DOI 10.17939/hushss2020.63.2.004

Using Sentiment Analysis to Analyze the Feedback of Students with Open-ended Questions†

Leem, Byunghak*; Kang, Hyunju**;Eum,Seongwon***; Srijan Kumar****

Abstract

This study measures students' preferences by analyzing their feedback about the lectures they have been attending based on open-ended questions on the course evaluation using the sentiment analysis method, one of the text mining methods. Based on this analysis, the quality of lectures was measured to suggest future changes made for improving teaching methods, directions for improving lecture quality, and its contents. This study aims to analyze qualitatively the feedback of student evaluation to examine student's perceptions about the lecture more directly and comprehensively. To this end, answers of open-ended questions were collected and classified with positive, neutral, negative opinions or polarity. As a result of sentiment analysis 422 words are extracted. Among them 327 words (82.2%) were positive as opposed to 44. To measure our model's accuracy, recall and precision and all of them are over 0.96 which indicates that machine learning model using NBC method successfully classified the data. This study could contribute to the course evaluation because it shows how to utilize students' feedback of open-ended questions using text mining and sentiment analysis. It also gives solid measurement of the subjective point of view of students.

Keywords: *Text Mining, Sentiment Analysis, Student Feedback, Course Evaluation, open-ended question*

---

† This work was supported in 2016 by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NNRF-2016S1A5B6913773)

* E-mail : bhleem@gmail.com
Ph.D. earned at: University of Texas at Arlington
Current position: Busan University of Foreign Studies
** E-mail : hyenhj@gmail.com
Ph.D. earned at: Korea University
Current position: Honam University
*** E-mail : sweum@bufs.ac.kr
Ph.D. earned at: Pusan University
Current position: Busan University of Foreign Studies
**** E-mail : srijanjnu@gmail.com
Ph.D. earned at: Kyunghee University
Current position: Busan University of Foreign Studies

Received : April 30. 2020, Revised : May 22, 2020, Accepted : May 25, 2020

감성 분석을 활용한 학생의 개방형질문 강의평가 피드백 분석

임병학, 강현주, 엄성원, 스리잔 꾸마르

국문요약

본 연구는 강의 평가에 대한 학생의 피드백을 분석하여 학생들의 선호도를 측정하였다. 학생들의 피드백은 강의평가에서 개방형 질문에 대한 답으로 구성되었으며, 분석은 텍스트 마이닝 방법 중 하나인 감성분석을 사용하였다. 이 분석을 기반으로, 본 연구는 강의에 대한 품질을 정성적으로 측정하여 강의 내용, 강의 방법, 강의 품질 등의 개선 방향을 제시하였다. 따라서 본 연구의 목적은 학생의 강의 평가 피드백을 질적으로 분석하여 강의에 대한 학생의 강의 인식을 보다 직접적이고 포괄적으로 탐구하고자 하였다. 이를 위해 개방형 질문에 대한 답변을 수집하고 긍정, 중립, 부정 극성으로 분류했다. 감성 분석 결과 422개의 문서가 추출되었고, 그 중 82.23% 인 327개의 문서가 긍정 극성을, 44개 문서가 부정 극성으로 나타났다. 기계 학습기반 나이브 베이지안 분류(NBC) 방법을 기반으로 한 본 연구의 극성 분류 모델은 정확도, 재현도 및 정밀도 측정에서 모두 0.96이상으로 성공적인 분류임을 보이고 있다. 본 연구는 텍스트 마이닝 및 감성 분석을 사용하여 개방형 질문에 대한 학생들의 피드백을 활용하는 방법을 보여줌으로서 오늘날 정량적인 강의평가를 보완하는 것으로 기여할 수 있다. 또한 학생들의 개방형 질문 강의 평가 분석을 정량적이고 정성적인 방법인 혼합방법을 제시하였다.

