Towards Improving Students' Attitudes to Lectures and Getting Higher Grades –With Analyzing the Usage of Keywords in Class-Evaluation Questionnaire–

Toshiro Minami

Kyushu Institute of Information Sciences
Dazaifu, Fukuoka Japan and
Kyushu University Library, Fukuoka Japan
Email: minami@kiis.ac.jp

Yoko Ohura

Kyushu Institute of Information Sciences
Dazaifu, Fukuoka Japan
Email: ohura@kiis.ac.jp

***Abstract*—The eventual goal of our study is to extract useful knowledge which will help students with improving their learning performance. Towards this goal, we have studied the methods for extracting useful information about students' attitudes to the lectures they take from the lecture-related data. Such studies will contribute to capture the real status of university students, and will be able to make a very useful tool for student development in the future. In this paper, we challenge the problem of outcome/grade estimation from the text data that have been written by the students in a term-end questionnaire. First, we introduce a new concept for rating a keyword which will contribute to increasing of the grades of the students who use them. Then, we use the contribution rates and estimate the grades of students. We change the weights and compare the estimated grades with the original ones, so that we can find the optimal weights of the keywords in the proposed framework. Finally, by using the optimal rates of keywords, we compare the usage of keywords between high-graded and low-graded students.**

Keywords–Text Mining; Weight of Word for Grade Estimation; Text Analysis; Educational Data Mining; Lecture Data Analysis.

I. INTRODUCTION

It has been pointed out in Japan that the students' academic skills and achievements have been decreasing. So, the universities have been putting a great amount of efforts in order to change the university professors to be able to do better lectures through the faculty development (FD) activities. However, students' academic skills do not improve accordingly.

Based on our observation, the biggest problem does not exist in the professors' teaching skills, nor in the students' academic skills and knowledge level. Main cause of this problem is rather on the students' attitudes to learning, such as eagerness to learn, curiosity to those that surround them, motivation to learning, and other mental tendencies. So, in order to pursue a solution to the problem of declining academic performances of students, it is insufficient to do efforts on FDs only. It is very important to take the matters of students (Student Development) into consideration.

Based on such recognition, our eventual goal of the study in this paper is to find out the most appropriate teaching/advising/leading methods for the students learn the most out of the lectures. In order to do this, we pursue the practical methods of extracting tips for improving lectures from data. We would help the students in learning more effectively by utilizing these tips. Our approach to this issue consists of two steps: (1) to make a student's learner model which includes attitudes to learning by proposing new concepts and measuring indexes for them, and see what we can find, and (2) to advise the student according to his or her learner model. This approach has an advantage in terms of understandability of humans. Even though the method we are developing is a naive one, we prefer to choose the understandable method rather than applying the established and more sophisticated methods if they are less understandable for us.

As a part of such an approach, we have been analyzing the answer texts to a question, which asked the students about their looking-back evaluation of themselves and the class [10]. Such data are considered to be appropriate to analyze the students' attitudes to the lectures. In the studies so far, we have found that the students with high examination scores use the words which indicate their wide point of view. On the other hand, the students with low grades use the words closely related to the lecture. An aim of this paper is to introduce a numerical method to text analysis and confirm this finding.

In this paper, we analyze the data obtained in the lectures. Such studies of educational data analysis have been conducted in the research field of Educational Data Mining (EDM) [12]. For example, Romero et al. [13] gave a comparative study of data mining algorithms for classifying students using data from e-learning system. Its major interest is on predicting the student's outcome. Our focus is on the student's psychological tendency in learning, such as eagerness, diligence, seriousness, etc. Many studies in EDM use the target data which are obtained from learning management systems (LMSx). On the other hand, we intend to obtain our target data in everyday lectures.

Goda et al. [2] proposed a method of text analysis, where texts are given by students as the reports in the everyday lectures. Our method, on the other hand, mainly uses such data as the homework, exercise, and term-end examination, which are obtainable in ordinary lectures.

