

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

# Discovering Novel Methods of Improving Middle Management Training through Text Mining

Kyoko Hayashi<sup>a,b,\*</sup>, Kazuhiko Tsuda<sup>a</sup>

<sup>a</sup>University of Tsukuba, 3-29-1 Otsuka Bunkyo-ku 112-0012, Tokyo, Japan

<sup>b</sup>GLOBIS University, 5-1 Niban-cho Chiyoda-ku 102-0084, Tokyo, Japan

---

## Abstract

As people enter the 100-year life period, they are working for longer periods of time, and acquiring new skills to stay up-to-date is becoming increasingly important. Moreover, educational institutions for working adults are conducting course satisfaction surveys to improve their educational programs; however, extracting useful information from the comments of several students each year is not easy. Therefore, in this study, a method was analysed to efficiently extract the targets of overall dissatisfaction and demands by conducting reputation and demand analyses using text mining based on the comments of a small group. As a result, we found knowledge that certain words can be used in a tool to efficiently extract relevant information.

© 2022 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

**Keywords:** Text Mining; Opinion; Satisfaction; Sentiment Analysis; Demand Expressions; Training Program; Human Resource Development

---

## 1. Introduction

As noted by Gratton and Scott [1], human life expectancy is increasing and people are working longer in the 100-year life expectancy era. In Japan, the population is ageing, and in 2021, laws were revised to extend the retirement age in the corporate sector. As people work longer, their business environment and roles change more frequently; moreover, the abilities required in the workplace change. Therefore, it is desirable for companies and employees to continue learning new skills and grow after entering the workforce; furthermore, education for working adults is becoming increasingly important.

---

\* Corresponding author. Tel.: +81-3-3942-6918

E-mail address: [kyoko.hayashi@globis.ac.jp](mailto:kyoko.hayashi@globis.ac.jp)

The educational institution where I work aims to continue providing high-quality educational programs. After each program, we conduct a survey with participants to determine their demands and satisfaction with the program, which we then use to improve the program. However, with numerous people enrolling in courses each year, considering that quantitative information can be processed efficiently and objectively, qualitative comment data tend to be interpreted and processed in an unorganised manner by teachers and staff members.

Therefore, this study aimed to explore methods to efficiently obtain information for improving educational programs. Using the text data of post-training questionnaires for managers, we propose a method to easily derive hints for program improvement by analysing the data and evaluate the proposed method by verifying the knowledge realised.

## 2. Previous studies

### 2.1. Text mining

Content analysis is used to analyse qualitative data, e.g. text, audio, and video. According to Higuchi [2], early attempts to contribute to the current content analysis were made in the 19th and early 20th centuries. However, qualitative data analysis was criticised as an impressionistic and non-scientific method. Subsequently, with the development of computerised natural language processing technology, quantitative text analysis methods were developed to ensure objectivity and reproducibility while retaining the advantages of qualitative research, and text mining research using large amounts of data became easier to conduct.

### 2.2. Research on satisfaction surveys

Several companies and organisations conduct satisfaction surveys with their customers after providing products. However, there has been extensive debate over how to interpret the quantitative results of satisfaction surveys; according to Oliver [3], satisfaction survey results tend to be biased toward higher scores. However, customer loyalty studies have shown that there is significant difference between the satisfaction of customers who exhibit the highest and next-highest scores. For example, according to Hestkett et al. [4], the repurchase rate of customers who gave Xerox's product a runner-up satisfaction rating of 4 was one-sixth that of customers who gave the highest satisfaction rating of 5. Takeya and Wataru [5] applied a similar mechanism to improve university lectures, thereby analysing and comparing the evaluation comments of the highest and second-highest groups in a course satisfaction survey; they found that the comments of the second-highest group included more information and led to better improvement.

### 2.3. Studies on polar expression, demand expression, and modality

Analysis to extract specific opinions from data containing people's opinions, e.g. survey comments, is called reputation analysis or sentiment analysis [6] [7]. Kanayama and Nasukawa [8] defined favourable and unfavourable opinions as 'polar expressions' and stated that extracting, classifying, and organising these opinions help provide useful information for the users and providers of the subject matter. Okada et al. [9] stated that polar expressions can be effectively extracted by focusing on sentence patterns based on expressions that express evaluation, e.g. adjectives and adjectival verbs, and functional expressions related to sentence structure, e.g. auxiliary verbs. However, in addition to information evaluation, questionnaires and related documents may include opinions regarding demands and demands, and Kanayama and Nasukawa [8] defined these opinions as 'demand expressions'. Demand expressions are not easily expressed in specific phrases, and in the case of Japanese, their intentions are often indicated by expressions at the end of sentences.

