



HAL
open science

Exploring the Students Feelings and Emotion Towards Online Teaching: Sentimental Analysis Approach

T. Praveenkumar, A. Manorselvi, K. Soundarapandiyan

► **To cite this version:**

T. Praveenkumar, A. Manorselvi, K. Soundarapandiyan. Exploring the Students Feelings and Emotion Towards Online Teaching: Sentimental Analysis Approach. International Working Conference on Transfer and Diffusion of IT (TDIT), Dec 2020, Tiruchirappalli, India. pp.137-146, 10.1007/978-3-030-64849-7_13 . hal-03701797

HAL Id: hal-03701797

<https://inria.hal.science/hal-03701797v1>

Submitted on 22 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Exploring the Students Feelings and Emotion towards Online Teaching: Sentimental Analysis Approach

Dr.T. PraveenKumar¹[0000-0002-1706-5674], Dr.A. Manorselvi² and

Dr.K. Soundarapandiyan³

¹Ramaiah Institute of Management, Bangalore, India

praveen@msrim.org

²St.Peters Institute of Higher Education and Research, Chennai, India

³Sri Sai Ram Engineering College, Chennai, India

Abstract

Data mining is a method to refine raw data to useful information. In education, data mining is a significant research part used to progress the value of education by observing students' performance and understanding their learning patterns. Real-time student feedback would empower faculty and students to comprehend the teaching and learning problems in the most user-friendly way for the students. This paper uses a Lexicon based sentimental analysis technique to analyze students' feelings and emotions through their feedback by correlating learning analytics to grounded theory. The sentiment analysis technique is a computational process to identify and classify subjective information such as positive, negative, and neutral from the source material. It can extract feelings and emotions from a piece of a sentence. Hence this paper aims to recognize the students' positive or negative feelings and distinguished emotions, towards online teaching. The methodology undertakes four processes. The first process is data extraction from the feedback collected from the students through open-ended questions (Text) and is used as source material and imported to R studio. The second process is data cleaning /data preprocessing, removal of annoying data, and separation of data. The third process is sentimental analysis, which divides the data into positive, negative, and neutral categories/groups. This lexicon-based method of sentimental analysis is used to classify the sentiments. The results were estimated using sentiment scores and emotional variance. The sentiment scores result found that students have positive sentiments/emotions towards online teaching and emotions vary concerning the online class timing.

Keywords: Learning analytics, Sentimental analysis, Text mining, Students Emotion and Learning

1 Introduction

Students Feedback is an essential component in the global education system, especially more so for the Indian higher education system. The apex accreditation agencies of the country like NBA, NAAC, etc. have included the students' feedback as one of the measuring criteria for assessing the quality of the teaching-learning process, students' feedback supports progresses in teaching (Ann Poulos et.al 2008). Students may have different issues concerning their class lectures (S. Cummins et.al 2012). Evidence of this is once a student does not know the content of the lecture or an exact case, the Students' feedback can explore what students understand and what they do not understand; what they like and dislike about the class. Moreover, students' feedback can support the faculty to realize the students' learning behavior (Toon Calders et.al 2012). Collecting the feedback is usually taken after a semester or course (Jeff Cain et.al 2007, Toon Calders et.al 2012). Students are requested to express various problems that they may have encountered in the course (Brijesh Kumar Baradwaj et.al 2011). But, students' feedback is highly helpful when it is taken in real-time, as it permits them to report problems promptly (D Rogers et.al 2013). Students' feedback is very significant as it can help the faculty to realize the students learning behavior. Now and then students do not understand what the faculty is tiresome to describe, students can indicate this to the faculty through their feedback. Student's feedback is mostly measured on Likert-scale through a closed-ended question. However, there are some open-ended questions where the responses cannot be analyzed structurally. Therefore, these descriptive answers are analyzed thoroughly. This analysis will become a tedious task when numbers are huge. In a higher education institution in India where the students mostly come from different states, there are differences in languages used when the students expressing their feelings for the semester that recently passed. So structural analysis is needed for open-ended answers, by adopting a sentiment analysis technique this study to detect the students' positive or negative feelings and distinguished emotions, texted feedback. The question posed in the present study is: How the student's real-time feedback helps in identifying the student's sentiments and emotion towards online teaching? we ask two research questions

RQ1. Would student's sentiments and emotions vary in online teaching?

