Sentiment Analysis of Students' Comment Using Lexicon Based Approach

Khin Zezawar Aung
University of Computer Studies, Mandalay
Mandalay, Myanmar
zezawar12@gmail.com

Nyein Nyein Myo
University of Computer Studies, Mandalay
Mandalay, Myanmar
vicky.mdy@gmail.com

***Abstract*— In education system, students' feedback is important to measure the quality of teaching. Students' feedback can be analyzed using lexicon based approach to identify the students' positive or negative attitude. In most of the existing teaching evaluation system, the intensifier words and blind negation words are not considered. The level of opinion result isn't displayed: whether positive or negative opinion. To address this problem, we propose to analyze the students' text feedback automatically using lexicon based approach to predict the level of teaching performance. A database of English sentiment words is created as a lexical source to get the polarity of words. By analyzing the sentiment information including intensifier words extracting from students' feedback, we are able to determine opinion result of teachers, describing the level of positive or negative opinions. This system shows the opinion result of teachers that is represented as to whether strongly positive, moderately positive, weakly positive, strongly negative, moderately negative, weakly negative or neutral.**

Keywords— opinion mining; sentiment analysis; teaching evaluation; students' comments; lexicon based

I. INTRODUCTION

Opinion mining is a process for tracking the mood of the people about any particular topic by review. In general, opinion may be the result of people's personal feelings, beliefs, sentiments and desires etc. Most of the opinion mining research has focused on business and e-commerce applications, such as product reviews, hotel review and movie ratings etc. This research work focus on students' feedback comments. Feedback analysis is more important to measure the performance of teacher. Analyzing students' comments using sentiment analysis approaches can classify the students' positive or negative feelings. Students' feedback can highlight various issues students may have with a lecture [8]. Sometimes students do not understand what the lecturer is trying to explain, thus by providing feedbacks, students can indicate this to the lecturer.

The answers of student evaluation consist of quantitative and qualitative data collected by two question types; the quantitative data was collected by close-ended questions as multiple choices and the qualitative data was collected by open-ended questions as comments and suggestions from student opinions in textual form. It is easy to analyze quantitative data in numerical rating scale by statistical techniques, but it is difficult to interpret the qualitative data from the students' comments[14]. Most researchers on teacher evaluation focused to quantitative data but ignored qualitative data. However, the personal answers from open-ended questions contain the sentences or phrases regarding the teaching methods. The processing of qualitative data analysis is very important and it can enhance the teacher evaluation effectiveness[9].

This paper is organized as follows. In Section 2, the relevant previous work on sentiment analysis in education domain is presented. The methods in sentiment analysis are discussed in section 3. Section 4 describes the construction of sentiment word database for teaching evaluation. Section 5 presents the architecture of our proposed system. The next section presents the case study and results of the teaching evaluation. Finally, we conclude the paper in section 7.

II. RELATED WORKS

Sentiment analysis has been studied in wide area of domain such as movie review, teaching review [1] [4] [13], product review [11], e-learning[2], hotel review and many more. Most scholars focused to quantitative data analysis. However, some studies have been done on qualitative data using sentiment analysis[14], we found six works that mentioned the idea of using opinion mining and sentiment analysis in education.

El-Halees [1] proposed course evaluation model using students' attitudes posted on discussion forum written in Arabic language to improve course evaluation. Naive Bays, k-nearest and Support Vector Machine are used in opinion classification. To extract the features for each course, the author used Gate Information Extraction tool and Rapidminer data mining tool to classify the students' posts. The experimental result shows that Naive Bays method has better performance than the other two methods. Their system calculates the positive or negative orientation by computing the difference between the number of positive and negative adjective words. The author didn't consider the score of opinion words. This work suggests the effectiveness of user-generated content to improve course evaluation.

978-1-5090-5507-4/17/$31.00 ©2017 IEEE
ICIS 2017, May 24-26, 2017, Wuhan, China

Rashid [3] presents frequent features and opinion words extraction from students' feedback dataset on faculty performance evaluation using two pattern mining algorithm; e.g., sequential pattern algorithms (Apriori) and Generalized Sequential Pattern (GSP). This study considered Noun and Adjective extraction. They compared two algorithms. Their evaluated results showed that GSP is more significant than Apriori for frequent features and opinion words extraction.