According to Inoue [10], the semantic content of a sentence comprises two parts: 'propositional content', which is the objective semantic content, and 'modality', which is the subjective semantic content. Modality expressions in Japanese include 1) predicate conjugations, 2) sentence adverbs (adverbs that correspond to the expressions that follow the sentence), 3) sentential verbs and interjections, and 4) intonation. Omori [11], after reviewing previous studies on modality, classified requirement attitudes into the following eight categories and provided examples of specific lexical expressions for each.

- A) Command: command, instruct, offer, impose, request
- B) Request: to ask, to request, to wish
- C) Prohibition: no, forbid, forbid, forbidden, not allowed, not permitted, forbidden, not allowed, prohibited, forbidden, strictly forbidden
- D) To invite: to invite to offer to do something
- E) Hope: to desire, wish, seek, expect, look forward to, pray for, want, desire
- F) Wishful non-decisional: may be
- G) Desirable, desirable, agreeable, welcoming, thankful for, pleased with, unwanted, necessary, important, better, appropriate, expedient
- H) Non-declaration: (the following succeeding words in the vocabulary of a demand) ~ would be, ~ may be, ~ seem to be

### 3. Proposed method

#### 3.1. Data to be used

The following data, from a post-training satisfaction questionnaire for participants (711 participants with an average age of 46) who participated in management training for section managers at a business school that provides management education for working adults from 2020 to 2009, will be used.

- (1) Quantitative data describing the level of satisfaction with training as a whole (the highest level of satisfaction was assigned a score of 5, and participants selected in the range of 1–5).
- (2) Text data to freely describe their impressions and requests regarding training as a whole.

The satisfaction levels from 1 to 5 and the ratios of the number of participants are listed in Table 1.

Table 1. Satisfaction and headcount ratio.

Satisfaction Level	Number of People	Percentage
5	441	62.0%
4	246	34.6%
3	21	3.0%
2	2	0.3%
1	1	0.1%
Total	711	100.0%

The average satisfaction rating for all participants was 4.58. As outlined in a previous study [3], the distribution of satisfaction was skewed toward the higher end of the scale.

#### 3.2. Research and analysis method

##### 3.2.1. Classification of data to be analysed

As outlined in a previous study [4], there may be a significant difference in attitude toward course satisfaction between the participants with a satisfaction rating of 5 and those with the next-highest satisfaction rating of 4. Moreover, it is expected that the comments of participants with a satisfaction level of 4 would include reasons why they did not provide the highest score, i.e. information that would contribute toward the improvement of the training program.

Although the total data of the participants who selected satisfaction levels 3–1 are small (3.38% of the total participants), negative reputations and demands manifest in their comments.

Moreover, there can be requests and suggestions for improvement in the comments of the participants who selected the highest satisfaction level of 5, although their satisfaction level is high. However, this group is the largest, i.e. 62.03% of the total participants, and it would be more efficient if a selection tool is considered to extract information from the data that will lead to improvement.

Therefore, for this survey, 711 participants were divided into three groups according to their attitudes toward course satisfaction: high- (Level 5), medium- (Level 4), and low-satisfaction groups (Levels 1 to 3). The data of the low-satisfaction group, which includes the smallest number of participants and is most likely to exhibit negative reputations and demands, were used as a sample for textual analysis. We will confirm whether the analysis of the other two groups as a leading tool will contribute toward efficiently searching for overall improvement requests.

The ratios of the number of people in the three groups are listed in Table 2.

Table 2. Ratio of the number of people in the three groups.

Group	Satisfaction Level	Number of People	Percentage
High-satisfaction	5	441	62.02%
Medium-satisfaction	4	246	34.60%
Low-satisfaction	1–3	24	3.38%
Total		711	100.00%

### 3.2.2. Analytical procedure

The flowchart illustrated in Fig. 1 shows the steps of the analysis in this study, and the details are described below.

KH Coder was used for text mining analysis.

