

Predicting Academic Emotion based on Brainwaves Signals and Mouse Click Behavior

Judith AZCARRAGA^{a*}, John Francis IBÁÑEZ^a, Ianne Robert LIM^a, Nestor LUMANAS Jr.^a, Rhia TROGO^a & Merlin Teodosia SUAREZ^a

^aCenter for Empathic Human-Computer Interaction, De La Salle University, Philippines

*jay.azcarraga@delasalle.ph

Abstract: Academic emotions such as *confidence*, *excitement*, *frustration* and *interest* may be predicted based on brainwaves signals. It is shown that the prediction rate can be improved further when the data from brainwaves signals are complemented by data based on mouse click behavior. Twenty-five (25) undergraduate students were asked to use a math tutoring software while an EEG sensor was attached to their head to capture their brainwaves signals throughout the learning session. At the same time, mouse-click features such as the number of clicks, the duration of each click and the distance traveled by the mouse were automatically captured. Using a Multi-Layered Perceptron classifier, classification using brainwaves data alone had accuracy rates of 54 to 88%. Prediction rates based purely on mouse features had accuracy rates of only 32 to 48%. When the two input modalities are combined, accuracy rates increased to up to 92%. Furthermore, the experiments confirmed that the predication accuracy rate increases as the number of feature values that deviate significantly from the mean increases. In particular, the prediction rates exceed 80% when at least 33% of the features have values that deviate from the mean by more than 1 standard deviation.

Keywords: Affect Recognition, EEG, Mouse Behavior, Tutoring Systems

Introduction

Intelligent Tutoring Systems (ITS) interact with learners through a computer tutor that acts like a human teacher. The computer tutor analyzes the responses made by the learners and guides them through the subject matter by providing appropriate learning materials based on their cognitive state. Recent works in the design of tutoring systems have attempted to make these systems more adaptive not only to the learners' cognitive state but also to their affective state. In such systems, also referred to as affective tutoring systems, the affective states of the learner may be recognized using the tutorial information and user profile [5] and sometimes in combination with signals from hardware sensors such as a camera, special mouse, microphone [10] and various other physiological sensors that capture EEG signals, EMG signals, skin conductance levels, heart rate, and respiration rate [1][6][7][8][9][16].

Brainwaves may be captured using an electroencephalogram (EEG) device that measures the electrical activity on the scalp induced by the electro-chemical processes related to the firing of neurons in the brain. Recent works in brainwaves analysis have

attempted to measure user alertness, cognitive engagement [20] and academic emotion [3][4][15].

Another device that is not much explored but may have the potential to detect affect is the standard mouse. This inexpensive device that has the closest contact with a computer user may yield features that can provide useful information about the user's behavior. Some studies have explored the potential of a biometric mouse to measure affect [23]. Recent studies have also investigated the potential of using brainwaves signals in combination with standard-mouse data for more accurate affect detection [3][4][15].

Based on features from both brainwaves and mouse behavior data, we try to predict and classify academic emotions such as confidence, excitement, frustration and interest. Moreover, we explore under what conditions would the prediction accuracy reach acceptable levels so that future designers of affective tutoring systems may use emotion prediction systems when such conditions or situations present themselves.

1. Affective Systems

Affective tutoring systems have studied the effect of emotions in the learning process of a learner. These emotions, also referred to as academic emotions, play an important role in the success of learning [19]. In a tutoring system scenario, these academic emotions may be recognized based on the learner's interaction with the system and/or on the physiological signals captured by hardware sensors.

Some systems are able to recognize affect, to some extent, without using any hardware sensor. In such systems, affect is detected based only on the recorded student's logged activities such as scores from the previous tasks, response time in performing tasks, frequency of getting hints, etc. Since emotions are naturally complex and are expressed in different modalities (i.e. face, voice, gesture, physiological signals), most affective tutoring systems have explored the multimodal approach for affect detection because the single modality approach poses some limitations. Some studies have shown improvement in performance with the combination of contextual information and physiological signals [9].

Recent developments in the study and design of tutoring systems have added several special hardware sensors to improve the accuracy rate in predicting academic emotions. This multimodal approach for affect detection in such systems has shown some promising results. For instance, in the work of Arroyo and his group [1], affective states such as confident, frustrated, excited, and interested were predicted with high accuracy using special devices such as a camera, posture chair, pressure mouse, and skin conductance sensor. Another similar multimodal system is the *learning companion* [16] that fuses information from camera, posture chair, pressure-sensitive mouse, skin conductance sensor and task state to help predict frustration and to determine if the user needs help. Likewise, *Autotutor* uses information from conversation cues, posture and facial features to be able to predict student boredom, flow/engagement, confusion and frustration [10].

