing, teaching or solving problems. From the previous questions and the teacher's courteous instruction, the students described their comments better and presented their situations and activities more detail and accurately.
The problem: Although several studies found that the cooperation had a positive impact on performance goals and learning motivation in the classroom (e.g. [27], [28]), students did not care about the cooperation attribute.

3) **The third interval:** The instructor gave some good points about the quality of student comments and encouraged them to collaborate together by discussing, teaching, solving problems, and helping their friends. The instructor explained the importance of cooperation to improve the students' level, to discover the mistakes and the problems in studies.

Table II presents the examples of comment data, and Figure 1 shows an example of TIs and feedback to improve the quality of student comments.

IV. Prediction Model of Final Student Grade

This section describes the procedures and methodology for building prediction models of final student grade.

A. Preliminary

1) Students' Grade: We consider predicting each student's results from his/her comments. Five grades: A, B, C, D and E were chosen instead of the mark itself as a student's result. The number of the students of grade E is smaller than that of

TABLE I: Instructions of writing the comments in the three intervals.

	Lessons (1-4)	Lessons (5-8)	Lessons (9-12)
Instructions	The teacher explained the advantages of comment data and gave the instructions of describing the current activity (C-comments) giving some examples.	The teacher employed the format of C-comment consisting of 4 questions.	The teacher provided advice to improve students' writing skills by presenting some examples and the importance of the cooperation with peers.
Case Study	Most of the students described the lesson's objectives and lecture subject; their comments copy a text from the book.	Most of the students described their current activity using six types of attributes (see Section IV-B2).	Most of the students cooperated with their friends
Feedback	Students had difficulty in expressing themselves although the teacher gave feedback after each lesson.	Although most of the students described their learning situations, some of them didn't write about the cooperation attribute.	Most of the students described their current activities including several attitudes.

TABLE II: Examples of comment data.

Question	Example
Lesson(1-4)	I knew the advantages of computer applications for teaching. It can help teacher to present the lesson using several ways: video, text and animation. The internet access also allows the students to get different educational resources ..,etc.
question 1	Today, I learned an interesting topic about the advantages of computer applications on Education.
question 2	I didn't understand how to employ the techniques with the chemistry and physics subjects.
question 3	I can apply several techniques for improving my skills in teaching. I am willing to learn those applications. Today, the teacher was very fast while explaining the simulation phenomenon.
question 4	I discussed with my friends about the advantages of simulation and virtual reality models for learning science subjects.