RQ2. Would students' sentiments and emotions indicate their satisfaction and engaged level in online teaching?

To answer these questions, we conducted a literature review, from which three objectives were formulated. The data used for the user is the student's feedback (text form) collected through google forms. Lexicon-based Sentimental analysis was used to obtain the result. R Studio was used in this study to do the analysis.

This paper is categorized as follows: first, based on the research question the literature is reviewed and discussed, second, the objectives identified, Then, the study methodology and analysis are described. Finally, results are obtained and discussed.

2 Related work and Research Gap

2.1 Sentimental analysis in education

Ortigosa et al (2014). studied the Facebook data collected from Spanish students, the sentiments in expressions of polarity (positive, negative, and neutral) was detected by using the two methods in sentimental analysis, lexicon approach and combine approach, noticed the polarity in terms of in the perspective of e-learning and pointy out that noticing polarity could be used in adaptive e-learning systems to support personalized learning. Troussas et al (2016). used a machine learning approach like Naive Bayes, Rocchio, and Perceptron classifiers to discovered students' divergence through Facebook in the circumstance of a language learning application. Troussas et al (2020) say that students who are functioning to manage with sentiments may not have appropriate resources accessible to engage in learning. Besides, emotions such as anger, anxiety, and sadness harm students' learning and can disturb students. Further happy students typically perform better. So noticing these feelings can help teachers to take action by concentrating on particular students. Song et al, (2018). also reconnoitered sentiment analysis on online learning systems to discover students 'opinions about courses, lecturers, and online learning systems, this study was conducted using Chinese language data to segregate polarity (i.e. positive and negative) from the sentiments through the SVM classifier machine-learning approach. Tian et al (2017). spotted e-

learners' emotions (i.e. Love, Joy, Anger, Frustration) from their texts in a chat system using POS-tags features and C4.5 decision tree as a classifier this study emphasized the benefits of identifying emotions through text, which is a less expensive method compared to others. Munezero et al (2014). discovered students' emotions from their learning diaries using lexicon-based sentimental analysis and identified the eight emotions are joy, sadness, fear, anger, anticipation, surprise, disgust, and trust. Utmost all the research on sentiment analysis in the educational area has been focused on the e-learning area. Though, there are differences between e-learning and classroom-based sentiment analysis given the words used differently in each area. Only one study was related to classroom-based sentiment analysis where the authors applied sentiment analysis to students' diaries Munezero et al (2013), Francis F. Balahadia, Ma. Corazon G. Fernando and Irish C. Juanatas (2016) developed a Sentimental analysis system to appraise English and Filipino course teacher performance. They estimated sentiment scores from student's ratings both in terms of qualitative and quantitative using a Naïve Bayes algorithm and vividly represented the percentage of positive and negative sentiments of students to help university administrators be aware of students' concerns. The literature exposes that there is limited research on sentiment analysis for students' real-time classroom feedback that uses machine learning techniques. But the student's feedback will have varied in the online mode of teaching so based on this the objectives of the study were formulated.

3 Objectives of the Study

- To recognize the students' positive or negative feelings, distinguished emotions, towards online teaching.
- To measure the sentimental word association
- To express the results and generate sentiment visualization.

4 Methodology

4.1 Data Selection and Extraction

The data used for the study was mined from student's feedback for online courses conducted in selected Top B Schools in Bangalore. The nonprobability sampling method of purposive sampling was used in this study. Purposive sampling technique is an approach where members conform to certain criteria. The main criteria for collecting students' feedback is the student must have attended at least 20 hours of online classes in a particular course handled by the concerned faculty. About 350 written feedbacks were taken for the study to do the sentimental analysis. Feedback from the students collected through Google forms.