Some researches have explored the potential of using electroencephalogram (EEG) devices for affect detection. In one study, the student's level of frustration, distraction and cognitive workload were observed while the student is engaged in different activities in a multimedia-learning environment [22]. Other research works have investigated the use of brainwaves to detect the affect of students while using a math software [2][3][4][15]. A similar work has explored the accuracy of using brainwaves signals and emotional dimensions in predicting the correctness of the student's answers [13]. Moreover, the use

of a biometric mouse to measure a user's emotional state and productivity was described in [23]. The study attempts to use the mouse for capturing motor behavioral information from skin conductance, amplitude of hand tremble, and skin temperature.

Most multimodal systems have focused on using expensive and sophisticated sensors for affect detection. To date, not much work has explored the use of a standard mouse which may have the potential to measure affect such as frustration [21]. Certain behavioral responses may be measured through mouse events such as mouse-clicks. Some patterns were observed in their mouse behavior when subjects were presented with frustration-eliciting events. Indeed, some studies have suggested that emotions and mood may have an effect on a person's motor movements [17]. Thus, it is possible that mouse events such as mouse clicks, frequency of mouse movement and duration of mouse clicks correlate with the grade on the valence and arousal dimensions of emotions. A user tends to click more when they are frustrated with the system (such as when there are lags and delays) [11].

Despite the positive results that were reported by such affective tutoring systems, much remain to be explored. In particular, the potential of combining physiological signals and mouse click data in order to improve the accuracy of predicting the affective state of the user has yet to be studied more extensively [3].

2. Experimental Set-up

Twenty-five computer science undergraduate students (14 male and 11 female) with ages from 17-21, all mentally healthy and right-handed, were recruited as the subjects in this experiment. All the subjects have already taken an intermediate algebra course. The participants were asked to learn a tutoring software called *Aplusix* which teaches algebra [18]. They were asked to solve 4 algebra equations of varying difficulty levels for about 15 minutes. While they were learning using the software, signals from an EEG sensor attached to their head were recorded. Also, the details of their mouse clicks, click duration and movement were automatically captured and stored in 2 different mouse log files - one for the clicks and duration and another for the movement.

The EEG device that was used in the experiment is the *Emotiv EPOC* sensor. Typically used for gaming purposes, the *Emotiv EPOC* sensor is equipped with 14 channels based on the International standard 10-20 locations. A service program was created to automatically capture the raw EEG signals coming from each of the channels.

Prior to the actual tutorial session, each subject was asked to close his/her eyes and relax for a period of 3 minutes in order to create the baseline EEG data while an EEG sensor was attached to the head. Brief instructions were then given on how to use the software. An observation module was developed to capture raw EEG signals and mouse data during the tutorial session. An emotion annotation window automatically pops up every 2 minutes. The level or intensity of each of the 4 emotions, *confidence*, *excitement*, *frustration* and *interest* can be specifically described by the participant using a sliding bar with values from 1 to 100 for each of the four emotions.

3. Data Preprocessing and Data Preparation

From the 25 subjects, only 16 were found to be useful, given the stringent conditions we had set in terms of balancing the data for all the four different emotions. Some were also not included due to lack of reported emotions.

Two EEG recordings were performed on each subject: one from the relax period

and one from the tutorial session. During the relax period, the values of each EEG channel for each subject were averaged. The average value serves as the baseline EEG data of that particular subject. The raw EEG channel values taken during the tutorial session were processed by computing the difference between the raw value of the channels and the mean value of corresponding channels from the baseline (relax state) data.

All the pre-processed EEG data, mouse data, and self-reported emotion tag were carefully synchronized, merged and uniformly segmented into 2-second windows with 1-second overlap. Each segment was treated as a single instance in each subject's dataset. The full dataset had a total of 17 features: 14 for the EEG channels and 3 for mouse behavior as summarized in Table 1. The self-reported emotion serves as the tag for each recorded instance.

Table 1. Features for Emotion Classification

EEG channels : AF3 F7 F3 FC5 T7 P7 O1 O2 P8 T8 FC6 F4 F8 AF4
Mouse Behavior : Number of Clicks, Distance Travelled, Click Duration
Self-reported Emotion : Frustrated, Interested, Confident, Excited

Six different datasets were formed based on the percentage of feature *outliers*. A feature value is considered an outlier if it exceeds 1 standard deviation from the mean of that particular feature and for that particular subject. Feature values that are outliers for each instance were counted. Based on this number, different datasets were formed as described in Table 2. The full dataset (Dataset 0) includes the instances from all the 16 subjects. Dataset 10 is composed of only those instances where at least 10% of the feature values are outlier values. Dataset 25 is composed of only those instances where at least 25% of the feature values are outlier values, and so on.