주제어: *텍스트 마이닝, 감성 분석, 학생 피드백, 강의 평가, 개방형 질문*

I. Introduction

Most educational institutes conduct course evaluations in which students are supposed to answer both objective and subjective types of questions. Thus, the course evaluation produces quantitative and qualitative data but only quantitative data is reported and used for improvement of the course because it is convenient. However, we should listen to the voice of students as verbalized in open-ended questions. In small classes, it is easy to reflect students' responses, but in the case of large classes, it is difficult to reflect students' opinions. Therefore, we look into students' feedback using text mining and sentiment analysis tools which enables us to examine students' perceptions about the lecture more directly and comprehensively.

To this end, the study is organized in the following procedures. Section II reviews related studies to derive the implications that could contribute to this study. Section III describes theoretical background of this study, which is text mining and sentiment analysis is presented. Then, Section IV explains how this study collects the data. Section V offers a conclusion summarizing the results of the study and discusses the study's limitations. Our purpose is to analyze students' feedback on open-ended questions; we do not discuss course evaluation scores.

Ⅱ. Literature Reviews

There is large amount of literature addressing the use and analysis of students' feedback instruments as a measure of teaching performance in higher education throughout the world. Students' course evaluation instruments most frequently consist of a combination of Likert scale responses to statistically valid questions and one or more open-ended questions (Denson, et al., 2010). Responses to Likert scale questions provide quantitative data that can be conveniently used for analysis of evaluation results. However, lecture evaluation items are limited and there is a possibility that various students' opinions cannot be collected. In order to improve class quality through lecture evaluation, it is necessary to collect opinions of various students in detail through descriptive course evaluation feedbacks in addition to quantitative evaluation results (Lee, et al., 2018).

For this reason, research has recently been conducted on how to collect and analyze text-based feedback written in students' class evaluation. We have reviewed relative articles on course evaluation using text mining tools.

Students' responses to open-ended questions are difficult to analyze because they are not directional. In order to make up for these shortcomings, the previous study looked at the characteristics of students' feedback on specific topics. Lee and Nam (2018) classified the top 30% of subjects by using the total scores of lecture evaluations, and analyzed the characteristics and patterns of good lectures for each major by using the narrative questions of lecture evaluations for each semester for 10 years. Choi and Ahn (2016) extracted the characteristics of lectures related to learner interaction through text mining, and compared the scores of these keywords with the evaluation scores obtained from the selective evaluation.

Shin and Choi (2019) analyzed the content most frequently mentioned by students in the subjective course evaluation using the text mining tools, and analyzed the characteristics of the content of the narrative course evaluation depending on the students and subjects. They confirmed that students' descriptive responses can be used as meaningful data for teaching improvement. Shin (2019) also studied the similar topic with similar tools and showed that

the ratio of noise data such as non-sentence, one letter, and no opinion was 10.6% of the total number of evaluations, and the percentage of short sentences of less than 20 characters was 75.6%. Also, there was a difference in the volume of responses depending on the characteristics of the students and lectures. Finally, the results of the study showed that the longer the sentence, the lower the score of the course evaluation.

The examples in Table 1 show that a method of clustering text was used to analyze the results of student course evaluations, while there have also been studies using sentiment analysis methods to investigate students' preferences. Jimenez, at al. (2019) compares the sentiment classification of student feedback questions at the sentence level and at token level for different classifiers.

Table 1 . Summary of the Research

Author	Approach	Industry	Data Source
Kwak, Min, Kim (2019)	Topic Modeling (Descriptive Lecture Evaluation)	Education	About 1,500 courses in the first semester of 2015
Lee and Nam (2018)	Text Mining	University	Additions from 47,177 courses from 2005 to 2014
Choi and Ahn (2016)	Text Mining and Opinion Mining	University	Descriptive lecture by two professors of the same subject
Jimenez, et al (2019)	Sentiment Analysis	University	From mid-2005 to the first semester of 2018, with a total of around 170,000 comments
Shin and Choi (2019)	Text Mining	University	For 9 semesters, starting from the first semester of 2014 to first semester of 2018 at a university in the metropolitan area.
Park (2019)	Text Network Analysis	University	Evaluation of Software Education Lectures from 2017 to 2018 at university
Shin (2019)	Text Mining	University	A total of 9 semesters starting from first semester of 2014 to first semester of 2018 at a university