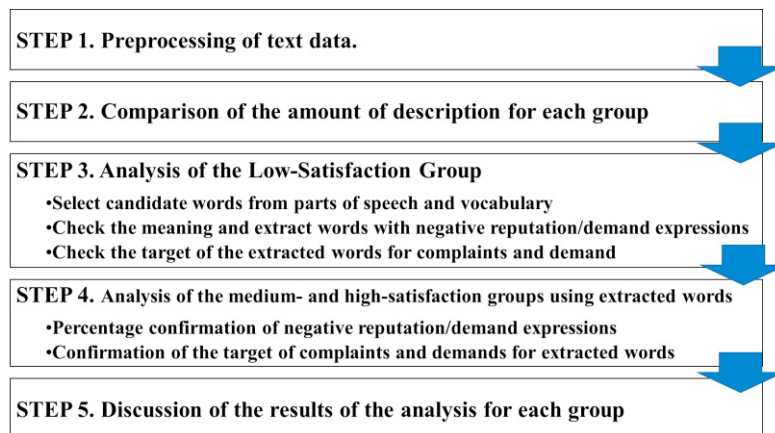


Fig. 1. Analytical procedure.

- STEP1. Pre-processing of the text data:

Using the morphological analysis system, MeCab, the morphological analysis of all text data is performed, and a list is created for each part of speech. Thereafter, the data are inspected, and any notational distortions are corrected to a unified description of the original data. Compound words are registered, such that they can be forcibly extracted by KH Coder.

- STEP2. Comparison of the amount of description for each of the three groups:

The number of comments written by each of the three groups is compared to check if there is any difference in the number of comments written based on the different satisfaction levels, which may interfere with the analysis results.

- STEP3. Analysis of the low-satisfaction group:

After extracting the survey comments from the low-satisfaction group and classifying them based on the part of speech using KH Coder, we narrowed them down to adjectives, adjectival verbs, and auxiliary verbs that are easily connected to reputation expressions and predicates (verbs: those in lexical expressions as exemplified by Omori [11], adjectives, adjectival verbs, and auxiliary verbs) and adverbs that are easily connected to demand expressions and created a list of candidate words. We then created a list of candidate words.

KH Coder's Key Word In Context function extracts 24 words before and after the above candidate words and checks whether they are used as negative reputation expressions or expressions of demand. Those with a probability of <50% are excluded, and a list of extracted words is created.

The target of each extracted word is checked and recorded.

- STEP4. Analysis of medium- and high-satisfaction groups using the extracted words:

Using the list of extracted words derived in STEP3, the text data of the medium- and high-satisfaction groups were analysed, and the percentage of words used as negative reputation/demand expressions was verified.

If the words are used as negative reputations or demands, the target of dissatisfaction or demand is identified and recorded.

- STEP5. Consider the results of the analysis for each of the three groups:

The words extracted from the low-satisfaction group are checked to see how similarly they were used in the text data of the medium- and high-satisfaction groups. It also compares what was the target of dissatisfaction and demands for each of the three groups. The results will be used to determine whether the use of the data of the low-satisfaction group as samples will contribute to deriving information that will lead to improvements based on the overall data. If a certain contribution is expected, we will derive knowledge on specific words to be used as keywords and specific procedures to be used to analyse the keywords for contributing toward the efficient improvement of educational programs.

## 4. Evaluation

### 4.1. Amount of information in the text data

After the pre-processing and morphological analysis of the data of 711 participants, the numbers of words and sentences appearing in each of the low-, medium-, and high-satisfaction groups are listed in Table 3.

Table 3. Number of people per group and amount of description.

Group	Satisfaction Level High	Satisfaction Level Medium	Satisfaction Level Low
Number of people (percentage)	441 (62.02%)	246 (34.60%)	24 (3.38%)
Total extracted word count	36,808 (63.71%)	19,014 (32.91%)	1,953 (3.38%)
Total number of sentences	1,328 (62.29%)	724 (33.96%)	80 (3.75%)

As the ratio of the number of people in each group and the ratio of the numbers of words and sentences that appear are almost equal, it was found that the number of comments written by each participant was almost the same among the three groups based on the satisfaction level.

#### 4.2. Text analysis of the low-satisfaction group

The text data of the low-satisfaction group were classified based on the part of speech and selected based on the part of speech and lexical expressions that are easily connected to reputation and demand expressions.

