Evaluation of the SA-E System for Analysis of Students' Real-time Feedback

Nabeela Altrabsheh[*], Mihaela Cocea[†], Sanaz Fallahkhair[‡] and Khaldoon Dhou[‡]

[*] College of Business and Economics, Qatar University, Doha, Qatar, Email: nabeela@qu.edu.qa

[†] School of Computing, University of Portsmouth, UK, Email: mihaela.cocea@port.ac.uk

[‡] School of Computing, University of Brighton, UK

[§] Department of Software and Information Systems, UNC Charlotte, Charlotte, USA

Abstract—Students' real-time feedback is acknowledged as an important source of information for teachers/lecturers to improve their teaching and address issues students may have, such as going deeper in some of the materials covered or providing more examples to understand an abstract concept. Previous applications collecting real-time feedback from students through clickers and mobiles typically collect limited information with pre-defined questions, while more recent applications using social media collect such a large volume of information that a lecturer cannot manually process it in real time. We developed the SA-E system for analysing students' real-time feedback provided via social media, and, in this paper, we present the evaluation of this system in real settings with lecturers and students. The results show that lecturers are highly satisfied with the proposed system. In contrast, although the participation of students in providing feedback was high, the students' opinions of the system were between neutral and dislike.

Index Terms—Sentiment Analysis, Technology Enhanced Learning, Students' Feedback

I. INTRODUCTION

Twitter was used previously as a tool to collect feedback in the classroom [1]. The lecturer had to read through all the students' tweets sequentially from the beginning to understand them, causing time loss; they would also require additional training to effectively use the tweets to provide feedback.

Our previous research addressed the issue of analysing textual feedback in real time by developing sentiment analysis models using machine learning techniques to provide two main outputs: (a) the distribution of feedback in terms of polarity, i.e. positive/negative/neutral, and (b) the presence in the feedback of particular emotions related to learning, e.g. excitement and boredom [2], [3]. We also created six visualisations to present the results to the lecturers. We integrated the sentiment analysis models and the visualisations into a PC-based system to be used in real lectures, and the opinions of the students and lecturers were collected and analysed. In this paper, we report the results of an evaluation study.

II. SYSTEM ARCHITECTURE AND EVALUATION STUDY

The system consists of the sentiment analysis models which predict polarity and emotion from students' feedback, and visualisations that present the results to the lecturer. The system was built and integrated using R language. The system is PC-based (including desktop and laptop-based), as feedback from lecturers indicated that most of them (92%) prefer these to tablets and smart phones (8%).

The results from the sentiment analysis models are fed into the visualisation module of the system, which displays the distribution of feedback according to polarity, as well as the presence of three emotions, i.e. amusement, excitement and boredom [3]. In addition, frequent words associated with each polarity and each emotion are also displayed.

The use of the system in real settings was investigated in terms of usefulness and usability; in addition, we were interested in the students participation levels, as an essential part in the use of the system. The participants consisted of 80 students and 4 lecturers from the computer department at an U.S. Institute, which cannot be named for ethical and legal purposes (i.e. they requested to remain anonymous). The data was collected during 6 lectures.

The purpose of this study was explained to the students by their lecturers. The students were asked to provide feedback via Twitter throughout the lecture or at set times during the lecture, depending on the lecturer's preference. They were provided with an anonymous Twitter account to tweet or they could use their own accounts.

The students' questionnaire consisted of four questions (see Table I). The first question was a Yes/No question asking students if they participated in providing feedback. The three remaining questions were Likert type questions asking how useful the system is and if they would adapt to it.

The lecturer questionnaire consisted of 13 Likert questions with the same 5-point scale as the students (see Table II). We also asked lecturers how often they used the system in each lecture ("beginning, middle and end", "only end" or "all throughout the lecture"). Lastly, we asked lecturers if they found any difficulties using the system.

III. RESULTS

The first question explored if students participated in providing feedback or not. The results show that one student answered 'No' to this question and two students left the answer blank; thus 85% of the students (out of all 91 students) declared that they provided feedback.

To find if the students liked or disliked the use of the system, one sample t-tests were used for both the composite score of the questionnaire (i.e. average across all questions),

TABLE I
ONE SAMPLE T-TEST RESULTS ON STUDENT'S ANSWERS.