5 Data Analysis and Discussion

The machine learning technique of Lexicon based sentimental analysis was used to do the sentimental analysis. The analysis of this study is separated into three-steps, the first is text mining, the next step is estimating sentiment score and the third step is the classification of student sentiment. R Studio was used in this study to do the analysis. R contains a huge set of Natural Language Processing (NLP) packages. The initial stages contain stacking the feedback (text file) into an R, then data preprocessing was done beforehand executing the sentimental analysis. This analysis demonstrates the following phases and analyses like Data preprocessing, Word Frequency, Word Cloud generation, Word Association, Sentiment Scores, and Emotion Classification

5.1 Text mining

Data Reading

The input case for this analysis has "Raw text" of replies (student's feedback), and is a text file. so here we used to read Lines function to read the data. The read Lines function merely abstracts the text from the Text document and consider each line as a character string. To read a subset of lines from the text document the n= argument logic was used

Data Preprocessing

Data preprocessing starts from cleaning the text data, which means removing distinct characters from the text. By using the `tm_map` function, the distinct characters like @, /, and | converted into space. The succeeding process is to eliminate the excessive space and convert the full-text data into lower case. Then the stop words are eliminated. The stop words are the most frequently occurring words in any language, example in English "the", "is", "are" have very minute value for NLP and it should remove for further processing. Then SnowballC package was used to do the stemming process. The stemming procedure shortens the word to its common source. The previous studies related to sentimental analysis in the educational domain were used this all mentioned preprocessing technique (J.M. Martin, A.et al 2013 & Erik Tromp, 2014).

Word frequency

After data preprocessing, the following phase is to build the document matrix, this table contains the frequency of words detected from the input text document, this frequency of words used to detect general or trending words in the feedback. By using the `TermDocumentMatrix ()` function from the text mining package, the below Document Matrix – a table was built. The resulting table1 shows the word frequency of input text document (students feedback).

Table 1. Word Frequency

S.No	Words	Frequency
1	Good	265
2	Pattern	40
3	Lecturer	39
4	Student	38
5	Interaction	38
6	Paper	35
7	Teacher	34
8	Exam	34
9	University	32
10	Check	32

Fig 1. Word Frequency

The maximum frequently occurring word is "good". Also noticed in the table there are no negative words like "not", which shows there are no negative prefixes to change the context or meaning of the word "good" (In short, this shows most replies don't comment

negative phrases like "not good"). Pattern, Lecture, student, interact are the subsequent repeatedly occurring words, which direct that utmost the students feel good about their lecture and pattern. Finally, the root “interact” and “lectur” for words like “interaction” and “lecturer”, etc. is also on the chart, and further analysis is needed to identify whether the perspective is positive or negative

Word Association

Word association means to find the relationship between the pairs of words, Correlation is a statistical technique used to find out the magnitude of the relationship between two variables. This tool was used to examine the association between the most recurrently occurring words in the student’s feedback and is used to view the student’s perspective about these words.

Table 2. Word Association

Words	goodlectur	lectur	student	style	understand
teach	0.35	0.3	0.25	0.29	0.28

Words	interact	teacher	lectur	delivery	Knowledge	proper	staff
punctual	0.54	0.52	0.48	0.44	0.3	0.25	0.25

The above table2. Shows the words which are most often associated with the selected three terms (The Lower corlimit = 0.25 is set). The output specifies that "teach" (which is the root for the word "teacher "or "teaching") has the highest of 35% association with good lecture and the lowest of 25% association with the student. So the most frequently occurring word ("teach") is positive. Similarly, another important word "Punctual" is extremely associated with the word "interact". These results show that most feedback of students is mentioning the faculty "Punctual and interaction" and can be taken from a positive perspective.

Sentiment Scores

Student sentiment towards online teaching was measured through sentiment scores. Here the sentiments are categorized as positive, negative, and neutral. And its denoted by a numeric scale. The numeric scale provides better details in identifying the unit of positive or negative strength of the sentiment enclosed in a student’s feedback. In this, we used the Syuzhet package for generating sentiment scores, which has four sentiment

dictionaries and offers a technique for retrieving the sentiment. This sentiment extraction tool (NLP) was developed at Stanford University. The get sentiment function takes two arguments: a character vector (of sentences or words) and a method. The designated technique decides which of the four available sentiment extraction methods will be used. The four techniques are syuzhet bing, affinn, and nrc. Each technique uses a diverse scale and hence results are marginally different.