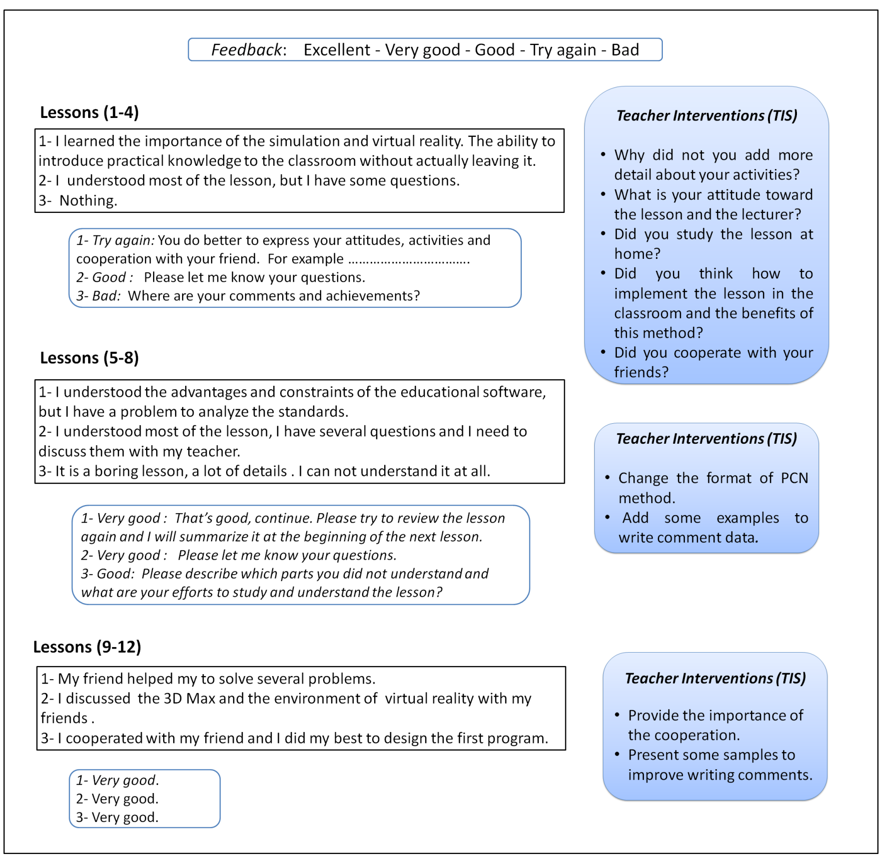

Figure 1: Example of TIs and feedback.

other grades. So we treated comment data of grade E as grade D. Table III shows the correspondence relation between the grades and the marks. The assessment to a student was done by considering the average mark of the student's exam assigned three times over the whole of the semester, the attendance rate, and students' activities in the classroom.

TABLE III: Grades categories.

Mark	Grade	Total #Student
100-90	A	30
89-80	B	48
79-70	C	46
69-60	D	21
59-0	E	5

2) Pre-Processing Data: Students described their comments in Arabic. To analyze their comments and extract words and parts of speech, they were translated into the English language. We used KH coder [29] to extract words and their parts of speech, including the occurrence frequencies of words in comments. In this study, only verbs, nouns, adjectives, and adverbs were used. Although student comments in Arabic were translated into English, the attribute type of the translated words in English correspond to the same one of the words in Arabic; the statistics of attribute type of the words in English are the same as those of the words in Arabic, thus the results of the two languages are the same.

B. Topic-based and Attribute-based Methods

1) Topic-based Method (TM): We create a word-by-comment matrix with extracted words. This word-by-comment matrix, is comprised of m words w_1, w_2,. . , w_m in n comments c_1, c_2, ...,c_n, where the value of each cell a_{ij} indicates the total occurrence frequency of word w_i in comment c_j. The number of words appearing in each comment averages 18 words, the number of words appeared in the comments is about 2200 in each lesson, and the number of words in all the comments without duplication is over 600 in each lesson. In this research, we employ LDA [30], which is used in statistical natural language processing to discover categories and topics from documents.

Here, we briefly outline the principle of the LDA model according to our experiment. Suppose we have N comments containing words from a vocabulary of size M words. The whole collection is summarized by an $M \times N$. In addition, there is a hidden (latent) topic variable z_k associated with the occurrence of a word w_i in a comment c_j. The model uses latent topic variables and probabilistic sampling techniques to generate the comments; comments are produced by sampling topics from a distribution of topics over comments and sampling words from a distribution of words over each topic [30]. The graphical representation of the LDA model is shown in Figure 2, where W_c is the number of words w_i in comment c_j. The goal is to maximize the following likelihood:

$$p(w|\phi,\alpha,\beta) = \int \sum_{z} p(w|z,\phi)p(z|\theta)p(\theta|\alpha)p(\phi|\beta)c\theta \quad (1)$$

where θ and ϕ are multinomial parameters over the topics and words, respectively and $p(\theta|\alpha)$ and $p(\phi|\beta)$ are Dirichlet distributions parameterized by the hyperparameters α and β. Since the integral is intractable to solve directly, we solve for the parameters using Gibbs sampling, as described in [31]. The hyperparameters control the mixing of the multinomial weights. When building the LDA model we run the MATLAB program to extract topics from comment data and evaluated the consequences of changing the number of topics T in the training phase by calculating the F-measure (F_1) results and confirming the optimal number of topics in each lesson. The optimal numbers of topics for our experiments were between 11 to 16 topics from lessons 1 to 12.

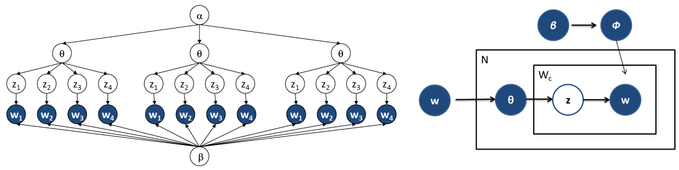

Figure 2: Graphical model representation of LDA [32].

