

Automatic Identification of Learning-Centered Emotions: Preliminary Study for Data Collection

Yesenia N. González-Meneses, Josefina Guerrero-García

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, Puebla, Mexico

{yeseniaglez0, joseguga01}@gmail.com

Abstract. The intention of this work is to achieve an automatic identification of emotions in educational environments using machine learning algorithms and physiological and behavioral signal acquisition technologies to identify relations between emotions and learning. Four of the main learning-centered emotions are considered [1]: engagement, boredom, confusion and frustration. It is proposed to make a fusion of data from four signal acquisition technologies with the objective of achieving the identification of emotions in the most precise manner. The development of an appropriate database for the study of emotions is a fundamental task. Therefore, considering the stages of the proposed methodology, the first of them is presented and the design of the experiment that will be executed for data collection with college students during a learning process.

Keywords. Automatic identification of emotions, learning-centered emotions, machine learning.

1 Introduction

From a computational approach, the detection and identification of emotions is a relevant problem within the area of affective computation research, since it is the starting point in the study and development of human-machine interaction systems sensitive to the emotions of human beings. This problem in the human-machine interaction area is known as classification of emotions, automatic detection or identification of emotions, as it is mentioned in [2] and [3].

The complexity of the problems that affective computation deals with lies in the fact it is an interdisciplinary area that encompasses computer science, psychology, and cognitive science [4].

So, to develop models of emotion recognition that generate satisfactory results with an acceptable degree of precision, it will be necessary to study these three areas in depth.

Particularly, in the computational area the challenge is selecting and testing machine learning algorithms that could be integrated in a complete model of recognition of learning-centered emotions. In this sense, two approaches have been used to recognize

human emotions [5]: The objective, using sensors or capturing images; as in the recognition of facial expressions, voice, heart rate, body language, body's thermal activity, muscular activity and brain waves. And the subjective, such as contextual analysis through direct observation, surveys, or interrogations of the individuals themselves. The automatic recognition of emotions and the human-machine interaction can be considered, then, from the use of physiological signal sensors that allow the acquisition of data through voice analysis, video images of the face, eyes, head, or body movements of the people [5]. Within the human-machine interfaces that capture physiological signals that can help in the recognition of emotions, as mentioned in [6], the brain wave diadems that send information to an electronic device in the form of an electroencephalogram (EEG) are found.

There are also cardiovascular wristbands, that measure heart rate and provide information in the form of an electrocardiogram (ECG); electrodermal activity sensors, that measure the level of conductivity of the skin through sweat on the hands; thermal cameras that allow the measurement of the temperature change of the human body associated with the different emotional states. There are also devices that measure muscle electrical activity in response to a nerve stimulation of a muscle, in the form of an electromyography (EMG). Regarding devices related to the identification of people's behavior -such as body postures and gestures- there are traditional video cameras, webcams or augmented reality that allow the recording of facial expressions and body movements, as well as eye tracking, important to the recognition of emotions. In this type of devices, there are also voice recorders, another medium used to identify emotions, as mentioned in [7].

The data obtained from different devices must be processed and classified based on a specific objective. In this case, they will be used for the recognition of learning-centered emotions and captured in real educational environments when students are executing a learning activity. These activities can be done through an electronic device and using an intelligent tutorial system or MOOC (Massive Open Online Courses), through an educational videogame, using a computer in a traditional way, whether watching a video, studying, investigating or reading.

Therefore, it is expected to contribute to physiological signal processing technics, from their capture, preprocessing, selection of characteristics and, finally, with the identification of emotions centered around learning. For this, we will use machine learning algorithms that will be implemented to process the different data coming from at least four signal acquisition technologies and captured in educational environments, to improve recognition accuracy and obtain a model of emotion recognition with an acceptable level of trust closer to reality within and educational environment.

The objective of identifying emotions during educational activities is to corroborate the relationship between learning-centered emotions and the level of learning obtained by students. This relationship is the base for approaching educational strategies that help to improve the levels of learning and, therefore, the educational level in our country.

