Lexicon Based Sentiment Analysis of Open-Ended Students' Feedback

Khin Zezawar Aung, Nyein Nyein Myo

Abstract: Sentiment analysis is helpful in finding the opinion of writer's feeling towards a specific topic. Teaching evaluation is a useful tool of assessment for teaching and courses at many universities, colleges and schools. Mostly close-ended questions and open-ended questions are used in teaching evaluation process. This paper used open-ended questions to provide the opinion result for teachers' effectiveness of teaching and over all course condition. In this paper, teaching sentiment lexicon, Afinn lexicon and Opinion lexicon are used to get the scores of opinion words in feedback comments. The students' feedback comments are analyzed by using three methods and display the opinion result as positive, negative and neutral class. According to the experimental results, the intensifier words are needed to consider in some feedbacks to get the correct opinion result. The accuracy of Method 1 using teaching sentiment lexicon is better than other two methods.

Index Terms: Lexicon Based, Opinion Mining, Sentiment Analysis, Students' Feedback.

I. INTRODUCTION

Sentiment analysis is sub field of opinion mining which analyzes the user generated data like reviews, people's opinions, evaluations, feedbacks, comments, and emotions towards topics such as products, services, entertainment, politics, social media, organizations, events, topics. Generally, opinion mining becomes the result of a people's feelings and sentiments, etc. So, opinion mining is popular topic in the current research area. The main objective of sentiment analysis is to classify the text into positive, negative and neutral category [2]. This system performs lexicon based sentiment analysis to display the opinion result for teachers by using open-ended students' feedback.

Generally, teaching evaluation [15] consists of close-ended question with fixed questions using multiple choices and open-ended questions. Most of the answers of close-ended question contain five levels: strongly agree, agree, neutral, disagree and strongly disagree. The answers of open-ended question contains the unstructured textual feedback and describes freely desire writing. The open-ended textual comments assist a chance to students [13] to point out some features which are not exactly covered by close-ended questions. Students' feedbacks are the important portion of teaching evaluation process. It can be collected by applying teaching evaluation survey form containing questions. It is important to survey students to obtain feedback regarding the teaching effectiveness and course over all condition.

The rest of the paper is structured as follows. Section 2 mentions the related works of sentiment analysis. The sentiment analysis approaches are discussed in section 3. The architecture of the proposed system is presented in section 4. The experimental set up of our research is mentioned in section 5. Conclusion of this paper is described in section 6.

II. RELATED WORKS

This section presents about the related works conducted for sentiment analysis of teaching reviews and product reviews. In 2017, Z. Nasim et al. [16] proposed sentiment analysis of student comments using hybrid approaches. Their paper presented a model to analyze the sentiment of students' textual feedback by applying TF-IDF and lexicon based approach. Their dataset contains 1230 comments extracted from Institute of Business Administration (IBA) located from Karachi, Pakistan. They compared their hybrid approach with other web based sentiment analysis tools such as Text Analytics API, Alchemy language API and Aylien Text API Their research work shows that their hybrid approach outperforms than other sentiment analysis tools.

M. Wen et al. [12] apply sentiment analysis on teacher's feedback from students enrolled in Massive Open Online Course (MOOC) in order to realize their opinion and decide whether there is a connection between emotions and student dropout rates. Their dataset contains three courses from Coursea.org using a screen scraping protocol. They find that the relationship between sentiment and student dropout rate makes sense once one examines practices for expressing sentiment within that specific course context. Student comments on quality and learning standards are considered as a strategy to enhance the teaching process.

In 2017, G. G. Esparza et al. [5] proposed a sentiment analysis of students' feedback using support vector machines. They describe a social mining model applying a corpus of real feedback comments in Spanish about teacher performance. They applied Support Vector Machines algorithm with three types such as linear, radial and polynomial, to determine a category of feedbacks in negative, positive or neutral. They evaluated performance of their system by calculating the accuracy, balanced accuracy, specificity, sensitivity and ROC curve. Their dataset contains 1040 comments of engineer students from Polytechnic University of Aguascalientes. The balanced accuracy of Support Vector Machines with linear kernel got 0.8149 and the sensitivity value is higher than all kernels. Support Vector Machines with polynomial obtained the lowest performance in all Metrics except in sensitivity.