Furthermore, Jagtap and Dhotre [6] demonstrate the effective method of automated sentiment analysis from teacher feedback assessment using Hidden Markova Model and Support Vector Machine. The authors only discussed the effective method concepts. There has been no experimental result.

LI Caiqiang and Ma Junming [10] perform online education teacher evaluation model based on opinion mining. They collect the students' comments written in Chinese language published in the LMS (learning management system) by using web crawler. They didn't indicate the words polarity strength in their work. If sentiment words of subjective text are not in the polarity word dictionary, they used pointwise mutual information (PMI) method to judge the polarity of it. This model gets an overall evaluation of each teacher.

Ortigosa, Martin and Carro [2] proposed other sentiment analysis approach for e-learning environment using a combined method of Spanish lexical based and machine learning techniques. In their experiment, combination of lexicon based techniques and SVM obtained the highest accuracy (83.27). In the context of e-learning, it is possible to extract information about the student's sentiments from the messages they write in Facebook, the students' sentiments towards a course can serve as feedback for teachers, especially in the case of online learning. However, this work still has the limitation of the sentiment analysis, all the words tagged as the same polarity get the same score, all positive words get the score = 1, all negative words get the score = -1 and all neutral words get the score = 0. They did not assign different weights to different words; bad and very bad are assigned the same score.

Similar to Pong-inwong and Rungworawut [9] proposed the construction of teaching evaluation sentiment lexicon for an automated sentiment polarity. Their work consists of two parts, first part is data preparation and second part is data modelling and evaluation. The authors chose ID3, Naïve Bayes and SVM to classify the teaching sentiment evaluation. Their experimental result shows that SVM has the highest accuracy of other two methods. In this work, the weight score of terms was defined by an expert who has experience in teaching evaluation. The sentiment weight was ranged from -1.00 to 1.00. Their proposed method couldn't consider the intensifier word and was constructed in Thai language.

As mention above it is possible to perform sentiment analysis in students' textual feedback. The application of sentiment analysis in students' comment was used in various objectives; teaching evaluation, course-online evaluation and teacher evaluation. The target of automatic sentiment analysis is improving the better accuracy result of sentiment classification and summarization. However, current works considered the sentiment analysis into three classes; positive, negative and neutral. None of the methods above considered the level of student's opinion.

III. Methods in Sentiment Analysis

There are two main approaches for sentiment analysis: machine learning based and lexicon based. Machine learning based approach uses classification technique to classify text. Lexicon based method uses sentiment dictionary with opinion words and match them with data to determine polarity. They assign sentiment scores to the opinion words describing how Positive, Negative and Objective the words contained in the dictionary are. Lexicon based approach involves calculating the orientation of words or phrases [16].

A. Lexicon Based Approach

In this approach, opinion words are divided in two categories. Positive opinion words are used to express some necessary things, and negative opinion words are used to describe unnecessary things.

B. Dictionary Based Approach

One of the simple techniques in this approach is based on bootstrapping using a small set of seed opinion words and an online dictionary, e.g., WordNet. The strategy is to first collect a small set of opinion words manually with known orientations, and then to grow this set by searching in the WordNet for their synonyms and antonyms. The newly found words are added to the seed list. The next iteration starts. The iterative process stops when no more new words are found. After the process completes, manual inspection can be carried out to remove and/or correct errors [7].

C. Corpus Based Approach

Using the corpus based approach alone to identify all opinion words, however, is not as effective as the dictionary-based approach because it is hard to prepare a huge corpus to cover all English words [7]. But it can help to find domain and context specific opinion words using a domain corpus which is the huge advantage of this method. The corpus-based approach is performed in statistical approach or semantic approach.

IV. Sentiment Word Database Construction

This section presents the construction of sentiment word database. The lexicon can be created either manually [12] or expanding automatically from a seed of words. In our lexicon the synonym is considered as a positive word and antonym is considered as a negative word [14]. In our work, the score of opinion words are defined by two experts who have experience in teaching evaluation. It is verified that the reliability of the

score of opinion words are given by a language expert from the Department of Languages, the University of Computer Studies, Mandalay.

There are 745 words in sentiment word database. Total intensifier words are 24, total positive words are 448 words and total negative words are 263 words, the remaining words are neutral words. The sentiment score ranges from -3 to +3. The score ranging from 1 to 3 are considered as positive; whereas any negative score ranging from -1 to -3. The sentiment score equals to zero is considered as neutral category. Some example words are shown in Table 1.