2 Problem Description

The problem involving automatic recognition of emotions has been an area of investigation highly active in the last years. Regardless of this, a clear solution, that is within reach of the most people, is still far away. Several drawbacks have influenced the construction of an appropriate solution from a computational point of view. On one hand, a factor that affects the performance of emotional recognizers in real contexts is the difficulty to generate databases with spontaneous emotions. Generally, works are made with actuated databases which provide portraits of emotions representing prototypical and intense emotions that facilitate the search of correlations and the subsequent automatic classification. This kind of databases are usually captured in a controlled environment, which decreases problems in the processing of information (noise, for example). In addition, it can guarantee a balanced number of samples per class. As consequence, there have not been good results when translating the knowledge extracted from these databases to real contexts [8]. In contrast, the data bases with spontaneous records show information with emotional content that does not belong to a single class, but a mixture of them. In other cases, there are samples with a very light emotional charge, close to a neutral emotional state. In addition, databases with spontaneous emotions are usually recorded in noisy environments, such as classrooms, study rooms, entertainment areas, offices, factories or in phone conversations, which leads to the inclusion of noise. Finally, because of the very nature of the problem, it is not possible to ensure a balanced quantity of examples per class.

Another challenge to be solved is the extraction and selection of a set of characteristics that allow recognizing emotions in the data captured spontaneously. Although progress in the area has been important, there is still much to be done in realistic contexts. Therefore, it is necessary to propose and explore other approaches that allow to reach a good performance of the recognition of emotions in real world applications. An evident aspect to consider is the fact that the area of application has an important influence on the accuracy of emotion recognition [7], since this may vary according to the context the age of the participating individuals, the hours of the day in which they work, the comfort with which activities are done, the degree of intrusion of the tools used, the specific type of activity that is being carried out, among others.

Considering the presented problems and the analysis of recent related works, we can identify that educational environments there is a lack of adequate methodologies to recognize the emotional state of students during learning processes through interaction with a computer. Preferably, that integrates diversified characteristics, obtained even from the fusion of data coming from the use of different technologies of acquisition of physiological and behavioral signals based on a model of emotions that allows to become closer to reality and, overall, to the cognitive learning process, which contributes to the analysis of the emotion-learning relation.

2.1 Research Objectives

The objective of this work is to recognize emotions in educational environments using machine learning algorithms and technologies of acquisition of physiological and

behavioral signals to identify emotion-learning relationships. To achieve the above, the following activities are proposed:

- Investigate, analyze and select the model of emotions and technologies for the acquisition of physiological data on which the proposal will be based.
- Analyze and select educational environments to make the physiological and behavioral data capture.
- Investigate, study and select machine learning algorithms for the selection of characteristics for the classification of emotions.
- Design a methodology for the identification of learning-centered emotions.

To validate the work done, tests will be carried out with college students and the recognition accuracy will be evaluated with metrics that allow comparing it with related works in the literature. With this we pretend to corroborate how the proposed methodology allows identifying emotions in educational contexts that can contribute to decision making in emotion-learning relationships.

3 Related Works

The analysis of the state of the art that is being carried out consists of a collection of articles focused specifically on the recognition of emotions in learning activities and the identification of the emotion-learning relationship. These investigations in automatic recognition of emotions represent approximately 20%, and the remaining proportion (approximately 80%) are works that recognize basic emotions (happiness, sadness, fear, anger, contempt, disgust and surprise [9]). There is a clear difference in the literature found. There are works that, although they make an identification of emotions, give greater relevance to the analysis of the emotion-learning relationship. Others emphasize especially the algorithms for automatic recognition of emotions and the recognition rate reached and, with less importance, analyze the emotion-learning relationship. Considering these two aspects, a review of the state of the art analyzed so far is presented.

3.1 Works Focused on the Computational Problem of Automatic Recognition of Learning-Centered Emotions

In the literature reviewed, these works represent approximately 63% of the total of papers reviewed that identify learning-centered emotions. Only the most recent research is mentioned here.

In the work of [10] they use a convolutional neuronal network for the recognition of learning-centered emotions. They run tests using three data-bases: RaFD, database of posed facial expressions containing images of 8 basic emotions and two spontaneous databases created by themselves, especially with content related to learning-centered emotions. The emotions they recognize are: Engagement, excitement, boredom and relaxation. The accuracy they achieve when using their databases is 88% and 74%, respectively.

