Exploring indicators from keyboard and mouse interactions to predict the user affective state

Sergio Salmeron-Majadas
aDeNu Research Group. UNED
Calle Juan del Rosal, 16. Madrid
28040. Spain
+34 91 398 93 88
ssalmeron@bec.uned.es

Olga C. Santos
aDeNu Research Group. UNED
Calle Juan del Rosal, 16. Madrid
28040. Spain
+34 91 398 93 88
ocsantos@dia.uned.es

Jesus G. Boticario
aDeNu Research Group. UNED
Calle Juan del Rosal, 16. Madrid
28040. Spain
+34 91 398 91 97
jgb@dia.uned.es

ABSTRACT
Following a low cost and non-intrusive approach, in this paper we discuss how prediction rates from 5 different data mining algorithms using 4 different emotional labeling approaches differ when exploring the usage of keyboard and mouse interaction sources for affective states detection in a math problem solving experiment.

Keywords
Data Mining, Affective Computing, Affective States, Human-Computer Interaction, Keyboard, Mouse.

1. INTRODUCTION
Due to the existing relations between emotions and cognitive processes in learning, there is a need to take into account the learners' affective state when supporting the learning process [7]. With this context in mind, in this paper we explore the potential of using mouse and keyboard interaction data as affective information sources, which are low cost and non-intrusive. We compare the results obtained from them with those provided by alternative data sources, such as sentiment analysis and physiological signals.

The most common approach reported in the literature regarding emotion detection is based on using a single data source as affective indicator [2, 12]. Usually, keyboard and mouse interactions as well as physiological sensors are used. Regarding keyboard, keystroke features extracted from single events are used to detect affective states [2], although combined keystroke events indicators have also been considered [4]. On the mouse side, some works have used features such as speed or direction to detect affective states [12]. A review of different studies carried out to detect emotions from keyboard and mouse interactions can be found in [6]. Physiological sensors have been widely used with affective purposes, but usually using intrusive ways to get data [5].

2. EXPERIMENT & RESULTS
A math problem solving experiment was carried in our lab with 75 participants (details in [11]) in order to research how to detect affective states with data mining [9]. To gather emotional data we used different data sources: keyboard interactions (***K***), mouse interactions (***M***), webcam recording, computer screen recording, Kinect recording and physiological recording (i.e. heart rate, skin conductance, breath frequency and skin temperature) (***P***). The experiment collected participants' emotional baseline. The mathematical tasks consisted of 3 series of 6 problems. For each problem, participants had to select one answer from a set of 4 possibilities and fill in the 9-point Self-Assessment Manikin (SAM) [3] scale to report their valence (i.e. pleasure) and arousal (i.e. activation) state. After each group of problems (task), participants had to type their feelings about it. Emotions were elicited by giving less time than required to do some tasks, or changing their difficulty level. All along the experiment each participant had an affective tutor, who supervised the progress and took timestamps on the physiological recordings on every task beginning.

Representing the affective states occurred during a session is an open issue [8], so several approaches to emotionally label interactions were considered: i) SAM scores provided by participants during the experiment (***Label 1***), calculating the mean and standard deviation for each task; ii) SAM scores provided for each task by two psychologists (with experience in motivational and educational issues) after reading the corresponding emotional reports (***Label 2***); iii) a categorical classification (positive, negative, neutral and positive-negative) provided by another expert (with 10 years of experience in supporting learners in e-learning platforms) when reading those emotional reports (***Label 3***), and iv) the average value from the 9-point SAM scores per task given by the participant and the psychological experts (***Label 4***).

For data processing, indicators were grouped by task. For keyboard interactions, depending on the event aggregation performed, the indicators generated were the following: i) number of key press events, ii) average time between press events, iii) average time between a press and its following release event and iv) number of times a certain key or a group of keys has been pressed (backspace key, delete key, alphabetical characters keys, etc), and v) the indicators proposed in [4], which were generated from creating combinations of two or three keystrokes events. On the mouse interactions side, indicators were as follows: i) number of clicks (per button and aggregated); ii) overall distance; iii) covered distance (distance the cursor has traversed) between two button press events, between a button press and its following release event, between a button release and its following press event and between two button release events; iv) the Euclidean distance in the four previously described cases; v) the difference between the covered and the Euclidean distances calculated; and vi) time durations between the proposed combinations of events.

When processing the physiological signals, differences between the values in each task and the baseline value for each signal were calculated and used to compute the average for all those values. Additionally, sentiment analysis (***S***) was used to automatically generate an affective score for each emotional report, counting the number of positive and negative terms according to the MPQA Opinion Corpus affective database.

Following previous works [10], our goal here was to predict the valence dimension as higher correlations were found with valence than with arousal. As suggested in the literature [1], the 9-point valence values were grouped into three categories (i.e., positive (>6), negative (<4) and neutral (4-6)). Different algorithms were used, namely C4.5 (C), Naïve Bayes (N), Bagging (B), Random Forests (R) and AdaBoost (A). Results in Table 1 show the best prediction rate depending on the labelling and the data source used and the algorithm applied to achieve that rate. The analysis was done on the data from 17 participants, who are the ones whose interactions have already been emotionally labeled with the four aforementioned approaches. When processing the data, some filtering decisions were taken, such as removing the registers with SAM values per task with a standard deviation higher than 2, as well as the registers corresponding to neutral and positive-negative categories.