2) Attribute-based Method (AM): The main objective of this research is to improve the quality of comment data and not only collect a text related to students' learning subject but also understand student sitations and attitudes in the learning environment, and build an understandable prediction model. The teacher can see students' learning situation, attitude to the lesson and their performance. We first assigned one of the six attribute types to each word extracted from student comments. The change of comment quality was evaluated by calculating the ratios of occurrence frequencies of attribute types to each student from lessons 1 to 12. The six attribute types are described as follows:

- **(a) Lecture's objective:** directly related to the general objective of each lecture (e.g., the advantages of computer techniques in education).
- **(b) Lecture subject:** related to the subject taken in the lectures, but not directly related to the objective (e.g., employ the skills of preparing educational lessons using learning theories, use educational communication skills via the internet and discuss the advantages of using computer applications in educational process).
- **(c) Understanding:** related to the understanding or learning the subject (e.g., understand, difficult, don't know, easy, clear).
- **(d) Attitudes:** related to students' positive attitudes or negative attitudes (e.g., can do, able to, interesting, self-dependence, enjoy, cannot do, unable, confuse, worry, boring).
- **(e) Cooperation:** related to the cooperation (e.g., cooperate, help, discuss, talk, ask, friend, explain).
- **(f) Others:** other words which do not appear in the previous types (e.g. think, go, item, fast). So, every word in a comment is assigned one of attribute types.

To show the effectiveness of word frequencies for the six types to all the students, we assign the weights to each attribute type in each lesson, where the weight represents the contribution of the type to each student who uses a word of this type. An example is shown in Figure 3.

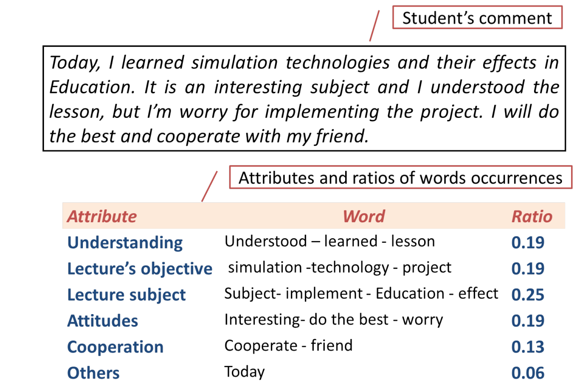

Figure 3: Correlation between extracted words with the attribute-value and the ratios.

The assignment is conducted based on the relation between student comments and the attribute type as follows:

Let s_i be the i^{th} students, $1 \leq i \leq 150$, G be the set of grades: $\{A, B, C, D\}$, and T be the set of attribute types: $\{a, b, c, d, e, f\}$. Let $k_{i,j}$ be the number of occurrences of attribute type $j (j \in \{a, b, c, d, e, f\})$ of the i^{th} student, s_i and $k^*_{i,j}$ is the ratio of occurrence of attribute type j of student s_i. Thus $k^*_{i,a}$, for example is the ratio of occurrence of attribute type a of student s_i. Here $\sum_{j \in a,b,c,d,e,f} k^*_{i,j} = 1$.

Figure 4 shows the ratios of attribute types to students in each grade in lessons 1, 5 and 9. In lesson 1, students in any grades didn't write about their cooperation. Further, students in grades A and B described their (d) Attitudes and (c) Understanding than other grades. In lesson 5, we can see the effect of teacher's interventions after formatting the comment data into four questions. The ratio of (c) Understanding, (d) Attitudes and (e) Cooperation was increased compared to that in lesson 1. In lesson 9, the ratio of (e) Cooperation attribute was increased and most of the students described their attitudes. Figure 5 shows the average ratios of attribute types a, c, d and e for lessons 1, 5 and 9. The ratios of (c) Understanding, (d) Attitudes and (e) Cooperation types were increased from lessons 5 and 9 than lesson 1. Students described their situations in addition to the lecture's objective. From lesson 9, students cooperated with each other more than in the previous lessons by getting instructions of their teacher.