In [11] they propose a computer-assisted method for special school instructors, where they teach students with mental disorders or emotional problems using a system that employs wearable sensor technologies and intelligent recognition of emotions. The emotion recognition module starts with the capture of the signal from brain diadems. The filtered data is sent to the characteristic extraction module. After extracting the characteristics, they are processed with two classifiers: support vector machine and near k-neighbors with a cross-validation of 10 iterations. The emotions they recognize are: happiness, calm, sadness and fear. An expression module shows the instructor suggestions for treatment according to emotional state detected.

In [12] they implement a binary local pattern for the recognition of learning-centered emotions. The purpose of the work is to build a database of spontaneous facial expressions corresponding to affective states in education to be used in different intelligent tutorial systems. The data capture technologies they use are video and EEG diadems (Emotiv-EPOC). The learning-centered emotions that they recognize are: frustration, boredom, engagement and excitement. They use a support vectors machine to do the classification of emotions. After applying the recognizer, its calculated precision was 80%.

The work of [2] explains the construction and validation of a database of facial expressions that they collect by taking pictures with a webcam. Each photograph is labeled with the emotions of the users obtained at that moment from the Emotiv-EPOC device. For the recognition of facial expressions use a technique based on geometry that measures the distances between the central point of the face and 68 other reference points. These measurements are transformed into characteristics to train a support vector machine. They obtain an accuracy, per emotion, of: boredom of 64%, engagement of 64%, excitement of 83% and frustration of 62%. The same model for the recognition of emotions is also used in [13], as part of an affective learning environment based on Web 3.0 to learn how to program in Java.

Finally, in [14] they propose different recommended activities to induce a certain mental state and capture the EEG response for each one of them. They intend to identify the ideal emotional state to learn. These activities are based on psychological research dedicated to measuring the level of attention, concentration and other functions. In the classification process, they use k-means and clustering with a total concentration rate of 96% of 3592 instances.

Of the totality of works reviewed we can conclude that the most popular devices for the capture of data are webcams, followed by EEG Emotiv diadems. From the use of these technologies and other less popular ones, 80% of the works create their own databases to train and test their recognizers. The most used algorithms for classification tasks are artificial neural networks, support vector machine (SVM), clustering techniques, Bayesian classifiers and nearer neighbors (KNN), among others.

Considering this analysis, we will begin the processing of the data collected using the aforementioned algorithms with the purpose of observing their behavior on the signals that we capture. Later we will make a proposal of our own that may include the use of hybrid algorithms.

3.2 Works Focused on the Analysis of the Emotion-Learning Relationship Based on the Automatic Recognition of Learning-Centered Emotions

Up next, the most recent research on the analysis of the emotion-learning relationship using computer techniques for the identification of emotions is listed. These works represent 37% of the literature reviewed so far, in terms of recognition of learning-centered emotions.

In [15], they make a study of the affective states that originate when students learn with technology, using a tutorial to learn the basics of programming in Python. The emotions they can identify are: engagement, confusion, frustration, boredom, curiosity (which were the most frequent affective states), anxiety, happiness, anguish, surprise, disgust, sadness and fear (which were the rarest). With the analysis of their results, they identify the emotion-learning relationship from five different approaches. They conclude that there is no co-occurrence of affective states that can be generalized; they corroborate the affective dynamics model of [1].

In [16] they present the development of an intelligent tutor with recognition and management of emotions for mathematics. And in [17] they develop an affective learning system for algorithmic logic by applying gamification. In both works, a module for the recognition of emotions is integrated. This process is based on the analysis of faces. In the second work they also add the capture of EEG data signals to label the facial images.

The analysis of these works helped us to identify the areas of opportunity regarding our research and to be able to define the design strategies of the methodology and algorithms for the selection of characteristics and classification to be used. This to validate the hypothesis that states that the proposed methodology allows us to identify emotions in educational contexts that contribute to the decision-making in the emotion-learning relationship.