	Mean	Std	p-Value
Q2: I feel that the lecturer changed in the lecture according to my feedback.	2.70	1.17	**0.025**
Q3: I like this new teaching style.	2.68	1.34	**0.033**
Q4: I think the system is useful for me.	2.70	1.40	0.059
Composite Score	2.69	1.21	**0.026**

TABLE II
ONE-SAMPLE WILCOXON SIGNED RANK TEST COMPARING THE
LECTURERS' ANSWERS TO THE NEUTRAL VALUE.

	median	mean	p-value
Overall, I am satisfied with the use of the system.	4	4.17	**0.02**
The system was easy to use.	5	4.83	**0.02**
The visualisations provided are easy to understand.	4	4.33	**0.02**
I feel that the system gave an accurate and correct judgement about the classes opinions.	4	4.33	**0.02**
I can integrate the system into my everyday teaching.	4	4.00	**0.01**
The system helps me understand students more.	5	4.83	**0.02**
The system in general is not a distraction in class.	4	3.50	0.18
I found the polarity detection useful.	4	4.33	**0.02**
I found the system useful.	5	4.83	**0.02**
Students using their mobiles are not a distraction in class.	4	4.16	**0.02**
I found the emotions detection useful.	4.5	4.50	**0.02**
I found the visualisations useful.	4.5	4.50	**0.02**
I changed my teaching based on the systems results	4	3.67	0.10

and the individual items. The one sample t-test allows the comparison between the mean of a sample (i.e. our students' answers) and a particular value – the middle of the Likert scale was used, i.e. the value of 3, which corresponds to a neutral position. The results show that the students have a relatively negative impression of the system. Significant differences can be observed for the composite score, Q2 and Q3. For Q4 (i.e. system is useful for them) the differences are not significant, indicating that the students have a neutral view of the usefulness of the system.

We also investigated the presence of variation across the 6 lectures, i.e. if students liked/disliked the use of the system in some lectures, but not in others. One sample Wilcoxon Signed Rank Test, a non-parametric alternative to the one sample t-test, was used due to the small amount of data distribution in each lecture. Similar to the previous experiment, the value 3 was used for the sample comparison.

For three of the lectures the median was 3, indicating a neutral view about the system; these lectures, i.e. lectures 1, 3 and 5, were all for students in year 2. The other three lectures, i.e Lectures 2, 4 and 6, the median was 2, indicating a relatively negative opinion about the system; these lectures were for students in year 1 and 3. For Lecture 2, with students in year 1, both the composite score and the individual questions were significantly different from the neutral value of 3, showing that the students in this lecture had the strongest negative view out of all lectures.

For all the other lectures the differences were not significant (with one exception discussed in the following), indicating that most students have a neutral opinion about the use of the system. In Lecture 4, a significant difference occurred for Q2, which is about the lecturer making changes in their teaching according to the feedback from the students; this difference suggests that the students did not see changes made in the lecture as a consequence of their feedback.

For the lecturer questionnaire, four lecturers answered the questionnaire for each lecture when they used the system. Two lecturers used the system in 2 subsequent lectures on the same topics. Consequently, there were 6 records for the responses to the lecturers' questionnaire.

For most questions, the difference from the neutral value is significant and positive, thus indicating that the lecturers had a positive view of the system. For two questions there was no significant difference – one referring to the system being a distraction and one about changes introduced in the lecture based on the information displayed in the system. These suggest that the lecturers have to allocate some of their attention to the system and that they made little changes to

their teaching based on the students' feedback.

The lecturers were also asked if they used the system: (a) throughout the lecture (i.e. allowing students to tweet at any time), (b) at the beginning, middle and end of the lecture, or (c) only at the end of the lecture. 17% used the system at the beginning, middle and end of the lecture and 83% chose to use the system all throughout the lecture. Lastly, the lecturers were asked if any part of system was difficult to use and 50% answered no, while the rest did not include any answer.

IV. CONCLUSION AND FUTURE WORK

This paper presented the evaluation of the SA-E system for analysis of students' real-time feedback. The system was evaluated in six lectures, and opinions from both the students and the lectures were collected via questionnaires.

The students answers showed that 85% of the students provided feedback via Twitter, showing their willingness to give feedback in real time. Their opinions about the use of the system were between neutral and dislike. In particular, they did not see any changes made by the lecturers according to their feedback, which was confirmed by the lecturers. This indicates that the "feedback loop" needs to close by informing the students how their feedback is used. In future research we will investigate different ways to feed back to the students.

The lecturers were very positive about the system – significant positive differences were obtained for 11 out of the 13 questions. The lecturers, however, did not make any changes according to the students' feedback, which could be due to predominantly positive feedback or the lack of familiarity with the system. In future research, we will investigate the use of the system on a longer term basis and the way lecturers process and act based on students' feedback.

REFERENCES

[1] J. Novak and M. Cowling, "The implementation of social networking as a tool for improving student participation in the classroom," in *ISANA Int. Academy Association Conf. Proceedings*, 2011, pp. 1–10.
[2] N. Altrabsheh, M. Cocea, and S. Fallahkhair, "Sentiment analysis: towards a tool for analysing real-time students feedback," in *IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI)*, 2014, pp. 419–423.
[3] ——, "Predicting learning-related emotions from students textual classroom feedback via twitter," in *The 8th Int. Conf. on Educational Data Mining (EDM)*, 2015, pp. 436–440.