Ⅲ. Theoretical Background

1. Text Mining

We are living in a world full of structured and unstructured data. As seen in Fig.1 unstructured data is database notes, e-mails, online reviews, social media contents, and so on. In fact, unstructured data makes up 80% of all data, and is growing at the rate of 55% and 65% per year. Business insights are hidden into these unstructured data. Therefore, we need a tool to analyse those unstructured data for predictive and prescriptive analysis. Text mining is a technique used to extract, analyze, and interpret hidden business insights from unstructured

textual elements of social media contents or online reviews. Generally, text analytics could be conducted by source extraction, text pre-processing, data transformation, and then text mining as shown in Fig 1.

The text analytics process starts with identifying the source of the text to be analyzed. Texts are dynamic, vast, diverse, multilingual, and noisy. Thus, finding the right source for text analytics is crucial point to gain practical implications. The next step is to parse, clean, filter the text, and create a dictionary of words using natural language processing (NLP), which is mostly based on machine learning techniques. In order to extract meanings from the text, the sentence structure and parts of speech are determined, stop words are removed, and spellings are checked. Before the application of analytical algorithms to any text, it should be transformed into computer-readable format for analysis using linear algebra-based technique and vector space models (VSM). The last step is to extract the needed business insights, using text mining algorithms, such as clustering, classification, association, and semantic analysis. In this step, we choose sentiment analysis.

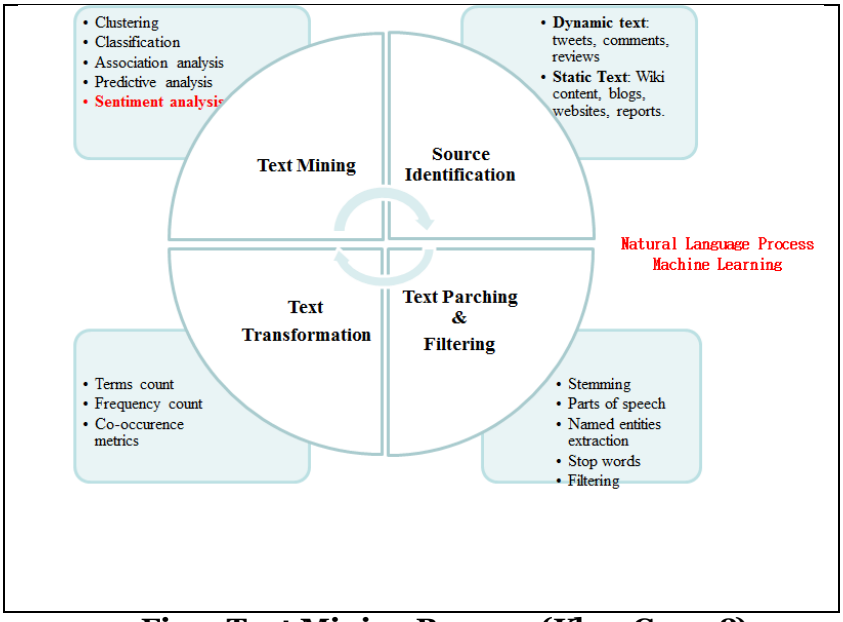


Fig 1. Text Mining Process (Khan,G.,2018)

2. Sentiment Analysis

Sentiment analysis is the automated process of determining whether a text expresses a positive, negative, or neutral opinion about a product or a topic. By using sentiment analysis, companies do not have to spend endless hours tagging customer data such as survey responses, reviews, support tickets, and social media comments. Sentiment analysis helps companies monitor their brand reputation on social media, gain insights from customer feedback, and much more. Sentiment analysis is to define automatic tools able to extract subjective information from texts in natural language, such as opinions and sentiments, so as to create structured and actionable knowledge to be used by either a decision support system or a decision maker. Sentiment analysis classifies text into objectivity/subjectivity and polarity like positive, negative, or neutral at the aspect, sentence, and document level as shown in Figure 3.