4 Proposed Methodology

General methodology of the investigation is shown graphically in Fig.1. The main stages are:

- Research, analysis and selection of technologies for acquisition of physiological data, emotions model, application context and machine learning algorithms, which will be proposed for the development of the methodology.
- Data processing; it includes the recording and acquisition of physiological and behavioral data to form the data base, the preprocessing required to prepare the data and the implementation of extraction, selection and integration algorithms of relevant characteristics derived from the different signals.
- Identification of emotions; the selected machine learning algorithms will be tested and trained for classifying emotions centered on learning.
- Validation; tests will be carried out to evaluate the complete methodology for the identification of emotions with metrics that measure their execution, precision and accuracy. Subsequently, the results will be interpreted to identify

and analyze emotion-learning relationships. Finally, a discussion of the objectives reached will be made.

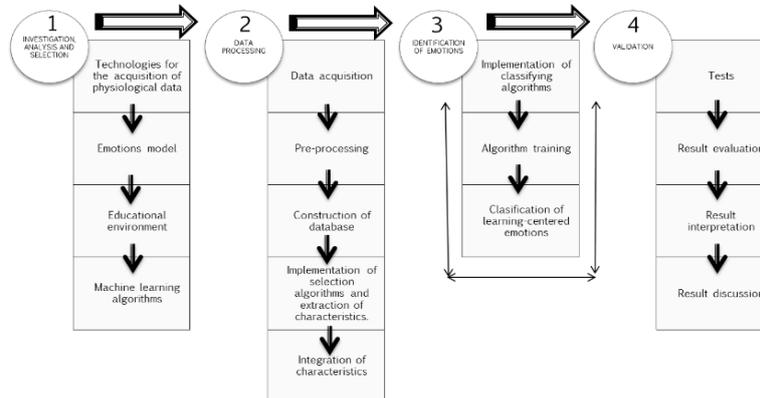


Fig. 1. General Methodology.

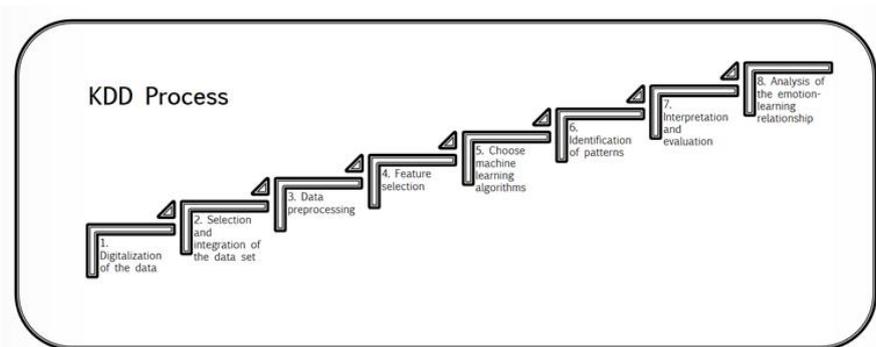


Fig. 2. Knowledge discovery in databases process for automatic identification learning-centered emotions.

Figure 2 show each of the stages of the KDD process (Knowledge Discovery in Databases) [18] which implicitly form part of the computational process for the automatic recognition of learning-centered emotions (these are part of the stages 2 and 3 in figure 1) that propose to implement:

- Recollection of data: correspond to the digitization of data and its storage.
- Data labeling: it consists in selecting and integrating all the data coming from multiple and heterogeneous sources.
- Preprocessing: if necessary, elimination of noise and isolated data. Use of prior knowledge to eliminate inconsistencies and duplicates. Choice and use of strategies to manage the missing information from the data sets.
- Transformation of the data: here the selection of useful characteristics to represent the data, the reduction of dimensionality or methods of transformation and fusion of the data is carried out.

- Data mining: is to choose machine learning algorithms. Choose the task of data mining and mining algorithms that treat all the criteria (classification, regression, clustering or mixed models).
- Identification of patterns: look for patterns of interest in a special way, that is, the recognition emotions.
- Interpretation and evaluation of mine patterns.
- Consolidation of discovered knowledge: discussion and analysis of the relationships of emotions identified with the level of learning.