Table 1. Best prediction rates depending on the labels and the input data sources. Best result per data labeling is bolded.

	Label 1	Label 2	Label 3	Label 4
K	0,65 (C)	0,74 (B)	0,58 (C)	0,67 (R)
M	0,65 (C)	0,74 (B)	0,57 (R)	0,67 (R)
S	0,82 (R)	**0,83 (A)**	**0,66 (A)**	0,81 (C,B,A)
K+M	0,67 (R)	0,74 (B,R)	0,59 (R,A)	0,56 (R)
K+S	0,75 (C)	0,74 (B)	0,64 (R)	**0,86 (C)**
M+S	**0,85 (C)**	0,74 (B)	0,6 (A)	0,81 (R)
K+M+S	0,75 (B,R)	0,74 (B)	0,62 (B)	0,77 (A)
P	0,67 (C)	0,74 (B)	0,52 (C,R)	0,53 (C)

3. DISCUSSION & FUTURE WORK

From Table 1, sentiment analysis seems to be the best data source, but its results can be improved when combined with keyboard or mouse. This suggest that combining different data sources would produce improvements, but this should be clarified with further experiments as using different prediction algorithms and alternative labeling approaches seem to induce significant differences in the results. Up to our knowledge, there are no works in the literature that report a deep comparison of the benefits of each labeling approach. Due to this, it seems of interest to study different approaches to label emotions by performing a comparative analysis using a large number of algorithms depending on their predictive features (using feature selection techniques). Another future step of interest to take is exploring the idea of mining these sources separately and then mining the obtained outputs, in search for a system that would be able to automatically choose the data source to be used depending on their individual success in the prediction.

ACKNOWLEDGMENTS

Authors would like to thank MAMIPEC project (TIN2011-29221-C03-01) colleagues for their support and the Spanish Government for its funding. Sergio Salmeron-Majadas acknowledges the FPI grant (BES-2012-054522).

REFERENCES

[1] Baran, B., Pace-Schott, E.F., Ericson, C. and Spencer, R.M.C. 2012. Processing of Emotional Reactivity and Emotional Memory over Sleep. *The Journal of Neuroscience*. 32, 3 (Jan. 2012), 1035–1042.

[2] Bixler, R. and D'Mello, S. 2013. Detecting Boredom and Engagement During Writing with Keystroke Analysis, Task Appraisals, and Stable Traits. *Proceedings of the 2013 International Conference on Intelligent User Interfaces* (New York, NY, USA, 2013), 225–234.

[3] Bradley, M.M. and Lang, P.J. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*. 25, 1 (1994), 49–59.

[4] Epp, C., Lippold, M. and Mandryk, R.L. 2011. Identifying emotional states using keystroke dynamics. *Proceedings of the 2011 annual conference on Human factors in computing systems* (2011), 715–724.

[5] Handri, S., Yajima, K., Nomura, S., Ogawa, N., Kurosawa, Y. and Fukumura, Y. 2010. Evaluation of Student's Physiological Response Towards E-Learning Courses Material by Using GSR Sensor. *Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on* (2010), 805–810.

[6] Kolakowska, A. 2013. A review of emotion recognition methods based on keystroke dynamics and mouse movements. *2013 The 6th International Conference on Human System Interaction (HSI)* (2013), 548–555.

[7] O'Regan, K. 2003. Emotion and e-learning. *Journal of Asynchronous learning networks*. 7, 3 (2003), 78–92.

[8] Porayska-Pomsta, K., Mavrikis, M., D'Mello, S., Conati, C. and Baker, R.Sj. 2013. Knowledge elicitation methods for affect modelling in education. *International Journal of Artificial Intelligence in Education*. 22, 3 (2013), 107–140.

[9] Salmeron-Majadas, S., Santos, O.C. and Boticario, J.G. 2013. Affective State Detection in Educational Systems through Mining Multimodal Data Sources. *6th International Conference on Educational Data Mining* (Memphis, 2013), 348–349.

[10] Salmeron-Majadas, S., Santos, O.C. and Boticario, J.G. 2014. An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context. *18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems* (Poland, 2014), in press.

[11] Saneiro, M., Santos, O.C., Salmeron-Majadas, S. and Boticario, J.G. 2014. Towards Emotion Detection in Educational Scenarios from Facial Expressions and Body Movements through Multimodal Approaches. *The Scientific World Journal*. 2014, (Apr. 2014), e484873.

[12] Tsoulouhas, G., Georgiou, D. and Karakos, A. 2011. Detection of Learner's Affective State Based on Mouse Movements. *Journal of Computing*. 3, 11 (Nov. 2011), 9–18.