- Adjectives: many, good, difficult, no, heavy, few, short, old, painful, low, wanting, hard
- Adjectival noun: unfortunate, enormous
- Auxiliary verb (negative): no
- Adverbs: especially, more, still, a lot, a little more, a little

Of the aforementioned 22 words, 17 appeared as negative reputation or demand expressions with 100% probability, and Table 4 lists the number of occurrences for each object of dissatisfaction or demand.

Table 4. Analysis of the low-satisfaction groups (selected those with 100% negative reputation/demand expressions).

Reputation and demand content	Difficult	No	Few	Old	Painful	Low	Thin	Wanting	Hard	Unfortunate	Enormous	Especially	More	Still	A lot	A little more	A little	Total	Percentage
	Adjective	Adjective	Adjective	Adjective	Adjective	Adjective	Adjective	Adjective	Adjective	Adjectival noun	Adjectival noun	Adverb	Adverb	Adverb	Adverb	Adverb	Adverb		
Not enough time, too much volume, want to learn carefully	3	1					1			2	1	2	3	2	1	1	1	18	51.43%
Lack of teaching materials, difficult to understand, outdated				1														1	2.86%
Requests for program and lecture content	1	2	2			1		2		1		1	1		1			12	34.29%
Less reflection of advance assignments in lectures, burden												1						1	2.86%
Form of online lectures									1					1				2	5.71%
Time zone, day of the week					1													1	2.86%
Total	4	3	2	1	1	1	1	2	1	3	1	4	4	3	2	1	1	35	100.00%

The third-most frequently appearing word out of the 22 words was ‘good’. This adjective ‘good’ is used in a positive sense; however, in 55.56% of the comments, it was used with an adverb in the Japanese modality expression (e.g. ‘more is better’) to express a demand.

In the total count of 22 words, the most common subjects of dissatisfaction and demands were 33 occurrences (40.74%) with the meaning ‘less time, more volume, or more careful learning’, and 30 occurrences (37.04%) with the meaning ‘requests for the program and lecture content’, with the third-most common being 4 instances (4.94%) of ‘little reflection of pre-tasks in lectures, burden’.

This time, the analysis focused on words that were expressed as negative reputations or demands more than 50% of the time in the pre-screening process, and among these, the number of times that they did not demonstrate negative reputations or demands was 13, or 16.05% of the total occurrences. In other words, the probability that a word was used as a negative reputation or demand expression was 83.95%.

#### 4.3. Application to the medium-satisfaction group

Using the list of 22 words with a 50% or higher probability of being used as negative reputational expressions or demands, which was derived from the analysis of the low-satisfaction group, we analysed their usage in the survey comments of the medium-satisfaction group.

The results showed that 83.95% of the words in the low-satisfaction group were used as negative reputational expressions or demands; however, approximately half (49.74%) of the words in the medium-satisfaction group were used in the same sense.

Of these, the 11 words that were used more than 50% of the time as negative reputations or demands were difficult, few, old, thin, wanting, difficult, disappointing, more, still, a little more, and a little. Table 5 lists the 11 words and subjects of the demands.

Table 5. Analysis of the medium-satisfaction group (words that were negative reputation or demand expressions with a probability of 50% or more were selected).

Reputation and demand content	Difficult	Few	Old	Thin	Wanting	Hard	Unfortunate	More	Still	A little more	A little	Total	Percentage
	Adjective	Adjective	Adjective	Adjective	Adjective	Adjective	Adjectival noun	Adverb	Adverb	Adverb	Adverb		
Time is too little, volume is too much, I want to learn carefully	4	1			1	0	2	5	2	15		30	21.90%
Insufficient materials, difficult to understand, outdated			6	1	1	3		3		6	1	21	15.33%
Requests for lecture content	2							1		4		7	5.11%
Insufficient learning and interaction among members						2	3	2		3		10	7.30%
Less reflection of pre-subjects in lectures/burden		1					2	1		5		9	6.57%
Too many students		1								5		6	4.38%
Lecture environment							1	1		4		6	4.38%
Online lecture format	4					3	7		2	5		21	15.33%
Lecturers					2		2	3		2		9	6.57%
Opportunities to speak, points to be made	1						1					2	1.46%
Time zone										2		2	1.46%
Responsiveness of the office						1						1	0.73%
Not a negative reputation or request	8	3				1		1				13	9.49%
Total	19	6	6	1	4	10	18	17	4	51	1	137	100.00%

In the total count of 22 words, the most frequent negative reputation and expression of demands for possible targets for improvement was ‘Less time, too much volume, want to learn carefully’ (76 times, 13.36%), followed by ‘Lack of teaching materials, difficult to understand, old’ (47 times, 8.26%), and the third-most frequent was ‘online lecture format’ (40 times, 7.03%).