Manuscript published on 30 December 2018.
* Correspondence Author (s)
   Khin Zezawar Aung*, Faculty of Information Science, University of Computer Studies, Mandalay, Myanmar,
   Nyein Nyein Myo, Faculty of Information Science, University of Computer Studies, Mandalay, Myanmar,

L. V. Avanco and M. G. V. Nunes [8] proposed the sentiment analysis of products review using lexicon based technique. The authors describe the lexicon-based sentiment polarity results of product reviews in Brazilian Portuguese language. They used three different lexicons such as Opinion Lexicon, a subset of LIWC and SentiLex lexicon. The performance of three different lexicons is evaluated using simple algorithm. The experimental results show that the better combination is the version of the algorithm that concerns with negation and intensity and apply the Sentilex sentiment lexicon. Their system got 0.73 of average F-measure.

III. SENTIMENT ANALYSIS APPROACHES

Three main approaches can be done for sentiment analysis: machine learning based [5], lexicon based [7], [8] and hybrid approach [3]. Most of the machine learning based approach apply training data and calculate the performance on the test dataset. This approach applies classification technique to classify text. Lexicon based approach of sentiment analysis uses a lexicon or a sentiment dictionary that contains words list concerning with sentiment polarity [2].

Lexicon based approach includes calculating the orientation for a document from the semantic orientation of words or phrases [11]. In generally, the lexicon can be created either automatically or manually [2]. The authors in [11] presented the lexicon based approach to define the words' orientation by using their constructed dictionaries. The authors in [10] used an online lexical resource WordNet to define the orientation of word. Hybrid approach of sentiment analysis applies sentiment lexicon and machine learning methods. Most of the hybrid approach applies sentiment lexicon for labelling the training data [3] and applies the machine learning classifier to classify the polarity of sentiment on the process of evaluation.

The lexicon-based technique mainly depends on opinion words that describe the positive or negative sentiments. Lexicon is very important part in sentiment analysis, enhancing the quality of lexicons can improve the efficiency and accuracy of sentiment analysis. The following section describes about the Opinion lexicon, Afinn lexicon and Teaching sentiment lexicon.

A. Opinion Lexicon

Opinion lexicon [6] composed of around 6800 English words classified into positive and negative classes. Liu et al. applied the antonym and synonym of adjective words in WordNet to define the adjective orientation of words. Firstly, small list of adjectives seed is manually labelled as positive or negative. This adjective seed list is domain independent. The list will be grown by applying WordNet, resulting in a list of 4783 negative terms and 2006 positive terms including slang words, morphological variants, and misspellings.

B. Afinn Lexicon

Afinn lexicon [1] was created in 2009 for tweets downloaded for online sentiment analysis. The old version called AFINN-96 contains 1468 different words, including a few phrases. The updated version called AFINN-111 contains 2477 words. The score in AFINN is ranging from -5 to +5. -5 means very negative and +5 means very positive.

The words were manually scored by the author. The word list in Afinn lexicon starts with a set of obscene words [4]. Most positive words were scored with +2 and most negative words were scored with –2, strong obscene words with either –4 or –5.

C. Teaching Sentiment Lexicon

Opinion words are defined in teaching sentiment lexicon. Small set of opinion words in open-ended students' feedback are collected manually and assigned the polarity scores by two skillful persons who have understanding in assessment of teaching. The opinion words are increased by searching in the thesaurus for their synonym words and antonym words. For example, synonyms for understand are "know, aware, realize, comprehend, apprehend, catch, etc." The antonyms for understand are "misunderstand, fail, neglect, disregard, etc." In this teaching sentiment lexicon synonym word is defined as a positive opinion word and antonym word is defined as a negative opinion word [14]. But scores of all synonyms are defined differently. The score of 'inspire' is stronger than the score of 'motivate'. The word 'motivate' has opinion score +2 but the word 'inspire' has opinion score +3. The reliability of the opinion words' scores are defined by a language expert from the Department of Languages.

This proposed approach is based on the idea of Taboada et al. [11]. Total words in teaching sentiment lexicon are 1437 words. Total number of positive opinion words is 958 words and total negative opinion words are 479 words. The score of sentiment ranges from -3 (strongly negative) to +3 (strongly positive). The score of positive opinion words ranges from +1 to +3 and the score of negative opinion words ranges from -1 to -3. There is a reason why we need to give the score of opinion word. Consider the following sentence.

"My teacher can solve the unseen problem".