Ames et al. [1] studied in the similar motivation to ours. They investigated the students' attitudes to the class, learning, etc., based on the answers to questionnaire items. However, their underlying data were obtained by asking the students to choose the rate from 1 to 5 for each question item. In our case, even though 2 of our question items are asking to rate from 0 to 100, other questions are asking to write the students' own thought in a free-text format.

Our data analysis style is also different from the major studies in EDM. Most of them somehow intend to analyze the big data, and the data obtained automatically as log data. On

 ISBN: 978-1-61208-386-5

the other hand, we would rather take the approach of dealing with small data, because our target data themselves may be very small [6][10]. Also, the data we deal with are somewhat represent human students, and we, as the staff in an educational organization like university, we have to educate all of them. Thus, we have to take attention to all the data as well, even if they are located in the far-away areas from the central area, because they represent one or more students.

We have been taking such an approach in library data analysis. In previous studies [4][7][8], we took library's loan records as the target data and analyzed them by proposing new concepts, such as expertise levels of books and library patrons. From these experiences we are convinced that such an approach is useful also for other types of small data. So, we take the same approach in our lecture data analysis.

In order to achieve our aim, the rest of the paper is organized as follows. In Section II, we describe the data we use for analysis. In Section III, we present our interesting findings in our previous studies, such as [10])[11]. In Section IV, we conduct the analysis by focusing attention to the words used by the students in the answer texts of a question about their final evaluation, which asked the students to evaluate their achievements of the class. We propose a process model for estimating the grades of students, and perform the process. Then, we change the parameter in order to find the optimal results. We discuss the usage of keywords in this optimal case. Finally in Section V, we conclude the discussions and findings in this paper.

II. Target Data

The data used in this paper came from the class of "Information Retrieval Exercise" in 2009 in a women's junior college [5][6][10]. The total number of students who attended the class was 35. They were year 2 students and going to graduate. The most important aim of the course was to let the students become expert information searchers so that they had enough knowledge about information retrieval, search, finding, and also had enough skills in finding appropriate search engine site and search keywords based on the understanding of the aim and the background of the retrieval.

The term-end examination of the course consists of 3 questions. The first question is to ask them to find the Web sites of search engine, and to summarize their characteristic features. The second question is on finding the Web sites on e-books and on-line material services. The third question is to find and argue about the information crimes in the Internet environment. The aim of these questions is to evaluate the skills on information retrieval, including the skills for planning and summarizing. These skills are supposed to have learned and trained in the course, through their exercises in the classes and in the homework assignments. We use the score of term-end examination as the measure for the student's achievement.

We also asked the students to answer the questions as the overall evaluation of them for the course. They are: (Q1) what the student has learned in the lectures, (Q2) good points of the lectures, (Q3) bad points that need to be improved, (Q4) to score the course as a whole with the numbers from 0 to 100, where the pass level is 60 as in the same way to the examination score, ..., (Q11) to score the student herself of her outcome and her attitude toward the course from 0 to 100 as in the same way as in Q4, etc.

III. Findings in Our Previous Analysis

We will illustrate what we have done and have found in our previous studies. Our study in Section IV is carried out based on these achievements.

A. Analysis of Numerical Items

As the first example of the results obtained in the previous analysis, Fig. 1 shows the correlation between the self-evaluation scores (x-axis) and the examination scores (y-axis). The former data are obtained as the answer to Q11. Actually, the data for about 60% of students are shown in the figure, because only 21 students gave answers to the scoring questions Q4 and Q11.

It is interesting to see that the students who have higher examination scores, i.e., those who locate above the line for the average score 71, evaluates themselves mostly in the range between 40 and 80. Thus, the range size of the self-evaluation is as wide as about 40.

On the other hand, the students who have lower-than-average scores evaluate themselves relatively high. With one exception of the student who has the lowest self-evaluation score, the rest of them mostly evaluate themselves more than average score.

In other words, we can say that the students who have high examination scores evaluate themselves from a very low score up to a very high one, which means that those students who evaluate low would have the self-image that "I am the person who can do better than what I have been doing." These students have a good desire of self-improvement.