#### 4.4. Application to the high-satisfaction group

Using the list of 22 words with a probability of being used as negative reputations or demands of 50% or more, which was derived from the analysis of the low-satisfaction group, we analysed their usage in the survey comments of the high-satisfaction group.

The results showed that 83.95% of the words in the low-satisfaction group were used as negative reputations or demands, and 17.27% of the words in the high-satisfaction group exhibited the same meaning.

The seven words that were used more than 50% of the time as negative reputational expressions or demands were ‘hard’, ‘old’, ‘wanting’, ‘unfortunate’, ‘a little more’, ‘more’, and ‘still’. Table 6 lists the seven words and objects of dissatisfaction and demands.

In the total aggregate for the 22 words, the most frequent negative reputation or expression of desire expressed as a target for improvement was ‘online lecture format’ (40 times, 5.19%), followed by ‘little time, too much volume, want to learn carefully’ (34 times, 4.42%), and the third-most frequent was ‘lack of teaching materials, difficult to understand, old’ (18 times, 2.34%).

Table 6. Analysis of the high-satisfaction group (words that were negative reputation or demand expressions with a probability of 50% or more were selected).

	Old	Wanting	Hard	Unfortunate	More	Still	A little more	Total	Percentage
	Adjective	Adjective	Adjective	Adjectival noun	Adverb	Adverb	Adverb		
Reputation and demand content									
Time is too little, volume is too much, I want to learn carefully		1	1		4	1	8	15	20.27%
Lack of teaching materials, difficult to understand, outdated	2		1		1		5	9	12.16%
Requests for program and lecture content					1			1	1.35%
Insufficient learning and interaction among members				3	1		2	6	8.11%
Lack of reflection of pre-tasks in lectures, burden					1		1	2	2.70%
Lecture environment							1	1	1.35%
Online lecture format			2	7	1	5	1	16	21.62%
Lecturers		1						1	1.35%
Negative reputation or no request	1	1	3	2	7	3	6	23	31.08%
Total	3	3	7	12	16	9	24	74	100.00%

#### 4.5. Comparison of the three groups of words most likely to lead to negative reputational or demand expressions

Table 7 lists the use of 22 words in the survey comments of the low-satisfaction group, which are more than 50% likely to lead to negative reputational expressions or demands, in the medium- and high-satisfaction groups.

Table 7. Comparison of the extracted words in the three groups.

Word	Part of speech	Low-Satisfaction Group			Medium-Satisfaction Group			High-Satisfaction Group		
		Total number of occurrences	Number of occurrences in negative reputation/ demand expressions	Percentage of occurrences with negative reputation/ demand expressions	Total number of occurrences	Number of occurrences in negative reputation/ demand expressions	Percentage of occurrences with negative reputation/ demand expressions	Total number of occurrences	Number of occurrences in negative reputation/ demand expressions	Percentage of occurrences with negative reputation/ demand expressions
Many	adjective	12	9	75.00%	35	13	37.14%	72	4	5.56%
Good	adjective	9	5	55.56%	112	33	29.46%	168	19	11.31%
Difficult	adjective	4	4	100.00%	19	11	57.89%	13	3	23.08%
No	adjective	3	3	100.00%	58	17	29.31%	100	10	10.00%
Heavy	adjective	2	1	50.00%	0	0	N/A	0	0	N/A
Few	adjective	2	2	100.00%	6	3	50.00%	7	0	0.00%
Short	adjective	2	1	50.00%	8	3	37.50%	9	2	22.22%
Old	adjective	1	1	100.00%	6	6	100.00%	3	2	66.67%
Painfull	adjective	1	1	100.00%	0	0	N/A	0	0	N/A
Low	adjective	1	1	100.00%	1	0	0.00%	0	0	N/A
Thin	adjective	1	1	100.00%	1	1	100.00%	0	0	N/A
Wanting	adjective	2	2	100.00%	4	4	100.00%	3	2	66.67%
Hard	adjective	1	1	100.00%	10	9	90.00%	7	4	57.14%
Unfortunate	adjectival noun	3	3	100.00%	18	18	100.00%	12	10	83.33%
Enormous	adjectival noun	1	1	100.00%	0	0	N/A	1	0	0.00%
No	negative auxiliary verb	21	17	80.95%	192	85	44.27%	282	41	14.54%
Especially	adverb	4	4	100.00%	23	6	26.09%	43	3	6.98%
More	adverb	4	4	100.00%	17	16	94.12%	16	9	56.25%
Still	adverb	3	3	100.00%	4	4	100.00%	9	6	66.67%
A lot	adverb	2	2	100.00%	3	2	66.67%	1	0	0.00%
A little more	adverb	1	1	100.00%	51	51	100.00%	24	18	75.00%
A little	adverb	1	1	100.00%	1	1	100.00%	0	0	N/A