5 Preliminary Study for Data Collection

The data collection is part of the first stage of the proposed methodology. This includes research, analysis and selection of data acquisition technologies, emotion model, education environment and computer learning algorithms. As part of the preliminary study, the first three elements have been selected. The study, analysis and selection of physio-logical and behavioral data acquisition technologies was carried out by choosing the following: traditional video camera, ICI 9320P thermal camera, heart rate sensor implemented with Arduino and Kinect for Windows. Regarding the model of emotions, we chose to work with discrete and continuous emotions. The learning-centered emotions to recognize will be: engagement, boredom, confusion and frustration. The education environment with which students will interact for the teaching-learning process will be a MOOC of basic algebra using the first topic (duration of 36 minutes).

The experiment for data collection will consist in the capture of behavioral signals and physiological signals in a laboratory of experiments of the Doctorate LKE (Language and Knowledge Engineering) at BUAP (Benemérita Universidad Autónoma de Puebla). We will work with students of the computer faculty at same University, the sample size will be 100 students, who will be asked to participate in a learning activity using the MOOC of basic algebra. The recorded data will be stored, processed and used to recognize the emotions that students present during the learning process. Behavioral signals will correspond to images of facial expressions and movement of the head and hands, stored in JPG and AVI files.

The physiological signals correspond to measurements of the heart rate and temperature of facial areas, stored in TXT, JPG, AVI and XLSX files. The capture of behavioral signals will be done through a video camera and the cameras of the Kinect 360. The capture of physiological signals will be done through a heart rate sensor which will be placed on the ring finger of the left hand. A thermal camera will also be used to capture the facial temperature. Every 10 minutes two tests will be applied to obtain a self-evaluation by the student. In the first test the student must select the emotion he is feeling (engagement, boredom, frustration or confusion). To measure the emotion in a continuous way, the manikin test will be applied, in which the level of each of the three variables involved in the presented emotion will be selected (valence, activation and dominance). In figure 3, photographs of experiments that have been done before starting with formal capture are shown.

Once the database is built, we can proceed to implement the remaining stages proposed in the methodology.

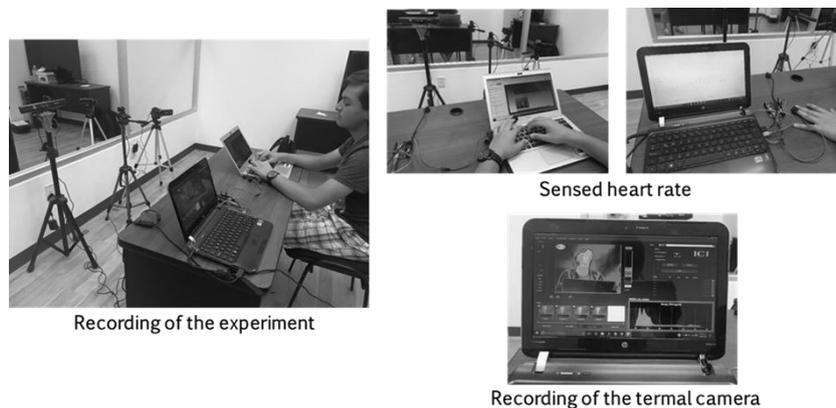


Fig. 3. Pictures of the first experiments.

6 Conclusion

The development of this project will contribute to the identification of gaps in the automatic recognition of emotions centered on learning. It will allow to identify and propose less uncomfortable technologies for students recommended for the collection of data in the identification of emotions centered on learning; we will start with the use of video cameras, thermal cameras, heart rate sensor and Kinect for Windows.

For the development of the proposed methodology, the use of machine learning algorithms is proposed, starting with neural networks and the use of fuzzy techniques. A preliminary analysis of the captured data will allow us to choose the best algorithms according to their characteristics and distribution.

A contribution is expected in the proposal to integrate data from different physiological and behavioral signal acquisition technologies. The data acquisition stage will be carried out through a controlled experiment in which college students will participate using a MOOC to learn algebra. Data of 100 students will be collected. We will try to use a cardiovascular wristband to capture the heart rate that is less obstructive and replace the finger sensor, since we detect discomfort to use the fingers of the left hand in which it is currently placed.