Syuzhet Vector Results

Table 3. Summary Results of Syuzhet Vector

S.No	1	2	3	4	5	
1	2.35	0.75	1.50	0.75	1.5	
S.No	Min	1st Qu	Median	Mean	3rd Qu	Max
1	-1.3000	0.5000	0.7500	0.8665	1.0000	5.35

From the above table 3., Summary results of the Syuzhet vector show the first component has a value of 2.35. It means the sum of the sentiment scores of all important words in the first response(line) in the text file, adds up to 2.35. The scale used for measuring sentiment scores ranges from -1(indicating most negative) to +1(indicating most positive). The summary statistics table showed that the median value is 0.7500 and the mean value is 0.8665, which is above zero and inferred as the overall average sentiment across all the responses (student's feedback) is positive. It means students' sentiment towards online teaching is positive.

Bing and Affin Vector Results

Table 4. Summary Results of Bing and Affin Vector

S.No	Vector	1	2	3	4	5	
1	Bing	1	1	2	1	1	
2	Affin	3	3	3	2	2	
S.No	Vector	Min	1st Qu	Median	Mean	3rd Qu	Max
1	Bing	-2	0	1.00	0.87	1.00	5.00
2	Affin	-5	1	3.00	2.52	3.00	10.00

The Sentiment scores scale for bing is a binary scale with -1 representing negative and +1 representing positive sentiment and affinn is an integer scale ranging from -5 to +5. From the above table 4. summary results of bing and affinn vectors also show that the Median value of Sentiment scores is above 0 and inferred as the overall average sentiment across all the student's responses (feedback) is positive.

Comparison of Three Vectors

Meanwhile, the study used three different methods and three diverse scales, it's well to change their output to a common scale before relating them. By using sign function in R, the basic scales were converted, the sign function converts all positive numbers to 1 and all negative numbers to -1, and all zeros remain 0.

Table 5. Normalize scale and compare three vectors

	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	0	1	1	1	1	1

Form the above table 5. it's Noted the first element of the first two rows (vector) is 1, indicating that all three methods have calculated a positive sentiment score, for the first response (line) in the text.

Emotion Classification

This study used the NRC Word-Emotion Association Lexicon to classify the emotion of students towards online teaching. The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive). `get_nrc_sentiments` function was used in this study. This function, which returns a data frame with each row representing a sentence from the original text. The data frame has ten columns (one column for each of the eight emotions, one column for positive sentiment valence, and one for negative sentiment valence). The data in the columns (anger, anticipation, disgust, fear, joy, sadness, surprise, trust, negative, positive) can be accessed individually or in sets. The following table shows the Data frame returned by the `get_nrc_sentiment` function

Table 6. Data frame

S.No	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Negative	Positive
1	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	2	0	3
3	0	1	0	0	1	0	1	1	0	1
4	0	1	0	0	1	0	0	2	0	2

5	0	1	0	0	1	0	1	1	0	1
6	0	1	0	0	1	0	0	1	0	2

From the above table 6, it inferred that Zero existences of words related to emotions of disgust, anger, sadness, and fear. 1 incidence each of the words linked with emotions of anticipation, joy, and surprise. And 2 incidences of words linked with emotions of trust. No incidences of words are linked with negative emotions. Maximum 3,2,1 existences of words in positive emotions. The subsequent phase is to generate charts to help visually examine the emotions of students. The below chart shows the entire amount of occurrences of words in the text, related to each of the eight emotions.