The word 'problem' has opinion score -1 and the word 'solve' has opinion score +2 in teaching sentiment lexicon. So the opinion result of this sentence is positive using teaching sentiment lexicon. But in Afinn lexicon, the word 'problem' has opinion score -2 and the word 'solve' has opinion score +1. Therefore the opinion result is negative applying Afinn lexicon. Opinion lexicon contains the word 'problem', it is negative opinion word and gives opinion score -1 but the word 'solve' isn't contain in Opinion lexicon. So the opinion result is negative using Opinion lexicon. The actual opinion result is positive. So the score of opinion word is very important to define the correct opinion result. Some of the words in Opinion lexicon are negative. But these words need to consider as positive word in students' feedback. Consider the following sentence.

"Our teacher is very concerned about student's well-being in the course."

The word 'concern' is needed to consider as a positive opinion word in teaching sentiment lexicon but negative opinion word in Opinion lexicon. Similarly, consider the following sentence.

"Our teacher makes a joke example and calls us back to her lesson."

In this sentence, the word 'joke' is a positive opinion word in teaching sentiment lexicon but negative opinion word in Opinion lexicon. The word 'funny' is also a positive opinion word in teaching sentiment lexicon and negative opinion word in Opinion lexicon. Similarly, consider the following sentence.

"Teaching speed is very fast in her teaching time".

In the above sentence, the word 'fast' is a negative opinion word in teaching sentiment lexicon but positive opinion word in Opinion lexicon. In educational domain, the word 'fast' is needed to consider as negative opinion word.

Different scores of some words in three lexicons are mentioned in Table I. The symbol, '_', means not present in corresponding lexicon.

Table I. Different Scores in Three Lexicons

Opinion Word	Score in Teaching Sentiment Lexicon	Score in Afinn Lexicon	Score in Opinion Lexicon
joke	2	2	-1
funny	2	4	-1
understand	2	_	_
fast	-1	_	1
overcome	3	_	_
care	2	2	_
knowledgeable	2	_	1
solve	2	1	_
solving	2	1	_
concern	2	_	-1
concerned	2	_	-1
familiar	2	_	_
enough	2	_	1
aid	2	_	_
approachable	2	_	_
challenging	1	_	-1
challenge	1	-1	_
problem	-1	-2	-1
complicate	-3	_	_
sleepy	-1	_	_

IV. PROPOSED SYSTEM ARCHITECTURE

This section presents the architecture of the proposed system. Figure 1 shows the architecture of lexicon based sentiment analysis for open-ended students' feedback.

This system consists of five main functions. These functions are categorization of comments, preprocessing, opinion word identification and assignment of polarity score using three lexicons, calculation of total and average polarity score, showing the opinion result using three methods. Firstly the teaching evaluation summary from University of Computer Studies, Mandalay (UCSM) is collected. The idea of Stanford University's teaching evaluation summary form is used to ask the questions for UCSM students. Four types of question are used to collect the feedback. These are:

1. Please comment on your instructor with regard to effectiveness.

2. Please comment on your instructor's attitude towards students.

3. Please comment on assignments and exams.

4. Do you have any additional comments on the course over all.
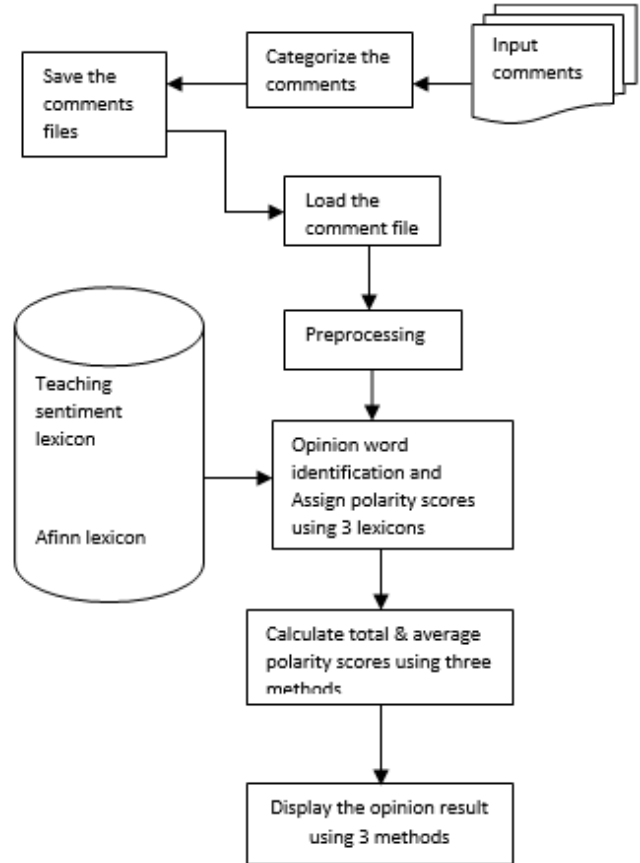
Figure. 1. System Architecture