On the other hand, the students who have poor performance seem to believe in themselves without evidence, and evaluate themselves something like, "I do fairly well in my study." Another possibility is that they actually recognize very well about their poor efforts and poor performance. Still, or maybe because of it, they wanted to believe themselves strongly, that they are not very poor in their efforts, instead of admitting their poor efforts. In this way, they could avoid facing what they really were, and keep their pride. As a result of such a phenomenon, the correlation coefficient between the self-evaluation score and the examination scores becomes a negative value of -0.1.

In our previous analysis, we have also investigated the relations of attendance and homework scores, together with the scores in the questions of Q4 and Q11. We have found that a notable number of students are just attending the lectures and

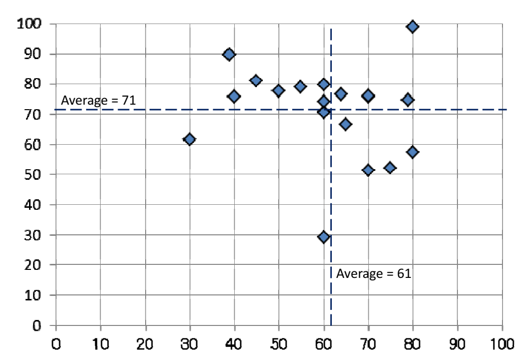

Figure 1. Correlation between self-evaluation scores (x-axis) and examination scores (y-axis).

do their homeworks. Probably because they do not intend to learn seriously, they could not learn in the real sense. Thus, quite a lot of students' efforts are rather superficial and they do not affect to the true improvement of their academic skills [9].

Furthermore, we have a result, which indicates that students concern more on their efforts than on their achievements, which is an important result for the lecturers to know. See the papers [5][6][9][10] for more findings in our analysis.

B. Analysis of the Relationship between Students' Viewpoint and Outcome from the Answers for Question 1

The first question of the term-end questionnaire was "Q1: What did you learn in this class? Did it help you?" It is the question to ask the students for summarizing what they have learned in the lectures of the course, and they express in their own words. Thus the answers to the question might express their understanding, recognition, point of view, and so on, in the course. Because the data are in free text format, it is more difficult to analyze and extract useful information out of them than the data in numerical format. However, at the same time, they are very appropriate to know about the students on their mental statuses that are normally hidden in their minds. We want to extract information such as which teaching materials attract them, in what attitudes they had in the class, how they felt by attending the class, etc.

As the first step to conduct the text data analysis, we need to transform them into the data that can be used in our analysis method. As an approach to the free text analysis, we took attention to the usage of words and phrases in the texts. The words used by the students might somewhat reflect their own views and attitudes to the course. Also, in order to obtain more subjective results, it is preferable to extract the words from the text data and analyze them than to extract the students' attitude data in a manual methods by us humans.

We chose the KH Coder [3] as the analysis tool. KH Coder is a free software equipped with the facilities of morphological analysis for Japanese language. This facility is very important in dealing with Japanese texts, because Japanese has no word divider, like the space in English. Thus word segmentation is a big issue in natural language processing. KH Coder can extract words together with doing statistical analysis including correspondence analysis. In our analysis, we took the answer text of each student as one document for KH Coder.

C. Extraction of Words which Appear in the Answers

Table I shows the words that appear in the texts more than 5 times and their number of occurrences, in the decreasing order of the number. We can see that the words related to the lectures appear in high frequencies. For example, the word "Search" appears 88 times in the answers to Q1, which is the most frequently used one among all words. Also the words "Information", "Library" and so on appear in the list. The lecture-related words are 6 (20%) among 30 words, whereas 4 (29%) among 14 words with frequencies more than 10.

D. Correspondence Analysis of Words and Students

It is important to know not only the words themselves but also their relations with others, such as between word and word, between word and student. Analysis of such association may give us more useful information about students and their attitudes to learning.