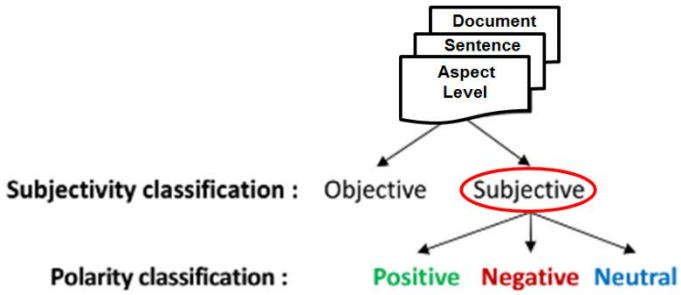


Fig 2. Sentiment Analysis

After classifying subjectivity and polarity in text, its results can be presented in different types of visual reports. Sentiment analysis has two kinds of approaches like machine learning and lexicon-based approach. In machine learning approach, there are two classes of techniques: supervised and unsupervised. In Lexicon-based approach, there are dictionary-based approach and corpus-based approach. In this study we analysed sentiment by using supervised learning and Dictionary-based approach.

Ⅳ. Research Method

Fig.3 provides the whole procedure of this study. First, we collected text data using open-ended questions from students in classes, and pre-processed the data using Netminer 4.0. Then, the sentiment analysis was conducted using Textom, and the results of analysis were reported in the end with our suggestions and limitations of this study. Here, netminer 4.0 is specialized in SNA (social network analysis) and text mining, and Textom also is specialized in text mining and sentiment analysis.

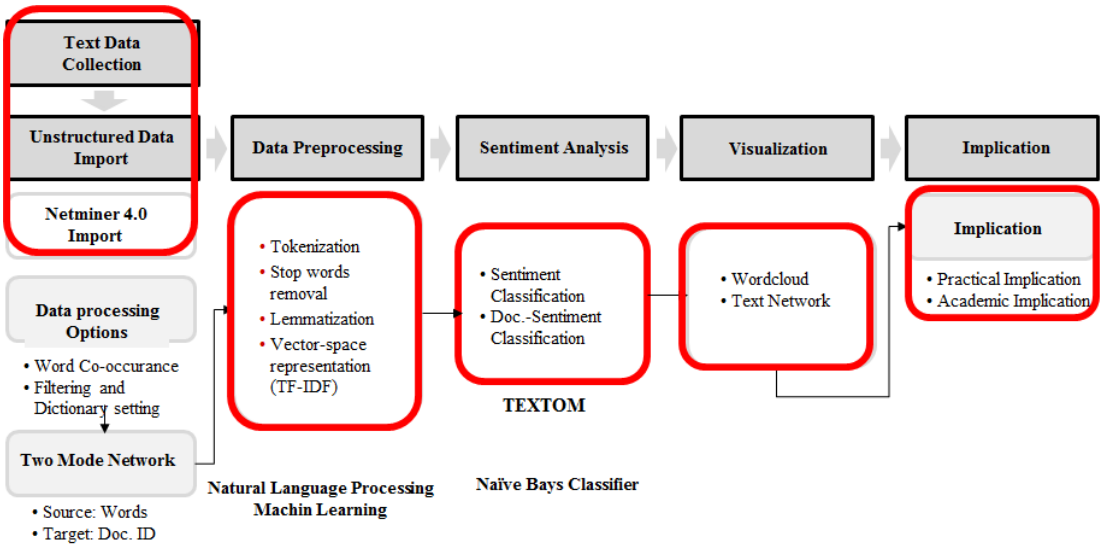


Fig 3. Research Procedure

Specifically, we collected course evaluation data from 9 classes from Nov. 25 to Dec. 2, 2019. The number of respondents were 258 students that made 422 documents for open-ended questions. For data cleaning, we imported the collected data into Netminer 4.0.