The feedback of first year, second year, third year, fourth year and master students are collected at the end of first semester in 2018. The input comments of all questions types are kept in csv file format. These input comments are categorized into four files according to four question types and then saved into Notepad file. To display the opinion result, one file is loaded.

Next step is preprocessing. The preprocessing step involves case conversion that changes all words into lower case, removing of stop words, split sentence, replace "n't" with "not" to catch the negation words, punctuation removal and tokenization.

Next step is opinion word identification and assign the polarity scores using three lexicons. The proposed system compares each tokenized word in the input comment with negation word or intensifier word or opinion word by using teaching sentiment lexicon, Afinn lexicon and Opinion lexicon. And then the corresponding polarity scores are assigned to each word by using three lexicons. After that the total and average polarity score of all sentences are calculated by using three methods. Method 1 represents the opinion result using teaching sentiment lexicon and calculates negation word, intensifier words and opinion word. The intensifier scores are not equally assigned in Method 1. In Method 2, the opinion result is calculated using Afinn lexicon and Method 2 calculation includes negation word and opinion word but not included the intensifier word. Method 3 calculates negation word,

intensifier words and opinion word using Opinion lexicon. In Method 3, positive word lists are assigned to +1 and negative word lists are assigned to -1. The intensifier scores are equally assigned +50% for booster or amplifier words and -50% for down toner. Finally, the opinion result using three methods is displayed. The following algorithm is used to calculate the opinion result using Method 1 and 3.

```
Input: students' feedback comment
Output: Opinion Result using Method 1 and Method 3.
Begin:
1. For each sentence in comment:
2.   For each word in sentence:
3.     If word==Negation then
4.       Neg=1
5.     If word==Intensifier then
6.       icount=icount + 1
7.       If icount==1 then
8.         Give iscore1
9.       If icount==2 then
10.        Give iscore2
11.    If word==opinion then
12.      opcount=opcount+1
13.      opscore = teaching lexicon or  Opinion lexicon
14.      If icount==1 and Neg==1 then
15.        opscore= iscore1* opscore* (-1)
16.        Neg=0
17.      Else if icount==2 and Neg==1 then
18.        opscore= iscore1* iscore2* opscore* (-1)
19.        icount=0; Neg=0
20.      Else if Neg==1 then
21.        opscore= opscore * (-1)
22.      Else if icount==1 and Neg==0 then
23.        opscore= iscore1* opscore
24.      Else if icount==2 and Neg==0 then
25.        opscore= iscore1* iscore2* opscore
26.        icount=0
27.    tot=tot + opscore
28.  Next word
29. Next sentence
30. avgScore= tot/opcount
31. If avgScore>0 then opinionResult="positive"
32. Else if avgScore<0 then opinionResult=" negative"
33. Else if avgScore==0 then opinionResult="neutral"
End
```

In the above algorithm, the negation word is searched firstly. Then, the intensifier words are searched. Finally, the opinion word is searched by using Teaching sentiment lexicon for Method 1 and by using Opinion lexicon for Method 3.

The algorithm for Method 2 is described as follow.

```
Input: students' feedback comment
Output: Opinion Result using Method 2.
Begin:
1. For each sentence in comment:
2.   For each word in sentence:
3.     If word==Negation then
4.       Neg=1
5.     If word==opinion then
6.       opcount=opcount+1
7.       opscore =Afinn lexicon
8.       If Neg==1 then
9.         opscore=opscore*(-1)
10.        Neg=0
11.      tot = tot + opscore
14.  Next word
15. Next sentence
16. avgScore =tot/opcount
17. If avgScore>0 then opinionResult="positive"
32. Else if avgScore<0 then opinionResult="negative"
33. Else if avgScore==0 then opinionResult="neutral"
End
```

In algorithm for Method 2, we search the negation word firstly and then opinion word is searched using Afinn lexicon. In Method 2, the intensifier words are not considered.