TABLE I. EXTRACTED WORDS AND THEIR OCCURRENCES (FREQ.> 5)

Word	Freq.	Word	Freq.	Word	Freq.
Search	88	Way	16	Think	8
Class	37	Examine	16	Do	8
Information	37	Keyword	13	Get	8
I think	34	Are various	11	Various	7
Library	33	Use ∗	10	Feel	7
Learn	32	Help	10	Function	7
Know	30	Necessary	9	Result	7
Myself	21	Use ∗	9	Important	7
How	21	Internet	8	Opportunity	6
Now	17	Personal Computer	8	This time	6

∗ Different words in Japanese

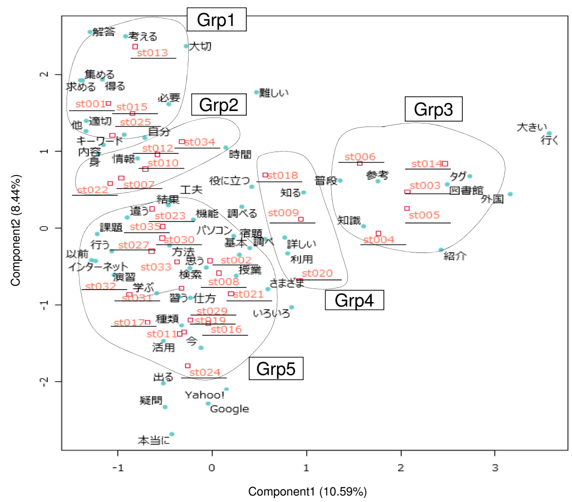

Figure 2. Correspondence analysis map of related words with students.

Fig. 2 shows the results of the correspondence analysis in a two-dimensional principle component space. The words in the figure are those occurring more than 2 times. The name in the form st0** represents a student. Then we divide the students into 5 groups manually, from Grp1 to Grp5. The total number of students who appear in the figure is 33 because two students did not answer to the question Q1. Table II shows the features of the groups, such as the numbers of students of the group, the member's average examination scores, and their variances. The numbers of the group members range from 3 (Grp4) to 16 (Grp5). For the average scores, Grp3 is the highest with the score 83.5, which locates in the upper-right part in Fig. 2. Grp5 takes the lowest average score with 59.3 which locates in the lower-left part. The P-value is 0.0469, and thus the assertion that there are differences between the averages of these examination scores of five groups is statistically significant at the 5% level [10].

E. Relation between Used Words and Examination Scores

Next, we deal with the correlations between the characteristic words that appear in the answers to Q1 and the examination scores. Table III shows the high ranked characteristic words of some of the students in the decreasing order in the Jaccard similarity measure. Note that the Jaccard similarity of student p and q is defined as the ratio of the number of words which

TABLE II. ANALYSIS OF VARIANCE TABLE OF 5 GROUPS.

Group No.	Number of Members	Average	Variance
Grp1	4	65.2	27.3
Grp2	5	70.5	98.8
Grp3	5	83.5	107.2
Grp4	3	69.8	68.7
Grp5	16	59.3	335.3

TABLE III. EXAMPLE CHARACTERISTIC WORDS OF STUDENTS IN GRP3 AND GRP5

Grp3

st005 99	st006 76	st003 77	st004 76	st014 90
Foreign●	Library▲	Put together	Show	Go
Library▲	Individuality	Country	Limit	Foreign country
Latest	Summary	Box	Interesting	Automatic
Effort	Take	HP	Photo	Especially
World	Relationship	Closed■	Introduction	Completely
See	Plus	Books■	Familiar	Lending■
IC■	Whole country	Appear	Japan	Electronic●
Tags■	At the same time	Root	Copyright●	Large
Various	Reference■	Tackle	Every time	Usually
Feel	Also	Province	Learn	Library▲

Grp5 (5 students out of 16)

st019 52	st030 27	st031 34	st032 29	st035 51
Old days	Result	Word	Internet ●	Draw
Question	Do	Since	Use of	Home
Really	Know	Approach	I	Work
Now	Information ▲	Motivation	Destination	Many
Think	Learn	Stimulus	Use	Get used to
Learn	Search ▲	Not at all	Job hunting	Listen
Respond		Different	Future	Touch
Answer		Frequency	Received	Utilize
Prior		Desk	Exercises	Vaguely
Current		A lot	Previous	Schoolchildren

are commonly used by p and q against the total number of words which are used either or both of p and q. The words marked with "●", "▲" and "■" in Table III indicate that they are classified as the general word, frequently used words that relate to the lecture, and the words closely related to the subject, respectively. The value in the right-hand side of (st0**) represents her examination score.