After the data preprocessing will start with the execution of the selected feature extraction and classification algorithms. It is expected to obtain results that contribute to improve the identification of emotions centered in the learning that can be considered for the decision making on pedagogical strategies and learning activities more adapted to each student. The implementation of the process of automatic identification of emotions can be integrated into intelligent tutorials within the student modeling module.

References

1. D’Mello, S., Graesser, A.: Dynamics of affective states during complex learning. *Learn. Instr.*, 22(2), 145–157 (2012)
2. Zatarain-Cabada, R., Barron-Estrada, M.L., González-Hernández, F., Rodríguez-Rangel, H.: Building a Face Expression Recognizer and a Face Expression Database for an Intelligent Tutoring System. In: *Proceeding IEEE 17th Int. Conf. Adv. Learn. Technol. ICALT 2017*, 2161–377X/17, 391–393 (2017)
3. Bosch, N., et al.: Automatic Detection of Learning-Centered Affective States in the Wild. In: *Proceeding 20th Int. Conf. Intell. User Interfaces IUI ’15*, 379–388 (2015)
4. Picard, R.W.: *Affective Computing*. (1997)
5. Fuentes, C., Herskovic, V., Rodríguez, I., Gereá, C., Marques, M., Rossel, P.O.: A systematic literature re-view about technologies for self-reporting emotional information. *J. Ambient Intell. Humaniz. Comput.*, 1–14 (2016)
6. Zatarain, R., Barrón, M.L., Luis, O.J., Martínez, J.A.: Reconocimiento automático y aspectos éticos de emociones para aplicaciones educativas. *Komputer Sapiens*, México, 27–31 (2014)
7. Cowie, R., et al.: Emotion recognition in human computer interaction. *IEEE Signal Process Mag.*, 18(1), January, 32–80 (2001)
8. Steidl, S.: *Automatic Classification of Emotion-Related User States in Spontaneous Children’s Speech*. Universität Erlangen-Nürnberg (2009)
9. Ekman, P.: *Emotions Revealed. Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 1st ed., New York: Henry Holt and Company (2003)
10. González-Hernández, F., Zatarain-Cabada, R., Barrón-Estrada, M.L., Rodríguez-Rangel, H.: Recognition of learning-centered emotions using a convolutional neural network. *Journal Intell. Fuzzy Syst.* (2017)
11. Mehmood, R. Lee, H.: Towards Building a Computer Aided Education System for Special Students Using Wearable Sensor Technologies. *Sensors*, 17(317), 1–22 (2017)
12. Zatarain-Cabada, R., Barrón-Estrada, M.L., González-Hernández, F., Oramas-Bustillos, R., Alor-Hernández, G., Reyes-García, C.A.: Building a Corpus and a Local Binary Pattern Recognizer for Learning-Centered Emotions. *Adv. Comput. Intell.*, II, 524–535 (2017)
13. Zatarain, R., Barrón, M.L., González, F., Reyes-García, C.A.: An Affective and Web 3.0-Based Learning Environment for a Programming Language. *Telemat. Informatics* (2017)
14. Arana-Llanes, J.Y., González-Serna, G., Pineda-Tapia, R., Olivares-Peregrino, V., Ricarte-Trives, J.J., Latorre-Postigo, J.M.: EEG lecture on recommended activities for the induction of attention and concentration mental states on e-learning students. *Journal Intell. Fuzzy Syst.* (2017)
15. Bosch, N., D’Mello, S.: The Affective Experience of Novice Computer Programmers. *Int. J. Artif. Intell. Educ.*, 27(1), 181–206 (2015)
16. Barrón, M.L., Zatarain, R., Hernández, Y.: Intelligent Tutor with Emotion Recognition and Student Emotion Management for Math Performance. *Rev. Electrónica Investig. Educ.*, 16, 88–102 (2014)
17. Zatarain-Cabada, R., Barrón-Estrada, M.L., Ríos-Félix, J.M.: Affective Learning System for Algorithmic Logic Applying Gamification. *Chapter Lect. Notes Comput. Sci.*, 576, no. August (2017)
18. Fayyad, U.M., Shapiro, G.P., Padhraic, S.: From data mining to knowledge discovery: An overview. *Adv. Knowl. Discov. Data Mining*, AAAI Press / MIT Press (1996)