V. EXPERIMENTAL SETUP

The section presents the experimental result of our research work. The open-ended students' feedbacks from University of Computer Studies, Mandalay (UCSM) are used to test the experiment. All feedback comments of all teachers are manually labelled to find the accuracy [9]. The performance of this system is tested using four evaluation values: Accuracy, Precision, Recall and F-Measure.

Accuracy (A) = (TP+TN)/ (TP+FP+FN+TN)

Precision (P) =TP/ (TP+FP)

Recall (R) = TP/ (TP+FN)

F-Measure (F1) = (2*P*R)/ (P+R) where, TP is True positive, FP is False positive, FN is False negative. TN is True negative. The feedback comments of UCSM at the end of first semester in 2018 are used to test the performance of this system. The opinion result is depended on the content of students' feedback.

The comparison of some opinion results using three methods are shown in Table II.

Table II. The Comparison of Some Opinion Results

Students' Feedback	Manual Opinion Result	Opinion Result using Method 1	Opinion Result using Method 2	Opinion Result using Method 3
She treats us like a mother to her children. She knows the weakness of her children.	positive	positive	neutral	neutral
She explains us the lecture to understand.	positive	positive	neutral	neutral
She is a very helpful teacher. She explains the misunderstanding problems again and again.	positive	positive	negative	negative
She can teach her lessons to understand easily.	positive	positive	neutral	neutral
She makes interactive discussion between us.	positive	positive	neutral	neutral

I can't understand her explanation.	negative	negative	neutral	neutral
We don't understand some chapter and I think that she is fast in teaching.	negative	negative	neutral	positive
She explains me how to solve problems and also solve the problems together.	positive	positive	negative	negative
We understand her teaching. She can explain everything that we don't know.	positive	positive	neutral	neutral
The textbook presented practice problems which were very helpful.	positive	positive	neutral	positive
Her teaching method is needed to be better. I feel sleepy during her teaching time.	negative	negative	positive	positive
She gives a difficult assignment but it is very useful in exam.	positive	positive	positive	positive
When we are sleepy, she makes a joke and calls us back to her lesson.	positive	positive	positive	negative
She is very familiar with her students.	positive	positive	neutral	neutral

The experimental result of students' feedback from UCSM is mentioned in Table III, IV, V, VI and VII respectively.

Table III. Experimental Result of First Year Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.86	0.96	0.91	85 %
Method 2	0.75	0.83	0.79	67 %
Method 3	0.76	0.83	0.79	70 %

Table IV. Experimental Result of Second Year Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.98	1	0.99	98%
Method 2	0.93	0.99	0.96	92%
Method 3	0.89	0.97	0.93	88%

Table V. Experimental Result of Third Year Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.97	1	0.98	97%
Method 2	0.9	0.95	0.92	85%
Method 3	0.82	0.96	0.88	79%

Table VI. Experimental Result of Fourth Year Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.98	0.96	0.97	94%
Method 2	0.83	0.95	0.89	80%
Method 3	0.83	0.92	0.87	77%

Table VII. Experimental Result of Master Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.99	1	0.99	99%
Method 2	0.88	0.88	0.88	78%
Method 3	0.87	0.96	0.91	84%

The average precision, recall and f-measure of students' feedback are mentioned in the Table VIII. Obviously, the average precision, recall and f-measure of Method 1 using teaching sentiment lexicon is higher than Method 2 and Method 3.

Table VIII. The Average Precision, Recall, F-Measure and Accuracy of Students' Feedback

Test Method	Precision	Recall	F-Measure	Accuracy
Method 1	0.96	0.98	0.97	95 %
Method 2	0.86	0.92	0.89	81 %
Method 3	0.83	0.93	0.88	80 %

According to the experimental result, intensifier word would consider in some feedback. "The textbook presented practice problems which were very helpful". In this sentence, the word 'problems' is a negative opinion in Afinn lexicon and it has opinion score -2. The word 'helpful' is a positive opinion in Afinn lexicon and it has opinion score +2. Method 2 doesn't consider the intensifier words. So the opinion result using Method 2 is neutral. According to Method 3, the intensifier scores are considered to calculate the opinion result. The word 'problems' has a negative opinion score -1 and 'helpful' has a positive opinion score +1 using Opinion lexicon. The intensifier word, 'very' is calculated using Method 3. Although the score of positive opinion word and negative opinion word are equal, the correct opinion result, 'positive', is displayed due to the score of intensifier word. If the intensifier score isn't used to calculate the opinion result, the opinion result is neutral using Method 3. "She gives a difficult assignment but it is very useful in exam". The opinion result of that sentence is positive using three methods. If the intensifier word, 'very', isn't considered in Method 3, the opinion result may be 'neutral'. So the intensifier score is needed to calculate to get the correct opinion result.