TABLE I. SAMPLE WORDS IN SENTIMENT WORD DATABASE

Example Opinion Words		
Opinion word	***Score***	***Description***
care	+2	verb
useful	+2	adjective
helpful	+2	adjective
clear	+2	adjective, verb
good	+2	adjective, verb
joyful	+1	adjective
marvellous	+3	adjective
brilliant	+3	adjective
ordinary	0	adjective
complex	-3	adjective
confuse	-3	verb
normal	0	adjective
complicated	-3	adjective
sleepy	-2	adjective
fast	-1	adjective
daily	0	adjective
most	+100%	Intensifier
slightly	-50%	Intensifier
really	+25%	Intensifier
little	-50%	Intensifier
very	+50%	Intensifier
easily	+25%	Intensifier

Same word in different domains can have different meanings. In our lexicon 'fast' is negative opinion and its polarity score is -1. For example 'I feel IPAD2 is fast running speed and higher screen resolution'. In this sentence 'fast' is the positive opinion in phone review domain but negative opinion in teaching review domain. For example 'I understand her teaching but speed of presentation is very fast for other students'. In this sentence 'fast' is negative opinion (-1) and 'very' depicts intensifier word that increased by 50%. The sentiment word database contains the scores of the polarity of every word. Some words are not included in existing lexicon. The polarity of some words written by students in teaching review needs to add in sentiment lexicon. Consider the following sentences in students' feedback.

"She can't explain the theory in detail."

"She doesn't have interactive teaching with her students while teaching."

"Our teacher is very helpful, knowledgeable. We can ask her about our lessons outside the class."

"She is a practiced teacher in C ++ programming."

"His calculation idea in programming is complex."

"We feel sleepy while he teaching."

"She always gives complicated questions in exam."

"Her teaching speed is fast for other student but I understand her teaching."

"Her voice is very low."

"Available outside the class"

"Over serious while explaining."

The words "explain, interactive, ask, understand, practiced, knowledgeable, available, slow, fast, talented, complex, complicated, sleepy, overcome, concern, etc." need to consider as opinion words to get the correct opinion result. Some words are not included in Afinn lexicon [5]and Vader lexicon.

1) Negation words: The negation words are the words which reverse the polarity of sentiment[15], changing good (+2) into not good (-2). (e.g. no, not, neither, nor, nothing, never, none) are very important in identifying the sentiments, as their presence can reverse the polarity of the sentence. In the proposed approach, whenever a negation word is encountered in a comment, the opinion score is reversed by a certain amount.

2) Blind negation words: (e.g. Need, needed, require, required) are also very important in identifying the sentiments. For example 'Her teaching method needed to be better', 'better' depicts a positive sentiment but the inclusion of the blind negation word 'needed' suggests that this sentence is really depicting negative sentiment. In the proposed approach whenever a blind negation word occurs in a sentence its polarity is immediately labeled as negative and assigned the opinion score to (-2).

3) Adjective, adverb, verb, noun words: Most of the opinion words are adjective. Our dictionary contains several different adjectives, adverb, verb and noun. For example 'She is knowledgeable. Her discussions are interesting. I understand her

teaching'. In this sentence 'knowledgeable' and 'interesting' are positive adjective opinion words, 'understand' is a positive verb opinion.

4) Intensifier words: They are classified into two major categories, depending on their polarity. Amplifiers (e.g., very) increase the semantic intensity of a neighboring lexical item, whereas down toners (e.g., slightly) decrease it [12]. For example, "Her explanation is really very good". In this sentence 'really and very' are intensifiers that increase the positive sentiment polarity.