C. Student Grade Estimation by SVM

SVM is a powerful solution to the classification problems [33]. For the TM, we have some pretreatments to prepare training data for SVM. From the previous phase, a number of topics and their corresponding words and probabilities were gotten. Then, all the words which one student used in his/her comment were found out and a topic vector was created according to the distinct words collected in the students comment. After obtaining a list of topic vectors of all students, the MATLAB LibSVM tool was used to predict final student grades. The SVM model with a radial basis function (RBF) kernel was employed to generate models from lessons 1 to 12 and predict a student grade as one of four grades: A, B, C and D, based on the analyzed comment data. For the AM, the input data to the SVM is the weight/ratios of the attribute types for each student. The SVM method with the RBF kernel was employed to generate models from lessons 1 to 12.

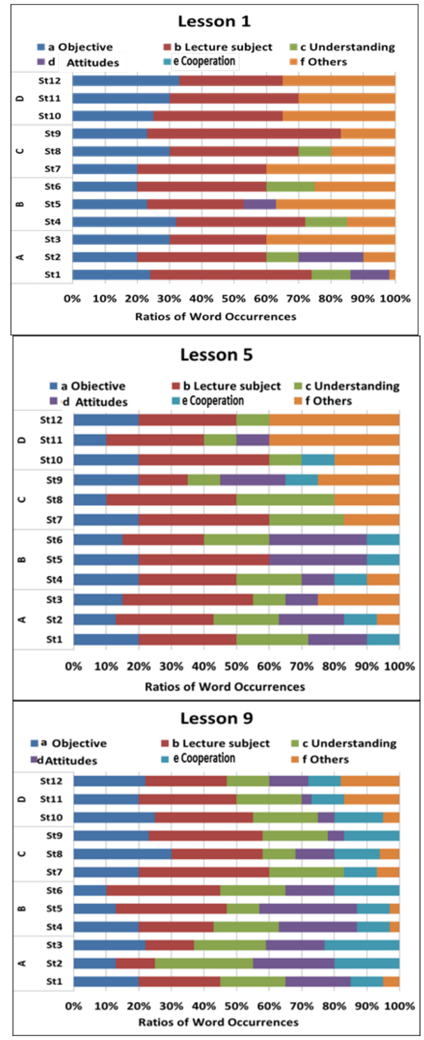

Figure 4: Ratios of word occurrences from type (a to f).

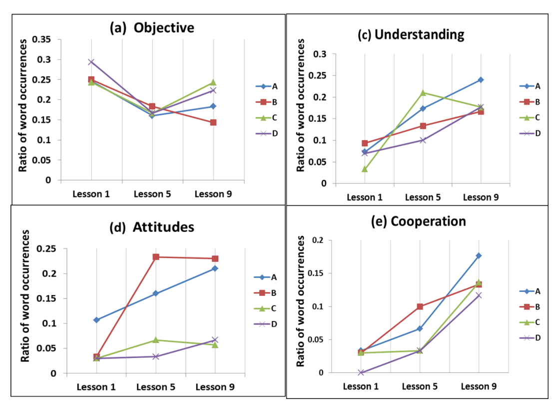

Figure 5: The relation between a,c,d and e types and the average ratios of word occurrences.

V. Evaluation Metrics

A 10-fold cross-validation approach is used to evaluate the proposed methods. Let G be the set of grades: $\{A, B, C, D\}$. Effectiveness for text classification was measured by accuracy (Acc) and F-measure (F_1) in each lesson.

Table IV is a confusion matrix whose entries are given as a function of two typical classes in comment-level classification, positive and negative comments. F-measure (F_1) is equal to the harmonic mean of Recall (R) and Precision (P) [34].

$$P = \frac{TP}{TP + FP} \quad , \quad R = \frac{TP}{TP + FN} \tag{2}$$

TABLE IV: Confusion matrix.

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

The Accuracy (Acc) score reflects the ability of a measurement to match the actual value of the quantity being measured [35]. The overall F_1/Acc score of the entire classification problem is computed by macro-average F-measure(F_1^{Ma}) and macro-average Accuracy (Acc^{Ma}).