Each dataset was balanced by ensuring that there are the same number of instances for each emotion. This is a critical step in dataset preparation as this would prevent any bias that would severely affect the multi-layered perceptron classifier. For Dataset 60, only 15 subjects were included since 1 subject did not have instances that had at least 60% of the features being outliers.

Table 2. Datasets for Emotion Classification

Dataset	No. of Outlier Features	No. of Students	Instances/Emotion
0	0 or more	16	3600
10	2 or more	16	2250
25	4 or more	16	650
33	6 or more	16	325
50	8 or more	16	260
60	10 or more	15	165

4. Results and Discussions

For each dataset, the accuracy of classifying the emotions of each modality, whether brainwaves or mouse, as well as of their combination was analyzed using the Multi-Layered Perceptron (MLP) classifier of WEKA, a machine learning tool for feature classification [12]. To test and validate the data, a 10-fold cross validation technique was employed.

Based on the results as shown in Table 3, we can compare the performance of each

modality. Classification based on brainwaves sensor data were consistently and significantly better than when based on just the data based on mouse behavior. The former had accuracy rates of 54% to 88% while the latter had accuracy rates of only 32% to 48%. The results of Table 3 also clearly show that the classification accuracy improves when data from both the EEG sensors and the mouse clicks are used. The classification accuracy goes up to a minimum of 61% and up to 92%.

The dataset preparation, based on outliers, was designed to confirm our hypothesis that when limited to instances where some feature values deviate significantly from their mean values for a given subject, the prediction accuracy increases. Concretely, feature values that deviate significantly from the mean are recordings of the EEG sensor when it is picking up something unusual, or when the mouse is handled or clicked somewhat differently. Table 3 clearly confirms our hypothesis. When at least 33% of the features are outliers, classification accuracy exceeds 80%, and accuracy rates even get to exceed 90%, when at least 60% of the features yield unusual (outlier) values. Tables 4 give the details of the precision, recall and f-measures according to specific emotion category while Table 5 presents the confusion matrices for the 6 datasets. Tables 4-5 reveal that the overall results of Table 3 are spread quite uniformly across all four emotions, except that the emotion *interest* can be predicted at a slightly higher rate compared to the other 3 emotions.

Table 3. Accuracy of emotion classification (percentage of correctly classified) in different modalities using Multi-Layered Perceptrons

Dataset	Brainwaves	Mouse	Brainwaves + Mouse
0	54.66	32.26	61.04
10	63.74	38.9	69.8
25	75.27	45.11	78.58
33	74.92	45.46	80.69
50	83.65	43.85	88.56
60	88.33	48.79	92.27

5. Conclusion and Future Study

Twenty-five (25) undergraduate students were asked to use a math tutoring software while an EEG sensor was attached to their heads to capture their brainwaves signals throughout the learning session. At the same time, mouse-click features such as the number of clicks, the duration of each click and the distance traveled by the mouse were automatically captured. The study reported here confirms that indeed, academic emotions such as *confidence*, *excitement*, *frustration* and *interest* may be predicted based on brainwaves signals. It is also shown that the prediction rate can be improved further when the data from brainwaves signals are complemented by data based on mouse click behavior. Using a Multi-Layered Perceptron classifier, classification using brainwaves data alone had accuracy rates of 54% to 88%. Prediction rates based purely on mouse features had accuracy rates of only 32% to 48%. When the two input modalities are combined, accuracy rates increased to up to 92%.

Furthermore, the experiments confirmed that the prediction accuracy rate increases as the number of feature values that deviate significantly from the mean increases. In particular, the prediction rates exceed 80% when at least 33% of the features have values that deviate from the mean by more than 1 standard deviation using MLP. Future tests need to be done to investigate performance rates of other classifiers. Moreover,

brainwaves signals may be processed using more advanced techniques such as Fast-Fourier transform (FFT) to extract more features that may be tested using the *outlier* detection approach presented in this paper.

Table 4. Precision, Recall, and F-measure values for each emotion

ALL INSTANCES / DATASET 0									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.485	0.311	0.379	0.311	0.147	0.2	0.559	0.4	0.466
Excitement	0.569	0.562	0.566	0.334	0.315	0.324	0.604	0.628	0.616
Frustration	0.511	0.603	0.553	0.295	0.19	0.231	0.556	0.657	0.602
Interest	0.597	0.711	0.649	0.329	0.638	0.434	0.712	0.757	0.734

DATASET 10									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.554	0.564	0.559	0.328	0.068	0.113	0.616	0.588	0.602
Excitement	0.596	0.648	0.621	0.343	0.326	0.334	0.686	0.678	0.682
Frustration	0.68	0.567	0.619	0.408	0.419	0.413	0.678	0.687	0.682
Interest	0.726	0.771	0.748	0.41	0.743	0.528	0.804	0.839	0.821

