Predicting students' performance using emotion detection from face-recording video when interacting with an ITS

Wilson Chango[1], Miguel Sanchez-Santillan[2], Rebeca Cerezo[2], Cristóbal Romero[3]
[1]Pontifical Catholic University of Ecuador, Ecuador
[2]University of Oviedo, España
[3]University of Cordoba, España
wilson.chango@pucese.edu.ec, sanchezsmiguel@uniovi.es, cerezorebeca@uniovi.es, cromero@uco.es

ABSTRACT
This research aims to predict the academic performance of students when interacting with an Intelligent Tutoring System (ITS) from emotions detection and analysis. We use data from 47 university students in a virtual learning environment. We have used data gathered form face recording of students' interactions with the system to detect students' emotions and determine to what extent they can predict the final students' performance during the learning session.

Keywords
Predicting performance, Emotion detection, Video analytics.

1. INTRODUCTION
Emotions are a critical component of learning and problem solving, especially when it comes to interacting with computer-based learning environments (CBLEs) [5]. Studies from affective computing literature suggest that facial expressions may be the best single method for accurately identifying emotional states [4]. The automatic detection of emotions technics are capable of isolating the mood of a learner by means of a facial recognition system through artificial intelligence and there are already tools that enable the processing of data in the form of video, such as the Microsoft Emotion API [1], FaceReader™, etc. However, we have not noticed previous studies testing to what extent the emotion recognition result of these tools is powerful enough to predict student´s performance. It could be potentially contributing to enhance the quality and efficacy of CBLEs (e-learning, multi-agent systems, intelligent tutoring systems, serious games, etc.) by including the learner's emotional states.
This research aims to test if student's emotions recognized by and API during a learning session with an ITS can be enough to predict the final student's performance.

2. EXPERIMENTS
Data were collected from 47 undergraduates enrolled at a public university in the north of Spain whom learned about a complex science topic while interacting with the ITS MetaTutorES [3] a multi-agent computerized learning environment. Participants represented a variety of disciplines, including psychology, education and engineering. The emotion data collected was naturally occurring, the emotions arouse from interactions with a the ITS MetaTutorES, designed to teach learners about the human circulatory system during a session ranging from 2:30 to 3:00 hours. During and at the end of the session, performance test about the circulatory system knowledge were taken for every subject, giving a final performance value ranging between 0 and 10, showing 10 the best performance. A pretest about previous circulatory system knowledge is taken at the beginning of the session and final performance is corrected based on that previous level. Videos from every learner' facial expressions were captured with a webcam and analyzed using automatic facial recognition software (Microsoft Emotion API [1]). The API classifies the facial expression in eight classes of emotion: anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. The analysis allows us to obtain at least one highest emotion during the learning session from every students´ frame having a high volume of frames (1 frame for second) for each student in every session. The confidence (values between 0 and 1) gives the likelihood for each class of emotion.

The first step of the experiment consists on check the correlations between the emotions detected and the students´ performance. The Pearson correlation test examines the relationship of each emotion with the student's performance obtanined during the learning session. The R value in Pearson's correlation coefficient goes from -1 to 1, meaning both values a high level of correlation and 0 a null level of correlation between variables (See Table 1).

Table 1: Pearson correlation test results

Emotion	R-Value
Anger	0.1295
Contempt	0.2165
Disgust	0.0882
Fear	-0.2415
Happiness	0.0459
Neutral	0.0463
Sadness	0.1546
Surprise	-0.1062

According to the results of table 1 none of the variables is highly correlated with the performance. However, based on the axes of emotions valence -positive emotions (happiness); negative emotions (anger, contempt, fear, disgust); non valence (neutral and surprise) [6] and looking at the positive or negative relationship, we can observe that only negative or non valence emotions are negatively related with performance.

Wilson Gustavo Chango Sailema, Miguel Sánchez, Rebeca Cerezo and Cristóbal Romero "Predicting students' performance using emotion detection from face-recording video when interacting with an ITS" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 578 - 580

In the second step, we applied several classification algorithms using the 8 emotions as input attributes for predicting the student's final performance. We used white box classification models (decision trees and rule induction algorithms) because the models they produce (tree and IF-THEN rules) are easy to understand [7]. In our experiments, we selected six well-known classification algorithms provided by WEKA [8]: three decision tree algorithms and three rule induction algorithms (see Table 3).

Table 3: Decision Trees classification algorithms.