For Grp3 (with the highest examination score) characteristically use the technical terms and those words from the broader point of view in comparing Japan and the world such as "Foreign," "National," and "Japan." It is interesting to see that the words which are relating to the homework assignments do not appear in Grp3. Thus, we can see that the students in Grp3 attended the lectures with the attitude of learning in a broad perspective.

The students in Grp5 (with the lowest examination score) use quite a lot of frequently-used general words, and do not use technical terms at all. It is interesting to see that many students use the words they have learned in the lectures; e.g., "Learn," "Master," "Study," "Useful," and "Use." So, we can guess they took too much attentions to the words themselves and did not pay much attention to their background, their relation to the related concepts, their values in our society, etc.

IV. GRADE ESTIMATION OF STUDENTS BY THE USAGE OF WORDS IN THE QUESTIONNAIRE ANSWER

As we have shown in the previous section, there exists some amount of relationship between the grades of the term-end examination and the usage of words in the answers to Q1 [10]. For example, the students in Grp3, the group with the highest average score, used the words that indicate their interest to the world, such as "Foreign," "Overseas," "Japan," and "National." They also used more technical terms and characteristic words than other group members.

On the other hand, the students in Grp5, the one with the lowest average score, used general words and popular words, such as "Remember," "Learn," "Useful," and they did not use technical terms. It is interesting to see that from the words used by the Grp5 students, they look like good students because they used the words directly related to the lectures. However, considering their poor achievements, they might have learned rather superficially in the lectures. They might thought it was very important in the lectures to attend the class regularly and remembered what the teacher talked about. They would not think it was important to have a wide view on the subjects they were supposed to learn and thought hard with their own brains. We investigated further on how much the word types affect to the students' achievements in [11].

A. *Grade Estimation of Students*

Next, we estimate the students' grades according to the framework illustrated in Fig. 3. The process consists of two steps. Step 1 is the process of assigning the weights to the keywords that appear in the texts. The weight is calculated so that it reflects the closeness of the keyword and the students who use it, and the weight becomes higher if the students took higher examination scores. The arrows in the left part of the figure illustrate this relationship.

Step 2 is the process of estimating the grades of the students. We also use the relationship between the words and the student in this step. If a student used the words which have higher weights, then the student's estimated score becomes higher. The most typical and simple method might be to distribute the weight of a keyword to the students who have used it in the equal rate, i.e., to give the keyword weight equally to its related students by dividing the weight by the number of students who used it. Then, we calculate each student's estimated score by summing up all the score values

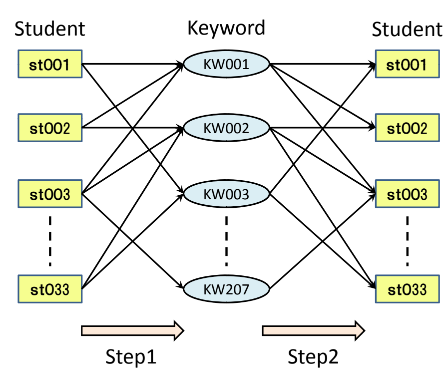

Figure 3. The method of the grade-estimation of students.

given by the words which the student used. We would like to make this simple definition more sophisticated.