In macro-averaging, F-measure for each lesson l_i is computed locally over each grade g first and then the average over all grades G is taken. R and P are computed, then F-measure for each grade g is computed and the macro-averaged F-measure is obtained by taking the average of F-measure values for each grade as follows:

$$F_{1_g} = 2 * \frac{P_g * R_g}{R_g + P_g} \quad , \quad F_1^{Ma} = \frac{\sum_{g \in G} F_{1_g}}{|G|} \tag{3}$$

The Accuracy (Acc). macro-average accuracy is computed as follows:

$$Acc_g = \frac{TP_g + TN_g}{TP_g + FP_g + TN_g + FN_g} \, , \, Acc^{Ma} = \frac{\sum_{g \in G} Acc_g}{|G|} \tag{4}$$

VI. Experiment Results

The following are the four research questions investigated in the paper:

- **(Q1):** Are there any correlations between the TIs and the quality of comment data to improve the prediction results?
- **(Q2):** Which model has the best results in predicting student performance, AM or TM?
- **(Q3):** Are there any difference between attribute types (e.g., understanding, attitudes, cooperation) in predicting final student grades?
- **(Q4):** Are there any clues to explain the prediction results obtained in each lesson?

To answer these research questions, the correlation of attribute types with the student grades for the AM was presented, then the comparison of the TM with the AM in predicting students' grades was displayed from lessons 1 to 12.

A. Correlation of Attribute Type with Student Grade

In order to investigate the effect of an attribute type to students' grades after TIs, Table V shows the correlation coefficient between attribute types from (a) to (f) and the student grade for lessons 1, 5 and 9. Figure 6 shows an example of the correlation between the cooperation and attitudes types and students' grades. It is easy to see that the overall correlation of lesson 5 was higher than lesson 1. In addition, types (c) (Understand), (d) (Attitudes) and (e) (Cooperation) had higher positive correlations than type a (Objective). On the other hand, types (f) Other and (b) Lesson subject showed the negative correlation.

TABLE V: Correlation coefficient between attribute type and student grade.

	Attribute Type					
Lesson	**a**	**b**	**c**	**d**	**e**	**f**
1	0.224	-0.15	0.174	0.235	0.336	0.131
5	0.368	-0.264	0.686	0.526	0.615	-0.233
9	0.543	-0.231	0.635	0.625	0.670	-0.233

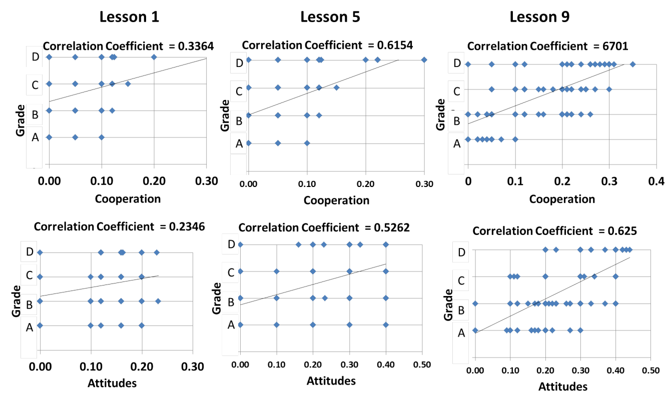

Figure 6: Correlation of the ratio of attribute types frequencies with student grades.

B. Overall Prediction Results

Tables VI and VII show the average overall prediction results for each grade using the AM and TM. Grade A scored the best results for the two methods. In addition, grade A for acde attribute types achieved the highest results. Figure 7 presents the F_1^{Ma} results between the two methods from lessons 1 to 12. The prediction performance of the AM achieved better results than the TM. The F_1^{Ma} were 74.2% and 76.9% for the TM and AM, respectively. To show the effect of the LDA model, the F_1^{Ma} with and without the LDA model was evaluated. The (SVM + LDA) model had the higher results than the SVM model only. On the other hand, to summarize the effect of attribute types on predicting students' grades, the F_1^{Ma} results with the different attribute types were shown in Figure 7. The prediction results for acde types had the best results among four cases whose F_1^{Ma} was 80.9%; this means that the selecting attribute types lead better prediction results compared to the results with all the types.

TABLE VI: Overall prediction results for TM.

