Data Sharing: Low-Cost Sensors for Affect and Cognition

Keith Brawner
United States Army Research Laboratory
12423 Research Parkway
Orlando, Florida, 32826
Keith.W.Brawner@us.army.mil

ABSTRACT

The Educational Data Mining (EDM) community has experienced many benefits from the open sharing of data. Efforts such as the Pittsburgh Science of Learning Center Datashop have helped in the development of learning data storage and standards in the educational community. In other fields, standards of comparison have been created through publication, sharing, and competition on identical datasets. This ability to share, compare, and grow as a field has proven to be a success. This paper presents a new and unique dataset, and shares it with the EDM community. Initial offline analysis results and secondary online analysis results are presented as benchmarks for comparison by future researchers.

Keywords

Data mining, machine learning, data sharing, affect, cognition

1. INTRODUCTION

The Army Research Laboratory (ARL) Learning in Intelligent Tutoring Environments (LITE) Lab has an interest in Intelligent Tutoring Systems (ITS) research, and has developed the Generalized Intelligent Framework for Tutoring (GIFT) [10] as an architectural output for research. GIFT is composed of several interoperable modules for the communication of sensor, learner, instructional, and performance information, with projects involved in each area. As part of a project involving sensors and learner data, an interesting and unique dataset was collected.

The purpose of this work is to share this collected data for the purpose of ITS development with the research community at large. Among the goals of the GIFT project is to be able to rapidly transition research into the community. Transition tools, authoring tools, and multiple programming language plugins have been constructed for this purpose, are curated to ensure overall stability and use, and are freely and publicly distributed [2]. The purpose of the research described as part of this data-sharing paper was to determine the effectiveness of low-cost sensors, and to test alternative modeling techniques. It is clear that this dataset can answer additional research questions of interest to the EDM community, and will be shared publicly at http://litelab.arl.army.mil/public.

2. HARDWARE

In total, measurements were collected via two Electroencephalography (EEG) systems (from Neurosky and Advanced Brain Monitoring (ABM)), a custom-made eye tracker, a Zephyr heart rate monitor, embedded Phidget pressure sensors within the chair, a Venier sonar sensor for distance from the computer, and emotional self-report measure. The self-report measure of EmoPro and the ABM headset have previously been validated to produce accurate measures of affective and cognitive states, respectively [5; 7]. A summary of the measures which these sensors produce is included in Table 1. Larger discussion on the relevance of each of these states to learning outcomes and validation of the baseline measurements is available in prior work [3; 4; 8].

Table 1. Summary of sensors used and states measured.

Sensor	Affective State	Cognitive State
*ABM EEG**		Attention, Engagement, Distraction, Drowsiness, Workload
Neurosky EEG		
Eye-tracker		Attention, Drowsiness, Workload
*EmoPro**	Anger, Anxiety, Arousal, Boredom, Fear, Stress	Attention
Heart Rate Monitor		
Chair Pressure Sensor (posture)	Arousal, Boredom, Frustration	Engagement, Flow
Motion Detector (posture)		
** Indicates Ground Truth Measurement*		

3. INITIAL ANALYSIS

The Logistic Model Tree (LMT) method of analysis [9] was selected for classifier construction on this data from among a series of methods considered [8]. Ten-fold cross validation at the class level was used in an effort to avoid overfitting. The created trees were found to have a single node, rendering this method similar to logistic regression. The measure of Area Under the Curve (AUC) of the Receiver Operator Characteristic (ROC) [6] is used to evaluate overall model quality. In general, the AUC ROC method produces a value in the range [0,1], with 1.0 representing perfect classification accuracy and 0.5 representing baseline levels. The overall finding is that there is significant room for improvement of generalized model quality, but that data trends are available to do so. These findings are summarized with Table 2 and Table 3.