Let $\{s_1, s_2, \ldots, s_m\}$ and $\{w_1, w_2, \ldots, w_n\}$ be the set of students and keywords, respectively, and let us define $A = (a_{i,j})$ so that $a_{i,j} = 1$ if and only if the word w_i is used in the texts provided by the student s_j, and $a_{i,j} = 0$ otherwise, for $1 \le i \le n$ and $1 \le j \le m$. For Step 1, we define a vector $S_1 = (g_1, g_2, \ldots, g_m)$ so as g_i is the examination score of the student s_i. We define a matrix $B = (b_{i,j})$ by setting $b_{i,j} = a_{i,j} / \sum_{1 \le k \le m} a_{i,k}$, so that the grade of a student is equally distributed to its related keywords, and $b_{i,j}$ is the distributed grade value to the word w_i from the student s_j.

Using this matrix we define the vector $W_1 = (c_i)$ of estimated grade value of keywords. Let k_i be set as the sum of the distributed grade values assigned to the keyword w_j; that is $W_1 = BS_1$, or:

$$c_i = \sum_{1 \le k \le j} b_{i,k} g_k \qquad \text{for } 1 \le i \le n \tag{1}$$

Table IV shows the list of words for the weight ≥ 0.0059 in its decreasing order. The weights here are normalized so that the total sum of the weights becomes 1. By using the weights of the words, we can estimate the grade of a student by summing up the weights of the words used by the student.

$$S_{E1} = A^T W_1 \times (t/u) \tag{2}$$

where t is the sum of the original grades, and u is the sum of the components of $A^T W_1$.

Table V shows the resulting estimated grades and their original grades of students in the decreasing order of the original grades. Fig. 4 shows the correlation diagram that shows the mutual relation between the original grades and the estimated grades of students.

TABLE IV. WEIGHTS OF WORDS. (RATE ≥ 0.0059)

No	Words	Rates	No	Words	Rates	No	Words	Rates
1	Newest	0.0071	8	Especially	0.0064	15	Layout	0.0059
2	Tag	0.0071	9	Feel	0.0064	16	Master	0.0059
3	World	0.0071	10	Library	0.0063	17	Recently	0.0059
4	Overseas	0.0068	11	Tackle	0.0063	18	Report	0.0059
5	Automatic	0.0064	12	See	0.0061	19	Server	0.0059
6	Big	0.0064	13	IC	0.0059	20	Start	0.0059
7	Electronic	0.0064	14	How to write	0.0059	21	Study	0.0059

TABLE V. ESTIMATED GRADES OF STUDENTS USING THE WEIGHTS OF WORDS.

No.	Student	Original	Estimated	No.	Student	Original	Estimated
1	st005	99	94	18	st013	69	49
2	st014	90	89	19	st017	66	63
3	st002	82	84	20	st018	66	66
4	st033	81	78	21	st020	64	60
5	st010	80	81	22	st025	61	66
6	st009	79	79	23	st001	60	66
7	st011	79	82	24	st016	57	62
8	st003	77	83	25	st023	56	51
9	st008	77	64	26	st029	55	63
10	st004	76	79	27	st034	54	60
11	st006	76	82	28	st027	53	18
12	st021	76	75	29	st019	52	53
13	st007	75	74	30	st035	51	57
14	st024	75	73	31	st031	34	44
15	st022	74	76	32	st032	29	43
16	st012	71	72	33	st030	27	29
17	st015	71	74				

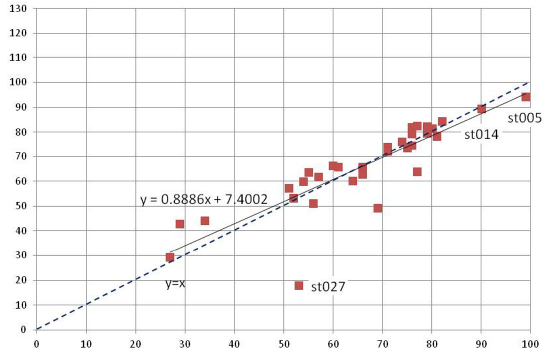

Figure 4. Correlation diagram between the original (x-axis) and the estimated (y-axis) grade.