DATASET 25									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.714	0.657	0.684	0.295	0.228	0.257	0.708	0.715	0.712
Excitement	0.74	0.717	0.728	0.514	0.488	0.5	0.788	0.788	0.788
Frustration	0.72	0.751	0.735	0.434	0.42	0.427	0.763	0.748	0.755
Interest	0.83	0.886	0.857	0.511	0.669	0.579	0.884	0.892	0.888

DATASET 33									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.642	0.689	0.665	0.427	0.28	0.338	0.795	0.8	0.798
Excitement	0.741	0.748	0.744	0.336	0.471	0.392	0.759	0.865	0.809
Frustration	0.744	0.705	0.724	0.384	0.351	0.367	0.78	0.729	0.754
Interest	0.883	0.855	0.869	0.696	0.717	0.706	0.906	0.834	0.869

DATASET 50									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.81	0.785	0.797	0.472	0.258	0.333	0.868	0.888	0.878
Excitement	0.777	0.846	0.81	0.352	0.288	0.317	0.859	0.912	0.884
Frustration	0.844	0.769	0.805	0.474	0.427	0.449	0.858	0.838	0.848
Interest	0.918	0.946	0.932	0.45	0.781	0.571	0.963	0.904	0.933

DATASET 60									
	BRAINWAVES ONLY			MOUSE ONLY			BRAINWAVES AND MOUSE		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Confidence	0.839	0.818	0.828	0.545	0.364	0.436	0.933	0.842	0.885
Excitement	0.89	0.885	0.888	0.547	0.594	0.57	0.943	0.909	0.926
Frustration	0.846	0.83	0.838	0.289	0.248	0.267	0.868	0.958	0.911
Interest	0.954	1	0.976	0.537	0.745	0.624	0.953	0.982	0.967

From the results, it can be claimed that designers of future learning software can employ emotion-prediction systems whenever 33% or more of the features have significantly different or unusual values. When there are fewer features that deviate significantly from the mean, the results of the emotion prediction system may not be dependable.

Table 5. Confusion matrices

Dataset 0	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	1119	605	1128	748	529	823	563	1685	1439	650	1080	431
Excitement	393	2024	614	569	430	1134	593	1443	398	2260	535	407
Frustration	452	566	2169	413	450	905	685	1560	442	528	2366	264
Interest	342	363	336	2559	293	529	481	2297	296	301	278	2725

Dataset 10	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	1268	454	279	249	154	620	698	778	1323	339	367	221
Excitement	372	1459	199	220	142	733	530	845	359	1526	257	108
Frustration	455	334	1276	185	101	423	943	783	337	237	1545	131
Interest	195	199	122	1734	72	364	143	1671	129	124	109	1888

Dataset 25	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	427	95	97	31	148	151	157	194	465	71	87	27
Excitement	87	466	61	36	112	317	130	91	74	512	44	20
Frustration	59	52	488	51	137	108	273	132	88	47	486	29
Interest	25	17	32	576	105	41	69	435	30	20	20	580

Dataset 33	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	224	44	43	14	91	115	77	42	260	33	24	8
Excitement	48	243	22	12	72	153	73	27	18	281	19	7
Frustration	57	28	229	11	43	135	114	33	37	38	237	13
Interest	20	13	14	278	7	52	33	233	12	18	24	271

Dataset 50	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	204	27	19	10	67	57	46	90	231	12	13	4
Excitement	26	220	10	4	35	75	68	82	16	237	7	0
Frustration	16	36	200	8	28	45	111	76	16	21	218	5
Interest	6	0	8	246	12	36	9	203	3	6	16	235

Dataset 60	BRAINWAVES ONLY				MOUSE ONLY				BRAINWAVES + MOUSE			
Classified as	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest	Confidence	Excitement	Frustration	Interest
Confidence	135	12	15	3	60	27	32	46	139	4	16	6
Excitement	8	146	10	1	18	98	36	13	6	150	8	1
Frustration	18	6	137	4	32	45	41	47	4	2	158	1
Interest	0	0	0	165	0	9	33	123	0	3	0	162

Acknowledgements

We thank all the subjects who participated in this experiment. Also, we thank the Center for Empathic Human-Computer Interaction and the Department of Science and Technology of the Philippine government for all the support in conducting this research.

References

- [1] Arroyo, I., Cooper, D., Burleson, W., Woolf, B., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. *Frontiers in artificial intelligence and applications*, 200, pp. 17–24.
- [2] Azcarraga, J., Inventado, P.S. & Suarez, M. (2010). Predicting the Difficulty Level Faced by Academic Achievers based on Brainwaves Analysis. In S.L. Wong et al. (Eds.) *Proceedings of the 18th*