Then, we pre-processed the data. First, documents were parsed and the data was chopped up into pieces called tokens. The POS (parts of speech) were tagged to each token and some stop words which would appear to be of little value in helping select documents matching a user need were excluded. All steps were processed by using Netminer 4.0. As a result of data pre-processing, we collected 703 words which were comprised of 490 nouns, 56 adjectives and 157 verbs.

Using these 703 words, we created the words cloud which is made up with high-frequency words as shown in Fig. 4. Big word means it has high frequency. And then, we analyzed text network by clustering co-occurrence words and found four major topics such as teaching environment, content, and teaching.

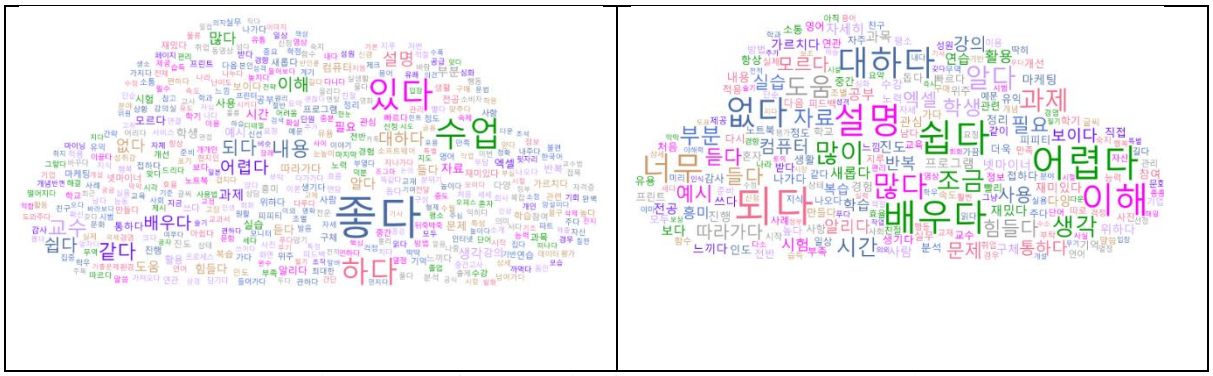


Fig 4. Before and After Text Preprocessing

Ⅴ. Results

V-a. Evaluating Accuracy: Precision, Recall and F-measure

To measure our model, we calculated the useful metrics (Powers 2011) of accuracy, recall and precision. All were over 0.96, which indicates that machine learning model using NBC method successfully classified the data. Accuracy = (TP + TN)/(TP + TN + FP + FN) = 0.96, Precision = TP /(TP + FP) = 0.98, Recall = TP /(TP + FN) = 0.98, F-value=2*(Precesion*Recall)/ (Precesion+Recall)=0.98.

Table 2. Sentiment Confusion Matrix

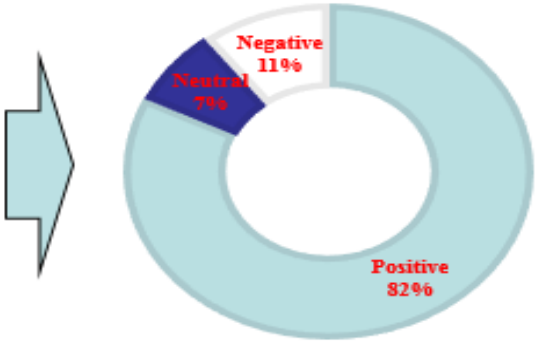
	Positive	**Negative**
Positive	**TP (285)**	**FP (7)**
Negative	**FN (6)**	**TN (38)**

TP: True Positive, FP: False Positive, FN: False Negative, TN: True Positive

V-b. Sentiment Analysis

Finally, we performed sentiment analysis using Textom with supervised machine learning and Naïve Bays Classifier (NBC). We built up 170 learning data that we then imported into Textom. In the resulting sentimental analysis 422 texts were classified, of which 327 (82.2%) were positive, and 44 negative.