B. Weight Adjustment for Optimal Estimation of the Students' Grades

The results obtained so far look good enough for estimating the students grades from the weights of keywords. However, the current weighting is proportional to the usage ratio of the word, and rather a simple method of weighting. Thus, it may be possible to get better result by taking different weighting method. We take an approach of introducing a parameter for changing the weighting, and find out the optimal value of the parameter for better weighting.

In order to pursue this scenario, we extend the definition of S_1 to the definition of S_n with the parameter n by: $S_n = (g_1^n, g_2^n, \ldots, g_m^n)$, so that the definition of S_1, which we have used so far, becomes a special case of S_n when $n = 1$.

Fig. 5 shows the change of correlation coefficient between the original and the estimated grades as n varies from 0 to 10. The optimal case in our framework comes when $n = 2.22$, and its correlation coefficient is 0.922.

C. Correlation between Original and Estimated Grades in the Optimal Case

Fig. 6 shows the correlation diagram of the grade-estimation of students when $n = 2.22$ and thus the correlation is optimal. Table VI shows the weights of keywords in this case. We can see that the keywords characteristically used by the students in Grp3 (with the highest average examination score) have high weight in general. For example, the word "Overseas" has the weight 0.0105, which is the 4th highest keyword weight. Other keywords "Japan" and "National" have 74th and 83rd highest among 206 keywords.

Figure 5. Change of correlation coefficient by the value of n.

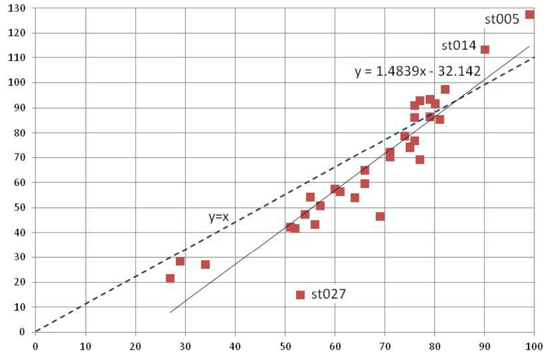

Figure 6. The optimal grade-estimation result (y-axis) represented as correlation diagram with the original grade (x-axis).

TABLE VI. WEIGHTS OF TOP MOST WORDS IN THE OPTIMAL CASE. (RATES ≥ 0.0069)

No	Words	Rates	No	Words	Rates	No	Words	Rates
1	Newest	0.0105	8	Especially	0.0085	15	Layout	0.0069
2	Tag	0.0105	9	Feel	0.0084	16	Master	0.0069
3	World	0.0105	10	Library	0.0083	17	Recently	0.0069
4	Overseas	0.0105	11	Tackle	0.0082	18	Report	0.0069
5	Automatic	0.0105	12	See	0.0075	19	Server	0.0069
6	Big	0.0085	13	IC	0.0074	20	Start	0.0069
7	Electronic	0.0085	14	How to write	0.0069	21	Study	0.0069

For the keywords preferably used by the Grp5 students, "Remember" has the ranking order 44, followed by "Useful," "Use," and "Learn" with the ranking order 100, 164, and 171, respectively. Thus, from the numerical data also, it has been clearly shown that the students in Grp3 used the words which have high weights, and those in Grp5 used the words which have small weight to the grades.

V. CONCLUSION AND FUTURE WORK

We have been studying the student's attitudes toward the lectures so far. In this paper, we conducted a study of estimating the grades of students in terms of the weights of words, which are calculated according to the frequencies of words used by students. Beginning from a simple method of linear weighting of words, we extended it to a more sophisticated method by using a parameter for adjusting the weighting. By changing the value of the parameter, we could find the value of the parameter which gives the optimal weighting of words for grade estimation. Then we discussed the words and their weights in the optimal case.

Our eventual goals in the research topic of this study are two-folds: The first one is to find new facts and tips for helping our students with more effective learning, and the other is to develop new concepts and measuring methods which can be used for the first goal. Thus, understandability is very important in our study. This is the reason why we rather choose naive methods of analysis than to use more sophisticated, but less human understandable methods.