	SVM				LDA+SVM			
Grade	P	R	F_1	Acc	P	R	F_1	Acc
A	0.662	0.712	0.686	0.622	0.719	0.810	0.759	0.695
B	0.652	0.750	0.698	0.661	0.761	0.746	0.755	0.709
C	0.613	0.584	0.598	0.642	0.731	0.701	0.718	0.671
D	0.611	0.563	0.586	0.543	0.762	0.718	0.738	0.735
Overall	0.635	0.652	**0.642**	0.617	0.743	0.744	**0.742**	0.703

TABLE VII: Overall prediction results for AM.

	all				abc				cde				acde			
Grade	P	R	F_1	Acc	P	R	F_1	Acc	P	R	F_1	Acc	P	R	F_1	Acc
A	0.758	0.844	0.798	0.728	0.778	0.811	0.794	0.768	0.818	0.841	0.829	0.782	0.833	0.881	0.856	0.808
B	0.785	0.781	0.778	0.741	0.776	0.792	0.784	0.767	0.798	0.811	0.804	0.742	0.814	0.824	0.819	0.781
C	0.731	0.737	0.738	0.699	0.762	0.756	0.759	0.754	0.785	0.728	0.755	0.743	0.812	0.763	0.787	0.769
D	0.767	0.761	0.761	0.753	0.771	0.776	0.773	0.722	0.785	0.726	0.754	0.703	0.816	0.734	0.773	0.743
Overall	0.761	0.781	**0.769**	0.730	0.772	0.784	**0.778**	0.753	0.797	0.777	**0.786**	0.743	0.819	0.801	**0.809**	0.775

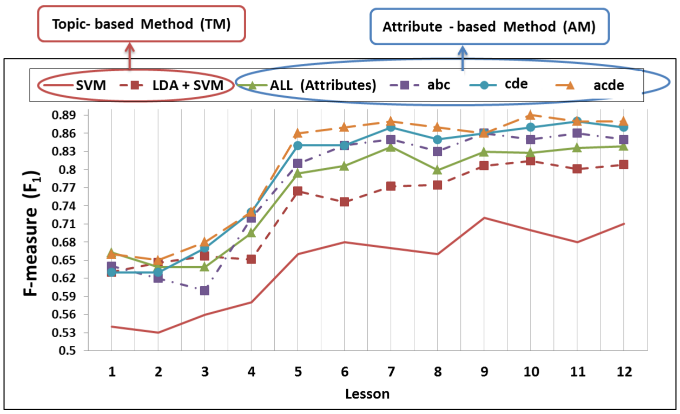

Figure 7: Overall prediction results for TM and AM.

TABLE VIII: Merits and demerits of TM and AM.

	AM	TM
Idea	Extract the main six attributes from all students.	Extract statistical topics.
Output	Attributes vectors.	Topics vectors.
Interpret the results	Users can interpret and understand the prediction results.	It is difficult to understand the prediction results.
Objective	Present an interpretable model by teachers, so that feedback can be provided.	Provide high prediction result, but it is quite hard to interpret it.
Usability	There is a persistent need to build a dictionary for extracting words related to the main six attributes.	Can automatically extract words and the related topics using the algorithm of the LDA model.

C. Discussion

By the end of this section, we can discuss the answer of the four research questions mentioned in Section VI.

- **(Q1):** For the purpose of collecting high quality comment data, we conclude that students can describe their attitudes and situations accurately, and the prediction results increased from lesson 5 after we introduced TIs and changed the format of comments collection. This finding will help the teacher to give feedback for improving students' achievements. The results are shown in Tables V to VII and Figures 4 to 7.
- **(Q2):** The results shown in Figure 7 and Tables VII, VI answer the research question that the AM had more significant results than the TM in predicting final student grades, although the optimal number of topics was defined in each lesson to improve the prediction performance. Table VIII illustrates the differences between two methods: TM and AM.
- **(Q3):** The third research question concerns the correlation of attribute types and prediction results. From Figure 7 and Table VI, we assumed that understanding, cooperation and attitudes with students' grades had better results for lesson 5 than lesson 1 and the prediction performance for (a) Objective, (c) Understanding, (d) Attitudes and (e) Cooperation (acde) types was higher than the other types.
- **(Q4):** The results displayed in Figures 4 and 7 illustrate the significant advantage in predicting final student grade over the period of the semester. From lessons 5 to 12, the quality of writing comments were improved; students managed to express their attitudes and situations compared to lessons 1 to 4.