Table 2. Summary of which sensor data was used to create Initial Emotional Models

Low-Cost Sensor	EmoPro Measurements		
	Anger	Anxiety/Fear	Boredom
HR			X
Eye Track			
EEG		X	X
Chair		X	
Distance		X	X
AUC ROC	**N/A**	**0.83**	**0.79**

Table 3. Summary of which sensor data was used to create Initial Cognitive Models

Low-Cost Sensor	ABM Measurements		
	Engagement	Distraction	Workload
HR	X	X	
Eye Track			
EEG			
Chair	X	X	X
Distance	X		X
AUC ROC	**0.80**	**0.81**	**0.82**

Later projects investigated a realtime signal approach to data processing for the classification of emotional states in realtime. There is some evidence that adaptable approaches among cognitive state data are able to model more accurately, but there remain few attempts to model states in this way [1]. Additionally, there is evidence that models created from bodily sensor data may fail generalization tests for reasons such as electrode drift, changes in default impedance, and other non-linear behavioral factors [1]. The core idea of this approach is that highly adaptable and individualized approaches to modeling would be better able to model emerging states at the student level. This was found to be true for affective measurements, but not for cognitive measurements (without further feature detection).

The total of these efforts is the development of realtime algorithmic approaches which are able to classify with very little labeling information. These approaches can be compared side-by-side to the binary classification, regression-based, logistic model trees created in the earlier study. Using methods for individualized realtime model construction on multiple individuals provides evidence to how well the model is likely to transfer to a new population, while having a comparison benchmark assures that it is possible to create a model at all. Attempts to model these cognitive states have thus far met with failure, while affective ones have been met with success [3]. There is interesting research in the improvement of the cognitive models, but this research line has been abandoned in exchange for other projects.

4. CONCLUSION

The dataset in this paper has been collected at expense to the Army, but is useful to a wider public. An initial project analyzed this dataset in order to determine if low cost sensors are able to mimic the performance of high cost sensors when supplemented with classification improvements. The finding was that they were able to, but that more work was needed in order to mimic the performance of the high cost sensors in a generalized fashion with the data available.

A secondary look at this dataset investigated a different research question. This study sought to examine whether highly individualized (not generalized) affective/cognitive models could be created with the same data available to previous classifiers. The answer to this question was that it could be done for affective models, but not be done with the raw cognitive data alone. Further work would need to be done to develop filters, feature extraction, and other, differing, methods of processing for these models.

Future efforts in this line of research will likely have to abandon the limitation on the initial streams of data through the development of feature detectors and other means of data processing. Future datasets for this line of research should look to include a checklist of features (Table 4) which would render it relevant to the learning problem area.

Table 4. Checklist of features for Low Cost Sensor dataset (recommended for other studies)

Does the dataset have…	?
Relevant states to learning	
Ability to be transferred, without modification, to another domain of instruction	
Relevant population	
Relevant cost for classroom inclusion	
Labeled data	
Initial benchmarks for research comparison	

5. REFERENCES

[1] ALZOUBI, O., CALVO, R., and STEVENS, R., 2009. Classification of EEG for Affect Recognition: An Adaptive Approach. *AI 2009: Advances in Artificial Intelligence*, 52-61.

[2] ARL, 2012. Generalized Intelligent Framework for Tutoring Release Page Army Research Laboratories, http://www.gifttutoring.org.

[3] BRAWNER, K.W., 2013. Modeling Learner Mood In Realtime Through Biosensors For Intelligent Tutoring Improvements. In *Department of Electrical Engineering and Computer Science* University of Central Florida, 500.

[4] CARROLL, M., KOKINI, C., CHAMPNEY, R., SOTTILARE, R., and GOLDBERG, B., 2011. Modeling Trainee Affective and Cognitive State Using Low Cost Sensors. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Orlando, FL.

[5] CHAMPNEY, R.K. and STANNEY, K.M., 2007. Using emotions in usability SAGE Publications, 1044-1049.

[6] HANLEY, J.A. and MCNEIL, B.J., 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology 148*, 3, 839-843.

[7] JOHNSON, R.R., POPOVIC, D.P., OLMSTEAD, R.E., STIKIC, M., LEVENDOWSKI, D.J., and BERKA, C., 2011. Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology*.

[8] KOKINI, C., CARROLL, M., RAMIREZ-PADRON, R., HALE, K., SOTTILARE, R., and GOLDBERG, B., 2012. Quantification of trainee affective and cognitive state in real-time. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* NTSA.

[9] LANDWEHR, N., HALL, M., and FRANK, E., 2005. Logistic model trees. *Machine Learning 59*, 1-2, 161-205.

[10] SOTTILARE, R.A., BRAWNER, K.W., GOLDBERG, B.S., and HOLDEN, H.K., 2012. The Generalized Intelligent Framework for Tutoring (GIFT).