V. System Architecture

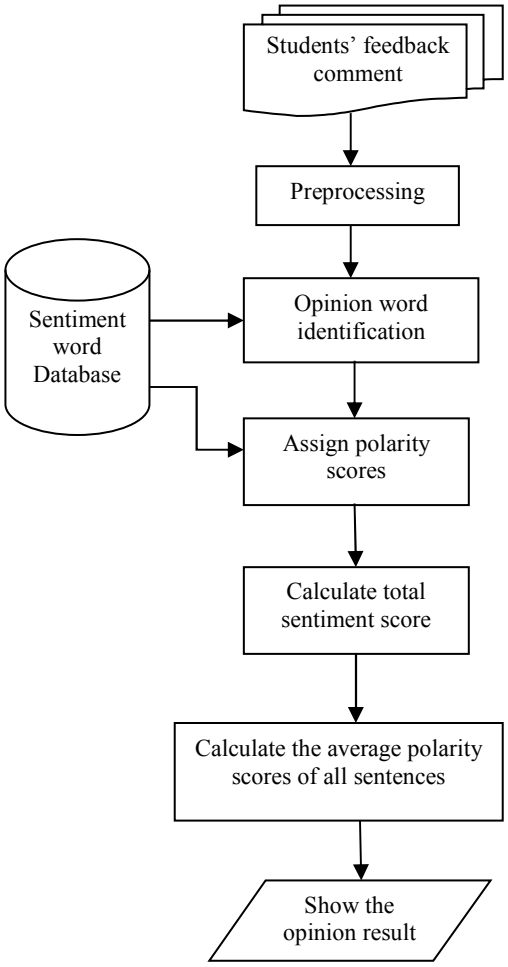

Fig. 1. The model of sentiment analysis

Fig. 1 shows the model for sentiment analysis of students' feedback comments using lexicon based approach. Initially, we gathered comments from the Faculty Evaluation Survey for University of Computer Studies, Mandalay, Myanmar. We used this comments as an input. Input texts were the students' comments for their teachers, which were composed of one or more sentences connected to a course and teaching style. The next step is preprocessing that removes the unwanted and noisy data. This step includes case conversion, split sentences, replace "n't" with "not", punctuation removal, tokenization.

The next step is opinion word identification. This system compares each tokenized word in the comment with blind negation word or negation word or positive opinion word or negative opinion word or intensifier word by using sentiment word database. And then the polarity scores are assigned to each word by using sentiment word database. The presence of the blind negation word indicates negative sentiment value (-2).

A heuristic technique is used to calculate the semantic orientation score of combining words for automated students' feedback comments analysis. In the following equations, 'Ws' is the semantic orientation score of combining words. 'Sinf' is the intensifier value of word based on 100%. 'Os' is the score of opinion word from sentiment word database.

$$Ws=Os \quad (1)$$

$$Ws=(100\%+Sinf) * Os \quad (2)$$

$$Ws=(100\%+Sinf) * (100\%+Sinf) * Os \quad (3)$$

$$Ws=Ws * (-1) \quad (4)$$

If there is only one opinion word in a sentence, the corresponding positive scores or negative scores are assigned using (1). If one intensifier word and one opinion word are found together, i.e. the location of intensifier word is adjacent with the location of opinion word, (2) is used to get the semantic orientation score of combining words. If two intensifier words and one opinion word are found in a sentence, moreover the index of first intensifier word must be the index by reducing two of the index of opinion word, (3) is used. If a negation word in front of the opinion word is found in a sentence, reversed polarity scores are given by (4).

The semantic orientation score of combining words in all sentences are summed up to get the total polarity scores by (5). In (5), 'PTs' is the total polarity score of all words in all sentences from one comment. 'Wsi' is the semantic orientation score of combining words for i^{th} term in one comment. 'i' is the order of combining opinion words from 1 to n. 'n' is the total number of combining opinion words in all sentences from one comment. 'T' is set of teaching sentiment terms from sentiment word database. 'PTsi' is the total polarity score of i^{th} term for all comments. 'N' is the total number of opinion words in all comments. 'P' is the average polarity scores of all comments.

$$PTs = \sum_{i=1}^{n} Wsi, (Wsi \in T) \quad (5)$$

$$P = \sum_{i=1}^{N} PTsi \ /N \quad (6)$$

The average polarity score of all comments can be calculated by (6).

If P>0 or <=1, the opinion result is defined as weakly positive. If P>1 or <=2, the opinion result is defined as

moderately positive. If P>2, the opinion result is defined as strongly positive. If P<0 or >= (-1), the opinion result is defined as weakly negative. If P< (-1) or >= (-2), the opinion result is defined as moderately negative. If P< (-2), the opinion result is defined as strongly negative. If P=0, the opinion result is defined as neutral.

VI. Case Study and Results

We tested the effectiveness of this system using the feedback comments of fourth year students from BIS (Business Information System) specialized students, HPC (High Performance Computing) specialized students, SE (Software Engineering) specialized students, KE (Knowledge Engineering) specialized students in our University. In this section, three comments are tested for sample calculation based on the same teacher. The opinion result of all comments is classified as strongly negative, or moderately negative, or weakly negative, or strongly positive, or moderately positive, or weakly positive or neutral.