VII. Conclusions

TIs have powerful motivators and supports for students in the learning environment. They allow teachers to encourage different aspects of student outputs. In the current study, the teacher offered recommendations to maximize the positive impact toward writing comment data that help to understand students' performance and provide vital feedback to improve their performance. On the other hand, the AM method was proposed to build understandable prediction models by classifying comment data into six attribute types that reflect students' attitudes and situations toward the lesson. Experimental results showed that the AM had better prediction results than the TM did. In addition, we can conclude that the AM model improved the overall prediction results using (a) Objective, (c) Understanding, (d) Attitudes and (e) Cooperation (acde) types than the other types.

Although this study was conducted on the course of computer applications, it can be applied to other subjects. The results indicated that the introduced attributes were useful to identify and improve student performance. We believe that some of the attributes can be applied to other subjects. Finally, our experiment is considered as a basis for understanding students more deeply in the classroom and it will help to build an e-learning management system that gives automatic feedback for improving students' performance.

VIII. Acknowledgments

This work was supported in party by JSPS KAKENHI Grant No. 26540183, 26350357 and 16H02926.

References

[1] M. Chalaris, S. Gritzalis, M. Maragoudakis, C. Sgouropoulou, and A. Tsolakidis, "Improving quality of educational processes providing new knowledge using data mining techniques," *Procedia - Social and Behavioral Sciences*, vol. 147, pp. 390–397, 2014.

[2] L. Limbu and L. Markauskaite, "How do learners experience joint writing: University students' conceptions of online collaborative writing tasks and environments," *Computers & Education*, vol. 82, pp. 393–408, 2015.

[3] S. k. Natek and M. Zwilling, "Student data mining solution knowledge management system related to higher education institutions," *Expert Systems with Applications*, vol. 41, pp. 6400–6407, 2014.

[4] W. He, "Examining students online interaction in a live video streaming environment using data mining and text mining," *Computers in Human Behavior*, vol. 29, no. 1, pp. 90–102, 2013.

[5] S. Sorour, K. Goda, and T. Mine, "Correlation of topic model and student grades using comment data mining," in *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. ACM, 2015, pp. 441–446.

[6] S. E. Sorour, K. Goda, and T. Mine, "Evaluation of effectiveness of time-series comments by using machine learning techniques," *Journal of Information Processing (JIP)*, vol. 23, no. 6, pp. 784–794, 2015.

[7] X. Wanli, G. Rui, P. Eva, and G. Sean, "Participation-based student final performance prediction model through interpretable genetic programming: Integrating learning analytics, educational data mining and theory," *Computers in Human Behavior*, vol. 47, pp. 168–181, 2015.

[8] C. Romero, M.-I. Lpez, ose Mara Luna, and S. Ventura, "Predicting students final performance from participation in on-line discussion forums," *Computers & Education*, vol. 68, pp. 458–472, 2013.

[9] Y. Gong, J. E. Beck, and N. T. Heffernan, "Comparing knowledge tracing and performance factor analysis by using multiple model fitting procedures," in *Proceedings of the International Conference on Intelligent Tutoring Systems*. Springer, 2010, pp. 35–44.

[10] M. Khajah, R. M. Wing, R. V. Lindsey, and M. C. Mozer., "Integrating latent-factor and knowledge-tracing models to predict individual differences in learning," in *Proceedings of the 7th International Conference on Educational Data Mining (EDM)*, 2014, pp. 99–106.

[11] P. I. Pavlik Jr, H. Cen, and K. R. Koedinger, "Performance factors analysis–a new alternative to knowledge tracing." *Online Submission*, 2009.

[12] A. Zafra and S. Ventura, "Predicting student grades in learning management systems with multiple instance genetic programming," in *Proceedings of the 2th International Conference on Educational Data Mining (EDM)*, 2009, pp. 307–314.

[13] T. Minami and Y. Ohura, "A correlation analysis of student's attitude and outcome of lectures investigation of keywords in class-evaluation questionnaire," in *Advanced Science and Technology Letters*, vol. 73, 2014, pp. 11–16.

[14] O. N. Sunday, "Relationships between students' task engagement and learning outcomes in chemistry," *Journal of Educational Research and Reviews*, vol. 1, pp. 1–8, 2013.