Even though our current status of study is in a very beginning stage, the methods developed in the studies so far have shown high potential of our methods. It will become a necessary knowledge management tool for student development [7] in the near future, because it is a very important topic for the institutional research (IR) for universities [6].

We have the following study topics for the future: (1) To develop a method to devise new ideas further, and to perform refinement of dedication to the study of student effort, and attitudes to learning, especially further analysis of the text. Use of other similarity measure like pointwise mutual information (PMI) is a possible candidate. Also, it is worth comparing our model with other types of models. (2) By collecting data from a different class, to analyze them, and to verify if the results of this study are also holds. Also, it is important to find out the characteristic features of each class by comparing them. It will be interesting to investigate what features are gender-specific. (3) To generalize the analysis methods, and to integrate them into an automated data analysis system.

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 24500318, 2014.

REFERENCES

[1] C. Ames, and J. Archer, "Achievement Goals in the Classroom: Students' Learning Strategies and Motivation Processes," Journal of Educational Psychology, Vol.80, No.3, pp.260–267, 1988.

[2] K. Goda, S. Hirokawa, and T. Mine, "Automated Evaluation of Student Comments on Their Learning Behavior," 12th International Conference on Advances in Web-Based Learning (ICWL 2013), Oct. 2013, pp. 131–140.

[3] K. Higuchi, "KH Coder." `http://khc.sourceforge.net/en/` [accessed: 2014-12-07]

[4] T. Minami, and K. Baba, "Investigation of interest range and earnestness of library patrons from circulation records," International Conference on e-Services and Knowledge Management (ESKM 2012) in IIAI-AAI 2012, IEEE CPS, DOI 10.1109/IIAI-AAI2012.15, Sep. 2012, pp. 25–29.

[5] T. Minami, and Y. Ohura, "An attempt on effort-achievement analysis of lecture data for effective teaching," Database Theory and Application (DTA 2012), in T.-h. Kim et al. (Eds.): EL/DTA/UNESST 2012, CCIS 352, Springer-Verlag, Dec. 2012, pp. 50–57.

[6] T. Minami, and Y. Ohura, "Towards Development of Lecture Data Analysis Method and its Application to Improvement of Teaching," 2nd International Conference on Applied and Theoretical Information Systems Research (2ndATISR 2012), Dec. 2012, 14pp..

[7] T. Minami, "Profiling of Patrons' Interest Areas from Library's Circulation Records – An Approach to Knowledge Management for University Students -," The Fifth International Conference on Information, Process, and Knowledge Management (eKNOW 2013), Feb. 2013, pp. 45–50.

[8] T. Minami, "Interest Area Analysis of Person and Group Using Library's Circulation Records," IADIS International Conference Information Systems (IS 2013), March 2013, pp. 215–222.

[9] T. Minami, and Y. Ohura, "Lecture Data Analysis towards to Know How the Students' Attitudes Affect to their Evaluations," 8th International Conference on Information Technology and Applications (ICITA 2013), July 2013, pp. 164–169.

[10] T. Minami, and Y. Ohura, "Investigation of Students' Attitudes to Lectures with Tex-Analysis of Questionnaires," 4th International Conference on E-Service and Knowledge Management (ESKM 2013), Sep. 2013, 7pp..

[11] T. Minami, and Y. Ohura, "A Correlation Analysis of Student's Attitude and Outcome of Lectures –Investigation of Keywords in Class-Evaluation Questionnaire–," Advanced Science and Technology Letters (ASTL), Vol.73 (FGCN 2014), Dec. 2014, pp. 11–16.

[12] C. Romero, and S. Ventura, "Educational data mining: A survey from 1995 to 2005," Expert Systems with Applications, Vol. 33, Issue 1, July 2007, pp. 135–146.

[13] C. Romero, S. Ventura, P. Espejo, and C. Hervas, "Data mining algorithms to classify students," 1st International Conference on Educational Data Mining (EDM 2008), June 2008, pp. 8–17.