Comment 1: She is very knowledgeable. Her discussions are very interesting and easily understand.

Comment 2: She is a helpful teacher. But her teaching speed is very fast. Her voice is also very low.

Comment 3: Her explanations are clear. Her class notes and outlines are very useful. She gives enjoyable discussion.

There are three opinion words in comment 1. The opinion score calculation of comment 1 is very knowledgeable (2× [100%+50%])=3 by (2), very interesting (2× [100%+50%])=3 by (2), easily understand (2× [100%+25%])=2.5 by (2). The total opinion scores in comment 1 is 3+3+2.5=8.5 by (5).

There are three opinion words in comment 2. The opinion score calculation of comment 2 is helpful =2 by (1), very fast (-1× [100%+50%]) =-1.5 by (2), very low (-2× [100%+50%]) = (-3) by (2). The total opinion scores in comment 2 is 2+ (-1.5) + (-3) = (-2.5) by (5).

There are three opinion words in comment 3. The opinion score calculation of comment 3 is clear =2 by (1), very useful (2× [100%+50%]) =3 by (2), enjoyable=1 by (1). The total opinion scores in comment 3 is 2+ 3 + 1 = 6 by (5).

So total opinion scores of all comments are 8.5+ (-2.5) +6=12 by (5) and divided by total number of opinion words in all comments. The result is 12/9=1.3333 by (6). So the opinion result of three feedback comments about the teacher is moderately positive.

We tested the students' comments of some teachers from our University using our lexicon and Afinn lexicon [5]. Among them, the opinion results of twelve teachers are displayed in Fig.2.

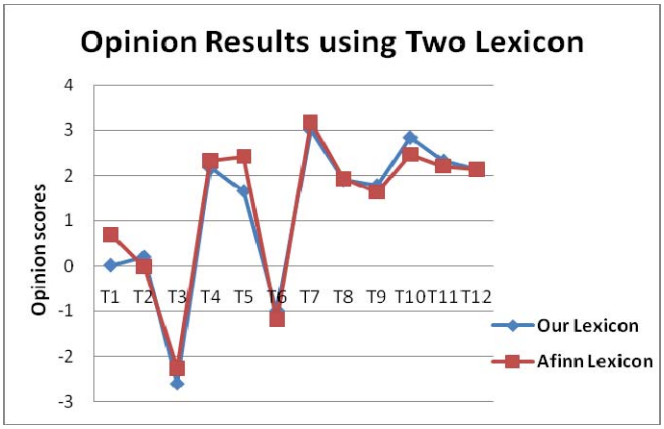

Fig. 2. Opinion result of teachers using two lexicon

Afinn is an English word list rated for valence with an integer between -5 (negative) and +5 (positive). The words in Afinn have been manually labeled by Finn Arup Nielsen in 2009-2011[5]. Afinn lexicon can't identify the students' comment including the words, "explain, interactive, ask, understand, practiced, knowledgeable, available, slow, fast, talented, complex, complicated, sleepy, overcome, etc.". Some of the above words are included in the reviews of teacher 2. So the opinion result of teacher 2 using Afinn lexicon is neutral. The opinion result of teacher 2 using our lexicon is weakly positive. The opinion result depends on the opinion words in students' comments. Some scores of opinion words in our lexicon and Afinn lexicon are a little different. If the scores of opinion words are the same in both lexicons, the opinion result is the same. If the scores of opinion words in both lexicons are different slightly, the average polarity score will be different. Although the average polarity scores of teacher 1,3,4,7,8,9,10,11 are slightly different from using two lexicons, the opinion results are same because these scores are found within the predefined ranges. The opinion result of teacher 2, 5, 6 using two lexicons are different because the scores of opinion words are different. Table 2 shows the average polarity score of twelve teachers using our lexicon and Afinn lexicon.