[15] D. Cheung, "Students ' attitudes toward chemistry lessons: The interaction effect between grade level and gender," *Res Sci Educ*, pp. 75–91, 2007.

[16] B. Şen, E. Uçar, and D. Delen, "Predicting and analyzing secondary education placement-test scores: A data mining approach," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9468–9476, 2012.

[17] T. Minami and Y. Ohura, "How student's attitude influences on learning achievement? - an analysis of attitude-representing words appearing in looking-back evaluation texts," *International Journal of Database Theory and Application*, pp. 129–144, 2015.

[18] S. M. M. Rubiano and J. A. D. García, "Formulation of a predictive model for academic performance based on students' academic and demographic data," in *Proceedings of the 45th Annual Frontiers in Education Conference (FIE)*. IEEE, 2015, pp. 148–154.

[19] M. M. Ashenafi, G. Riccardi, and M. Ronchetti, "Predicting students' final exam scores from their course activities," in *Proceedings of the 45th Annual Frontiers in Education Conference (FIE)*. IEEE, 2015, pp. 372–380.

[20] D. Petkovic, M. Sosnick-Pérez, S. Huang, R. Todtenhoefer, K. Okada, S. Arora, R. Sreenivasen, L. Flores, and S. Dubey, "Setap: Software engineering teamwork assessment and prediction using machine learning," in *Proceedings of the 45th Annual Frontiers in Education Conference (FIE)*. IEEE, 2014, pp. 1299–1312.

[21] K. Goda and T. Mine, "Analysis of students' learning activities through quantifying time-series comments," in *Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2011), Part II, LNAI 6882, Springer-Verlag Berlin Heidelberg*, 2011, pp. 154 –164.

[22] K. Goda, S. Hirokawa, and T. Mine, "Correlation of grade prediction performance and validity of self-evaluation comments," in *Proceedings of the 14th annual ACM SIGITE conference on Information technology education*. ACM, 2013, pp. 35–42.

[23] J. Luo, S. E. Sorour, K. Goda, and T. Mine, "Predicting student grade based on free-style comments using word2vec and ann by considering prediction results obtained in consecutive lessons," in *Proceedings of the 8th International Conference on Educational Data Mining (EDM)*, 2015, pp. 396–399.

[24] Y. Goldberg and O. Levy, "word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method," *arXiv preprint arXiv:1402.3722*, 2014.

[25] S. E. Sorour, T. Mine, K. Goda, and S. Hirokawa, "Predicting students' grades based on free style comments data by artificial neural network," in *Proceedings of the 45th Annual Frontiers in Education Conference (FIE)*. IEEE, 2014, pp. 2475–2483.

[26] S. Sorour, K. Goda, and T. Mine, "Estimation of student performance by considering consecutive lessons," in *Proceedings of the 6th IIAI International Conference on e-Services and Knowledge Management (IIAI ESKM)*, 2015.

[27] S. Attle and B. Baker, "Cooperative learning in a competitive environment: Classroom applications," *International Journal of Teaching and Learning in Higher Education*, vol. 2007, Volume 19, Number 1, 77-83 http://www.isetl.org/ijtlhe/ ISSN 1812-9129, no. 1, 2007.

[28] R. B. Iyer, "Relation between cooperative learning and student achievement," *International Journal of Education and Information Studies*, vol. 3, no. 1, pp. 21–25, 2013.

[29] K. Higuchi, *KH Coder*, Retrieved 3/22/2016, from. [Online]. Available: http://khc.sourceforge.net/

[30] D. M. Ble, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[31] T. L. G. M. Steyvers, "Finding scientific topics," in *Proceedings of the National Academy of Sciences*, vol. 101, 2004, pp. 5228 – 5235.

[32] N. C. T. Hai, K.-I. Kim, and H.-R. Park, "Svd-lda: A combined model for text classification." *JIPS*, vol. 5, no. 1, pp. 5–10, 2009.

[33] M. Pontil and A. Verri, "Support vector machines for 3d object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 637–646, 1998.

[34] Y. Yang and X. Liu, "A re-examination of text categorization methods," in *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 1999, pp. 42–49.

[35] C. Metz, "Basic principles of roc analysis." *Semin Nucl Med*, vol. 8, no. 4, pp. 283–298, 1978.