\*Shaded: 50% or more likely to be used as a negative reputational or demand expression

Of these 22 words, 11 words (50.00%) in the medium-satisfaction group and 7 words (31.82%) in the high-satisfaction group were used as similar expressions more than 50% of the time.



However, when the number of negative reputation/demand expressions used in the low-satisfaction group was reduced to 17 words (100%), 64.70% of the words in the medium-satisfaction group and 41.18% of the words in the high-satisfaction group exhibited a similar expression with a probability of 50% or more.

Furthermore, the seven words, ‘difficult’, ‘old’, ‘wanting’, ‘unfortunate’, ‘a little more’, ‘more’, and ‘still’, which appeared as negative reputations and demands more than 50% of the time in the high-satisfaction group, were found to exhibit >90% of the same rate of expression in the medium-satisfaction group.

#### 4.6. Comparison of the three groups of the target content of negative reputational and demand expressions

Table 8 lists a comparison between the subject of dissatisfaction and demands in each satisfaction group.

Table 8. Comparison of the extracted words in the three groups.

Reputation and request content	Low-Satisfaction Group		Medium-Satisfaction Group		High-Satisfaction Group	
	Appearances	Percentages	Appearances	Percentages	Appearances	Percentages
Not enough time, too much volume, want to learn carefully	33	40.74%	76	13.36%	34	4.42%
Lack of teaching materials, difficult to understand, outdated	1	1.23%	47	8.26%	18	2.34%
Requests for program and lecture content	30	37.04%	13	2.28%	9	1.17%
Insufficient learning and interaction among members	1	1.23%	20	3.51%	12	1.56%
Less reflection of pre-signups in lectures, burden	4	4.94%	27	4.75%	7	0.91%
Too many people	0	0.00%	11	1.93%	2	0.26%
Lecture environment	0	0.00%	8	1.41%	5	0.65%
Online lecture format	2	2.47%	40	7.03%	40	5.19%
Lecturers	0	0.00%	23	4.04%	4	0.52%
Opportunities to speak up, points to speak up	0	0.00%	9	1.58%	0	0.00%
Time and day of the week	1	1.23%	5	0.88%	0	0.00%
Support from the office	0	0.00%	4	0.70%	2	0.26%
Total	81	100.00%	569	100.00%	770	100.00%

\*Shaded: Top 3 for each group

As a result, in the medium- and high-satisfaction groups, the top three ranked items were the same: ‘Less time, more volume, and want to learn carefully’; ‘Lack of teaching materials, difficult to understand, and outdated’; and ‘Online lecture format’, although their rankings were different. In the low-satisfaction group, the first-ranked item, ‘Less time, more volume, want to learn carefully’, was consistent with the other two groups; however, the second-ranked item, ‘Desire for program/lecture content’, was ranked seventh in the medium-satisfaction group and fifth in the high-satisfaction group. The third-ranked item, ‘Little reflection/burden of prior assignments in lectures’, was ranked fourth by the medium-satisfaction group and fifth by the high-satisfaction group, indicating a slight difference in the awareness of the problem from the other two groups.

## 5. Results and discussion

### 5.1. Based on the textual analysis of the three groups

The 17 words used for the negative reputation/demand expressions with a 100% rate in the low-satisfaction group were found to be used as similar expressions in the medium-satisfaction group with a 64.71% rate (11 words) and in the high-satisfaction group with a 41.18% rate (7 words), for a probability of over 50%. Among the seven words that also appeared in the high-satisfaction group, the percentage of negative reputation/demand expressions in the medium-satisfaction group exceeded 90%.

Thus, using the keywords analysed and detected in the low-satisfaction group, which is a considerably small sample (3.38% of the total), we understood that, to some extent, the present analysis procedure will facilitate the discovery of information leading to the improvement of educational programs from the data of other groups.