VI. CONCLUSION

Sentiment analysis field consists of several applications that have been addressed in large variety of research studies. The sentiment analysis of open-ended students' feedback for teacher has been used in educational institutions around the world to provide the quality of teaching. Today, it is very important tools used to inform teachers who want to know the understanding of their teaching. The goal of this paper is to display the suitable opinion result for teachers by applying the open-ended feedback of students at the any University. Three methods are used to analyze the students' feedback. The precision, recall, f-measure and accuracy of Method 1 using teaching sentiment lexicon is better than Method 2 and 3. The research addresses some opinion words related with educational domain to display the correct opinion result. The intensifier words are needed to calculate to provide the correct opinion result in some feedbacks.

REFERENCES

1. AFINN word database an affective lexicon by Finn Årup Nielsen, http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6010/zip/imm6010.zip.
2. B. Liu "Sentiment analysis and opinion mining". Synthesis Lectures on Human Language Technologies, May 2012, pp. 1–167.

3. C. Pong-inwong and W. Songpan (Rungworawut), "TeachingSenti-Lexicon for Automated Sentiment Polarity Definition in Teaching Evaluation". 10th International Conference on Semantics, Knowledge and Grids (SKG), Beijing: IEEE, 2014, pp. 84 – 91.

4. Finn Arup Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", ESWC2011 Workshop on Making Sense of Micro posts: March 2011, pp. 93-98.

5. G. G. Esparza, Alejandro de-Luna, Alberto Ochoa Zezzatti, Alberto Hernandez, Julio Ponce, Marco Alvarez, Edgar Cossio and Jose de Jesus Nava, "A Sentiment Analysis Model to Analyze Students Reviews of Teacher Performance Using Support Vector Machines", 14th International Conference on Distributed Computing and Artificial Intelligence (DCAI 2017, June, 2017,), pp. 157-164.

6. http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

7. K. Z. Aung and N. N. Myo, "Sentiment Analysis of Students' Comment Using Lexicon Based Approach", 16th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2017), May, 2017, pp. 149-154.

8. L. V. Avanco, M.G.V.Nunes, "Lexicon-based Sentiment Analysis for Reviews of Products in Brazillian Portugues", 2014 IEEE Brazillian Conference on Intelligent Systems, pp. 277-281, 2014.

9. M. El-Masri, N. Altrabsheh, H. Mansour, A. Ramsay, "A web-based tool for Arabic sentiment analysis", 3rd International Conference on Arabic Computational Linguistics, ACLing 2017, November, 2017.

10. M. Hu and B. Liu, "Mining and summarizing customer reviews," ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2004, pp. 168-177.

11. M. Taboada, J. Brooke, M. Tofiloski, K. Voll and M. Stede, "Lexicon-based methods for sentiment analysis", Association for Computational Linguistics, 37(2), 2011, pp. 267-307.

12. M. Wen, D. Yan and C. P. Rose, "Sentiment Analysis in MOOC Discussion Forums: What does it tell us?", Proceedings of Educational Data Mining, 2014, pp. 1-8.

13. N. Altrabsheh, M. Cocea and S. Fallahkhair, "Sentiment analysis: towards a tool for analyzing real-time students feedback", 2014 IEEE 26th Internationaal Conference on Tools with Artificial Intelligence, 2014, pp.419-423.

14. P. Kaewyong, A. Sukprasert, N. Salim and F.A. Phang, "The possibility of students' comments automatic interpret using Lexicon based sentiment analysis to teacher evaluation", 3rd International Conference on Artificial Intelligence and Computer Science (AICS2015), 2015, pp. 179-189.

15. Q. Rajput, S. Haider and S. Ghani, "Lexicon-Based Sentiment Analysis of Teacher's Evaluation", Applied Computational Intelligence and Soft Computing, Hindawi Publishing Corporation, vol. 2016, September 2016.

16. Z. Nasim, Q. Rajput and S. Haider , "Sentiment analysis of student feedback using machine learning and lexicon based approaches", International Conference on Research and Innovation in Information Systems (ICRIIS), Langkawi, Malaysia, July, 2017.