	Frequency	%
Total	422	100
Positive	347	82.23
Neural	31	7.35
Negative	44	10.43



Naïve Bays Classifier (NBC)

Figure 5. Textom Sentiment Analysis Results

Table below is the report of sentiment analysis by keywords using sentiment dictionary on Textom. Positive words frequency represents 64.23%, but negative 35.77%. As a result of more careful analysis of this result, the words related to good feeling, interest, and joy were frequently mentioned in the positive case, and words related to disgust, sadness, and anger were mentioned in the negative case as shown below.

Table 3. Polarity Classification

	keywords	Frequency	Percent
Positive (64.23)	Interest	52	10.28
	Good Feeling	260	51.38
	Joy	13	2.57
Negative (35.77)	Pain	3	0.59
	Sadness	80	15.81
	Anger	9	1.78
	Fear	7	1.38
	Disgust	82	16.21

V-c. Sentiment Score of Each Class

Finally, we computed the sentiment score for each class based on the sentiment analysis using Si = (Npi-Nni)/((Npi+Nni) *100, where i: Class Name, Npi: # of positive Doc. in Class i, Nni: # of negative Doc. in Class i . This result can be used to complement current quantitative evaluation method. One interesting result was that classes using computers had a low score. That means that the class teacher should improve teaching environment, method, or contents. In the case of subjects with a small number of participating students, such as Korean language classes, the results are considered to be overestimated.

Table4. Sentiment Score

Course	Positive	Negative	Neutral	Score*
Consumer Behavior	15	2	1	76
Excel A	56	16	4	56
Excel B	59	8	1	76
Supply Chain Management	44	6	14	76
Social Media Marketing	31	8	2	59
Service Operations Management	41	1	4	95
Indian Culture	66	2	3	94
Korean	10	0	1	100
Hindi	24	1	1	92

Ⅵ. Conclusion and Limitations

In this study, open-ended question responses of students to course evaluations were analyzed using text mining tools. Among the text mining tools, this study selected a method of registering and analyzing learners' responses in the sentiment dictionary through sentiment analysis method that could then be converted into scores**.**

This study can contribute to the course evaluation because it shows how to utilize students' feedback of open-ended questions using text mining and sentiment analysis. It also gives solid measurement in the subjective point of view of students. However, the number of students participating in this study was not abundant. And we need to study more on cynical and paradoxical expressions such as 'I wish that~(*eumyon jokette* in Korean).

References

Choi,Jungwoong and Ahn,Dongkyu (2016), A Study on Descriptive Lecture Evaluation Using Data Analysis, Journal of Digital Convergence, Vol 14,11.

Denson, N., Loveday, T., & Dalton, H. (2010). Student evaluation of courses: what predicts satisfaction?. Higher Education Research & Development, 29(4), 339-356.

Jiménez, H. G., Casanova, M. A., Nunes, B. P., & Do Couto, A. C. N. F. (2019). Course Observatory: Sentiment Analysis of Comments in Course Surveys. In 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT) (Vol. 2161, pp. 176-178). IEEE.

Kwak, Minho., Min,Hyeree. and Kim,Meereem. (2019) Analysis of Students' Open-Ended Course Evaluation Using Topic Modeling, Asian journal of education, vol 20,2.

Lee,Haedeum and Nam,Minwoo, (2018). Analysis of Good Class Characteristics by Major of Department through Text Mining of Results of University Lecture Evaluation, Korean journal of early childhood education, vol 20,2

Park,Geumju.(2019), The Study on the Software Educational Needs by Applying Text Content Analysis Method: The Case of the A University, Korean Academic Industrial Society,Vol20,3

Powers, D.M.W., (2011). Evaluation: from precision, recall and f-measure to ROC, informedness, markedness and correlation. J. Mach. Learn. Technol. 2 (1), 37–63.

Shin,Jongho(2019), Analysis of Student Response status and Characters of Descriptive Course Evaluation – Focus on A University Case, CNU Journal of Educational Studies.

Shin,Jongho. and Choi,Jaewon.(2019). Text mining Analysis of College Students' Descriptive Course Evaluation, Learner-centered curriculum education, Vol 19.16