However, it was found that the adjective ‘good’, which is generally used as a positive reputational expression, may lead to negative reputational expressions or demands when it appears in a sentence with adverbs such as ‘more’, ‘a little more’, and ‘better’, and that the meaning should be read carefully.

## 5.2. Subject of negative reputation and demand and satisfaction

The top three complaints and demands were the same for the medium- and high-satisfaction groups; however, the low-satisfaction group had only one complaint in common with the other two groups. The low-satisfaction group exhibited one item in common, indicating a difference in perception from the other two groups. As the low-satisfaction group was small in this survey (3.38% of the total), if the opinions of this group were to be reflected as is, there would be a risk of compromising the satisfaction of the majority of other participants. Thus, it was confirmed that using the text data of the medium-satisfaction group, which has the second-smallest number of participants, is appropriate while searching for events to be considered for the revision of the educational program; this is because it overlaps with the demands of the high-satisfaction group, which has the largest number of participants.

## 6. Conclusion

This study aimed to examine the data derived from post-training satisfaction questionnaires to efficiently retrieve information that will help improve educational programs.

We concluded that the information to be improved can be efficiently retrieved by analysing the data of the medium-satisfaction group using the following keywords: ‘difficult’, ‘old’, ‘want’, ‘disappointing’, ‘a little more’, ‘more’, and ‘still’; these keywords were extracted via the data analysis of a small part of the low-satisfaction group. The aforementioned methods were used to determine the side of the education provided and can be beneficial to education providers and the recipients of their services.

However, the aforementioned procedure might have limitations, as the analysis in this study was limited to only one questionnaire for the training of section chiefs. Further testing with other training programs is necessary to confirm the usefulness of this method.

Furthermore, as this study was focused on training satisfaction, it is not clear whether learner satisfaction is related to the aims and learning comprehension of the learning provider. To clarify this issue, it will be necessary to analyse the syllabus, textbooks, other teaching materials, and student-written reports, which will be the subject of future studies.

## References

- [1] Gratton, Lynda, and Andrew J. Scott. (2016) *The 100-Year Life: Living and Working in the Age of Longevity*, London, Bloomsbury Publishing
- [2] Higuchi, Koichi. (2014) *Quantitative Text Analysis for Social Researchers: A Contribution to Content Analysis*, Nakanishiya-Shuppan.Co., Ltd.
- [3] Oliver, Richard L. (2010) *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd ed., Armonk, M. E. Sharpe
- [4] Heskett, James L., Thomas O. Jones, Gary W. Loveman, W. Earl Sasser, Jr., and Leonard A. Schlesinger. (1994) “Putting the service-profit chain to work.” *Harvard Business Review* 72 (2): 164–174.
- [5] Taketani, Keigo, and Hironori Watari. (2015) “Analysis of course evaluation for improving on-demand academic writing class: text mining approach from the perspective of customer satisfaction analysis.” *Kyoto University Researches in Higher Education* 21: 1–14.
- [6] Liu, Bing. (2012) “Sentiment analysis and opinion mining.” *Synthesis Lectures on Human Language Technologies* 5 (1): 1–167.
- [7] Yi, Jeonghee, Tetsuya Nasukawa, Razvan Bunescu, and Wayne Niblack. (2003) “Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques.” *In Proceedings of the Third IEEE International Conference on Data Mining* 427–434.
- [8] Kanayama, Hiroshi, and Tetsuya Nasukawa. (2005) “Extraction and organization of demand expressions.” *11th Annual Conference Proceedings of The Association for Natural Language Processing* 660–663.
- [9] Okada, Makoto, Kazuhiro Takeuchi, and Kiyota Hashimoto. (2015) “An investigation of effectiveness of estimation string pattern to extract sentiment information from customer reviews of tourists.” *The 29th Annual Conference of the Japanese Society for Artificial Intelligence* 1–3.
- [10] Inoue, Masaru. (2006) *Chapter 4: Modality, Dialect Grammar, Dialectology Series 2*, Tokyo, Iwanami Shoten, Publishers, pp. 137–178
- [11] Ohmori, Akira. (2009) “A Linguistic Foundation for Requirements Extraction: A Definition of the Concept of Requirement, and Mental Attitudes of Requirements.” *In Proceeding of the 8th Forum on Information Technology* 167–174.