TABLE II. Comparison of Average Polarity Score Using Two Lexicon

Average Polarity Score of all Comments		
Teacher Name	***Average Polarity Score Using Our Lexicon***	***Average Polarity Score Using Afinn Lexicon***
Teacher 1	0.02 (weakly positive)	0.69 (weakly positive)
Teacher 2	0.2 (weakly positive)	0.0 (neutral)
Teacher 3	-2.6 (strongly negative)	-2.25 (strongly negative)
Teacher 4	2.17 (strongly positive)	2.33 (strongly positive)
Teacher 5	1.65 (moderately positive)	2.42 (strongly positive)
Teacher 6	-1 (weakly negative)	-1.17 (moderately negative)

Average Polarity Score of all Comments		
Teacher Name	***Average Polarity Score Using Our Lexicon***	***Average Polarity Score Using Afinn Lexicon***
Teacher 7	3.0 (strongly positive)	3.17 (strongly positive)
Teacher 8	1.9 (moderately positive)	1.93 (moderately positive)
Teacher 9	1.78 (moderately positive)	1.64 (moderately positive)
Teacher 10	2.83 (strongly positive)	2.46 (strongly positive)
Teacher 11	2.32 (strongly positive)	2.21 (strongly positive)
Teacher 12	2.14 (strongly positive)	2.14 (strongly positive)

VII. CONCLUSION

In this paper, lexicon based sentiment analysis is used to evaluate the level of teaching performance from students' textual feedback comment. A database of English sentiment words is constructed to identify the polarity of words as a lexical source. Our sentiment word database contains the opinion words concerning with the academic domain to achieve the better result. Every opinion word in the database is given a value. The sentiment value is ranged from -3 to +3. This paper proposes the level of teaching evaluation method based on lexicon based approach. This method analyzes automatically the students' feedback comments to strongly negative, or moderately negative, or weakly negative, or strongly positive, or moderately positive, or weakly positive or neutral category using two lexicons. The level of opinion result for any teacher is given out from students' feedback comments.

References

[1] A. El-Halees, "Mining opinions in user- generated contents to improve course evaluation," Software Engineering and Computer Systems, pp. 107-115, 2011.

[2] A. Ortigosa, J.M. Martín and R.M. Carro, "Sentiment analysis in Facebook and its application to e-learning," Computers in Human Behavior, 31(1), pp. 527-541, 2014.

[3] A. Rashid, "Feature level opinion mining of educational student feedback data using sequential pattern mining and association rule mining," International Journal of Computer Applications (0975-8887). 81 (10), pp. 31-38, 2013.

[4] A. Sagum, Jessiree Grace M. de Vera, Prince John S. Lansang, Danica Sharon R. Narciso and Jahren K. Respeto, "Application of Language Modelling in Sentiment Analysis for Faculty Comment Evaluation," Proceedings of the International Multi Conference of Engineers and Computer Scientists, vol.1, March 2015.

[5] AFINN word database an affective lexicon by Finn Årup Nielsen, http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6010/zip/imm6010.zip.

[6] B. Jagtap and V. Dhotre, "SVM and HMM based hybrid approach of sentiment Analysis for Teacher Feedback Assessment," International Journal of Emerging Trends of technology in computer science. (IJETCS). 3 (3), pp. 229-232, 2014.

[7] B. Liu, "Sentiment analysis and subjectivity," Handbook of natural language processing, pp. 627-666, 2010.

[8] C. Pong-inwong and W. Rungworawut, "Teaching Evaluation Using Data Mining on Moodle LMS Forum" , 6th International Conference on New Trends in Information Science, Service Science and Data Mining (Issdm2012), Taipei: IEEE, 2012, pp. 550–555.

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

[10] L. Caiqiang, and M. Junming, "Research on online education teacher evaluation model based on opinion mining," National Conference on Information Technology and Computer Science (CITCS), 2012.

[11] 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, 2014, pp. 277-281, 2014.

[12] M. Taboada, J. Brooke, M. Tofiloski, K. Voll and M. Stede, "Lexicon-based methods for sentiment analysis," Association for Computational Linguistics, vol. 37, Issue 2, pp. 267-307, June 2011.

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

[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), pp. 179-189, 2015.

[15] P. Palanisamy, V. Yadav, H. Elchuri, "Serendio: Simple and Practical lexicon based approach to Sentiment Analysis," Seventh International Workshop on Semantic Evaluation (SemEval 2013), (2), pp. 543-548, 2013.

[16] P.T.Ngoc, M.Yoo, "The Lexicon-based Sentiment Analysis for Fan Page Ranking in Facebook," 2014 IEEE International Conference on Information Networking (ICOIN), Phuket, pp. 444-448, 2014.