Type	Algorithms	Description
Trees	**J48**	Java implementation of C4.5.
	Reptree	Fast tree learner that uses reduced-error pruning.
	Randomtree	Construct a tree that considers a given number of random features at each node.
Rules	**Jrip**	RIPPER algorithm for fast, effective rule induction.
	Nnge	Nearest-neighbor method of generating rules using generalized exemplars.
	Part	Obtain rules from partial decision trees built using J4.8

We executed each algorithm using stratified 10-fold cross-validation in which the dataset is randomly divided into 10 disjointed subsets of equal size in a stratified manner. We have compared the test results using the Accuracy and ROC Area evaluation measures (see Table 4).

Table 4. Results produced by all algorithms.

Algorithm	% Accuracy	ROC Area
Jrip	63,8298	0,5820
Nnge	53,1915	0,5290
Part	63,8298	0,6590
J48	**63,8298**	**0,6770**
Reptree	48,9362	0,5170
Randomtree	59,5745	0,5950
Avg	58,8653	**!Error de sintaxis,**)0,5932

Table 4 shows that the best results (highest values) were produced by J48 (63,8298%Acc and 0.6770 AUC). Next, we show in Table 5 the obtained decision model by J48 algorithm.

Table 5. J48 decision tree.

```
    Contempt <= 0.126904: Pass
    Contempt > 0.126904
    |  Disgust > 0.137741
    |  |  Sadness <= 0.1977232
    |  |  |  Fear <= 0.1551857: Pass
    |  |  |  Fear > 0.1551857: Fail
    |  |  Sadness > 0.1977232: Fail
    |  Disgust <= 0.137741: Pass
    Number of Leaves  :   5
Size of the tree :   9
```

The Table 5 show us a decision tree that let us learn some interesting information from. On one side, students who Pass show lower values than an umbral of emotions contempt, disgust, fear and disgust, and students who Fail show higher values than a umbral of these emotions.. On the other side, we can observe that negative emotions have more prediction power on performance than positive or non-valence emotions. And finally, negative emotions values over 0.15 (15% of the session time) are definitory to a Fail ending.

3. CONCLUSIONS

There was an assumption that emotions experienced during complex learning will impact learning and problem solving [2], and therefore, achievement. However, in this study, we observe that student's emotions when interacting with an ITS are not enough for predicting students' final performance. The results give us some information the relationship of each emotion with the student's performance. However could be necessary to refine and redefine the API emotions classification based on an educational psychology framework for some close emotions (e.g attention, engagement, hope, pride, etc.).

Finally, we purpose as a future prospect adding other different variables/attributes from the interaction with the ITS such as log files, eye tracking, etc. in order to obtain higher accuracy values to predict students´ performance. We also want to use more classifiers algorithms, particularly deep learning which would perform significantly better than classic methods.

4. ACKNOWLEDGMENTS

The authors acknowledge the financial subsidy provided by Spanish Ministry of Science and Technology TIN2017-83445-P and Plan de Ciencia Tecnología e Innovación del Gobierno del Principado de Asturias (FC-GRUPIN-IDI/2018/000199).

5. REFERENCES

[1] Arora, R. et al. 2018. Microsoft Cognitive Services. *International Journal of Engineering Science and Computing*. 8, No.4 (2018), 17323–17326..

[2] Buder, J. and Hesse, F.W. 2017. *Informational environments: Effects of use, effective designs*. Springer.

[3] Cerezo, R. 2018. MetaTutorES: Evaluación e intervención en metacognición desde una perspectiva multimodal. Symposium conducted at the *IX International Congress of Psychology and Education* (Logroño, Spain, Jun. 2018).

[4] D'Mello, S. and Kory, J. 2012. Consistent but Modest: A Meta-Analysis on Unimodal and Multimodal Affect Detection Accuracies from 30 Studies. *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (New York, NY, USA, 2012), 31–38.

[5] Harley, J.M. et al. 2015. A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*. 48, (Jul. 2015), 615–625.

[6] Pekrun, R. 2011. Emotions as drivers of learning and cognitive development. *New perspectives on affect and learning technologies*. Springer. 23–39.

[7] Romero, C. et al. 2013. Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*. 21, 1 (2013), 135–146. DOI:https://doi.org/10.1002/cae.20456.

[8] Witten, I.H. et al. 2011. *Data Mining:Practical Machine Learning Tools and Techniques second edition*.