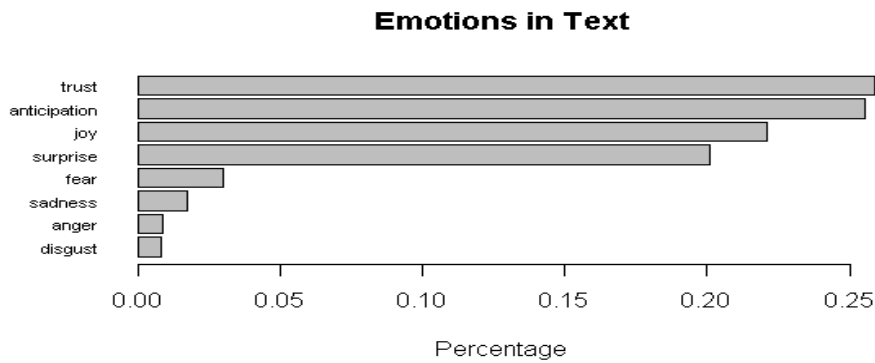


Fig.3. Emotions in Text – Bar plot

The above fig 3, bar plot permits a quick and easy assessment of the proportion of words linked with each emotion of students towards online teaching. The emotion “trust” has the lengthiest bar and displays that words linked with this positive emotion found just over 25% of all the expressed words expressed by the students in their feedback. The next emotion is “disgust” has the shortest bar and displays that words linked with this negative emotion found less than 1% of all the expressed words by the students in their feedback. Overall, words linked with the positive emotions of “trust” and “anticipation” account for almost 50% of the significant words in the text, and it can be interpreted as a good sign of emotion towards the online teaching.

6 Insights from the Analysis

After applying machine learning algorithms and estimate sentiment score and classified the emotion of students towards online teaching., the teaching activities to be examined were identified. Potential sentimental words or related words to the online teaching identified were: good, pattern, lecture, student, interact, paper, teacher, exam, the universe, check. Detecting these words, the objective is to explore which words impact students' sentiment scores the most, and which ones provide insight into student emotion towards online teaching. Good is a vastly said word, occurring in many student responses, in positive sentiment. This could mean that online teaching is good concerning the particular faculty and hour in emotional classification Trust is the highly mentioned word, this could mean the students trusted the content and faculty. Over some time, the emotion falls under the positive category also.

7 Conclusion

This study identifies and recognizes the students' positive and negative feelings, distinguished emotions, towards online teaching. and also measured the sentimental word association and Emotional variance over online teaching. Using a lexicon sentimental analysis method and affin, bing, and syuzhet vector concept the sentiment was classified into positive and negative. The results showed from all techniques the sentiment falls under the positive side. It necessity be noted here that student feedback data are only one source of data in general towards online teaching, but syllabus/module content and delivery, and should be used in combining with other teaching and learning evaluation strategies for further study.

8 References

- Ann Poulos and Mary Jane Mahony. Effectiveness of feedback: the student's perspective. *Assessment & Evaluation in Higher Education*, 33(2):143–154, 2008.
- Brijesh Kumar Baradwaj and Saurabh Pal. Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications*, 2(6):63–69, 2011.

- D Rogers, P Coughlan, and K Kearney. Playschool for postgrads: facilitating creativity and innovation in doctoral education using digital resources in, the digital learning revolution in Ireland: Case studies in practice from ndlr. pages 165–179, 2013.
- Munezero, M., Montero, C. S., Sutinen, E., & Pajunen, J. Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, 5(2), 101-111,2014
- Myriam Munezero, Calkin Suero Montero, Maxim Mozgovoy, and Erkki Sutinen. Exploiting sentiment analysis to track emotions in students' learning diaries. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research*, pages 145–152. ACM, 2013.
- Ortigosa, A., Martín, J. M., & Carro, R. M. Sentiment analysis in Facebook and its application to e-learning. *Computers in human behavior*, 31, 527-541,2014
- S. Cummins, L. Burd, and A. Hatch. Using feedback tags and sentiment analysis to generate sharable learning resources investigating automated sentiment analysis of feedback tags in a programming course. *Advanced Learning Technologies (ICALT)*, 10:653–657, July 2010.
- .Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., & Sun, C. Facebook sentiment: Reactions and emojis. In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media* (pp. 11-16),2017
- Toon Calders and Mykola Pechenizkiy. Introduction to the special section on educational data mining. *Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining.*, 13(2):3–6, May 2012
- Troussas, C. *Advances in Social Networking-based Learning: Machine Learning-based User Modelling and Sentiment Analysis* (Vol. 181). Springer Nature,2020
- Troussas, C., Espinosa, K. J., & Virvou, M. Affect Recognition through Facebook for Effective Group Profiling Towards Personalized Instruction. *Informatics in Education*, 15(1), 147-161,2016
- Yoo, S., Song, J., & Jeong, O. Social media contents based sentiment analysis and prediction system. *Expert Systems with Applications*, 105, 102-111,2018