

Building a Corpus of Phrases Related to Learning for Sentiment Analysis

María Lucia Barrón-Estrada, Ramón Zatarain-Cabada,
Raúl Oramas-Bustillos, Sandra Lucia Ramírez-Ávila

Tecnológico Nacional de México Instituto Tecnológico de Culiacán,
Culiacán, Sinaloa, Mexico

{lbarron, rzatarain, raul.oramas, sramirez}@itculiacan.edu.mx

Abstract. Learning-centered emotions unlike basic emotions emerge during deep learning activities and they have an important relation to cognitive processes of students. In this paper we present the creation process of a corpus of phrases (opinions) related to learning computer programming. Opinions (textual phrases), are categorized in different emotions related to learning such as frustrated, bored, neutral, excited, and engaged. The main contribution of this work is the creation of the corpus of sentences related to learning in Spanish that can be used to analyze text dialogs in order to detect the student's emotions and improve the teaching process adapting the content to the cognitive needs and also, to the affective state of the student. We present the study and the results obtained about the corpus of textual opinions.

Keywords: opinion mining, sentiment analysis, intelligent tutoring system, intelligent learning environments, learning-centered emotions.

1 Introduction

Emotions play a fundamental role in education. Learning is traditionally performed only within the classrooms in schools and was considered only the cognitive status of students, however, different technologies have been developed to address the need to integrate the emotions in the teaching-learning process. These technologies include some Intelligent Tutoring Systems (ITS) and Intelligent Learning Environments (ILE), which were designed to capture and identify the emotions of the users, but most of these conventional systems work only with basic emotions [1].

Humans express emotions in different situations of daily life. Basic emotions like anger, happiness, sadness, and fear have been studied by Ekman [2]. However, there are other types of emotions that emerge during deep learning activities, among which are emotions like confused, boredom, frustration, engagement, and excitement; these are known as emotions centered on learning [3, 4]. Emotions centered on learning play an important role in students because they affect different aspects such as cognitive

mechanisms and retention of information [5, 6]. There are many different ways to recognize emotions (face image, voice, text, corporal expression, heart rates etc.) but one of the most used today is based on the textual opinions expressed by users in several social networks. Currently, most of the tools to recognize emotions in text work with phrases written in English and those phrases are not focused on the learning field. In order to include an emotion recognition module based on textual opinions in an Intelligent Tutoring System we decided to create a corpus of phrases written in Spanish which are related to learning.

In this paper, we describe the development of an Educational Resources Assessment System, which is used to show the student several learning objects (LO) in order to capture textual opinions about them. The system generates a database (corpus), which contains the opinions provided by the students during the learning process when he/she access the educational resources. This corpus is used to train a classifier to recognize emotions and this classifier can be used to incorporate changes and improvements to the contents of the courses and other didactic elements in both ITS and ILE.

The main contribution of this work is the creation of a new database of phrases written in Spanish related to learning in the field of computer program development; to achieve this, a system was developed to capture and store students' opinions about learning objects; and finally, the implementation of a classifier algorithm to recognize learning emotions based on sentences written in Spanish.

This article is organized as follows: Section 2 presents the related work, Section 3 describes the Educational Resources Assessment System. Section 4 shows the process of creating the corpus of textual opinions and finally, conclusions are presented in section 5.

2 Related Work

In the last decade, the proliferation of internet sites where users express their opinions provoked as result the need to process these opinions automatically to obtain relevant information that could be used to make decisions so arose the opinions mining area. For Feldman “Sentiment analysis (or Opinions Mining) is defined as the task of finding the opinions of the authors about specific entities” [7]. Medhat, Hassan, and Korashy [8] distinguish Sentiment Analysis (SA) from Opinion Mining (OM) defining SA as “the computational study of the opinions, attitudes, and emotions of the people towards an entity”. They said that OM extracts and analyzes the opinion of the people about an entity while SA identifies the sentiment expressed in a text and then analyzes it. Thus, the objective of SA is to find opinions, identify the feelings they express and then classify its polarity.

Medhat, et al [8], propose three main levels of SA classification: document level, sentence level, and aspect level.

- *Document Level SA*: considers the whole document as a basic unit of information and assumes that the document contains the main opinion expressed by the author.

- *Sentence Level SA*: this level is intended to classify the feeling expressed at the level of each sentence.
- *Aspect Level SA*: intends to classify the sentiment with respect to the specific aspects of the entities.

Sentiment Analysis is important in different areas like Business Intelligence where tracking public viewpoints could lead to prediction in sales as well as to better reputation management and public relations [9], but in education can be used to collect students' opinions on topics or learning strategies and this information will serve to perform the analysis of feelings and its outcome may be used to improve course materials, teaching strategies, the personalization of contents, among others as well as can be used within Intelligent Tutoring Systems (ITS).

Sentiment analysis has been used in the educational field in several works as presented below. Altrabsheh, Gaber, and Cocea [10] present Sentiment Analysis for Education (SA-E), where they describe an architecture to analyze the feedback from the students using the sentimental analysis on Twitter. Here the students used Twitter to express their opinions about the learning material of the course and those opinions were used by the teacher to adjust the teaching style according to the results.

Ortigosa, Martin, and Carro [11] implemented a hybrid method for sentiment analysis in SentBuk which is a Facebook application to retrieve users' messages and classify them according to their polarity. It also provides several features like detecting emotional change, and user classification according to their messages, among others.

The SA method combines two techniques: lexical-based and machine learning, and reach an accuracy of 83.27%. This tool can be used in an e-learning context allowing the personalization of activities for the user according to his/her emotional state.

There are other researchers who have developed works related to this: Rowe [12] in which provides an overview of the few developments in the field of student emotions in relation to feedback based on recent research from social psychology and education. Altrabsheh, Cocea, and Fallahkhair [13] examined different methods so they could use to learn the feeling of feedback from students. Munezero, Montero, Mozgovoy, & Sutinen [14], present a functional system to analyze and visualize the student's emotions expressed in the learning diaries (instruments in which students reflect on their learning experience). In all these works, all the tools developed to analyze student's emotions were based on sentences written in English.

3 Technology

This section presents the development of the Educational Resources Assessment System (ERAS), which was used to create the corpus of textual opinions written in Spanish for the programming area. It also describe the module of analysis of feelings that uses the corpus of opinions for training. This module also performs the classification of opinions by determining the positive, negative or neutral polarity of the student opinion.

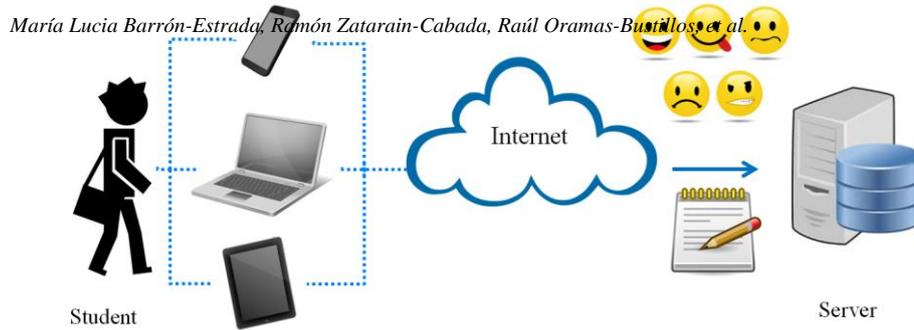


Fig. 1. Functioning of Educational Resources Assessment System.

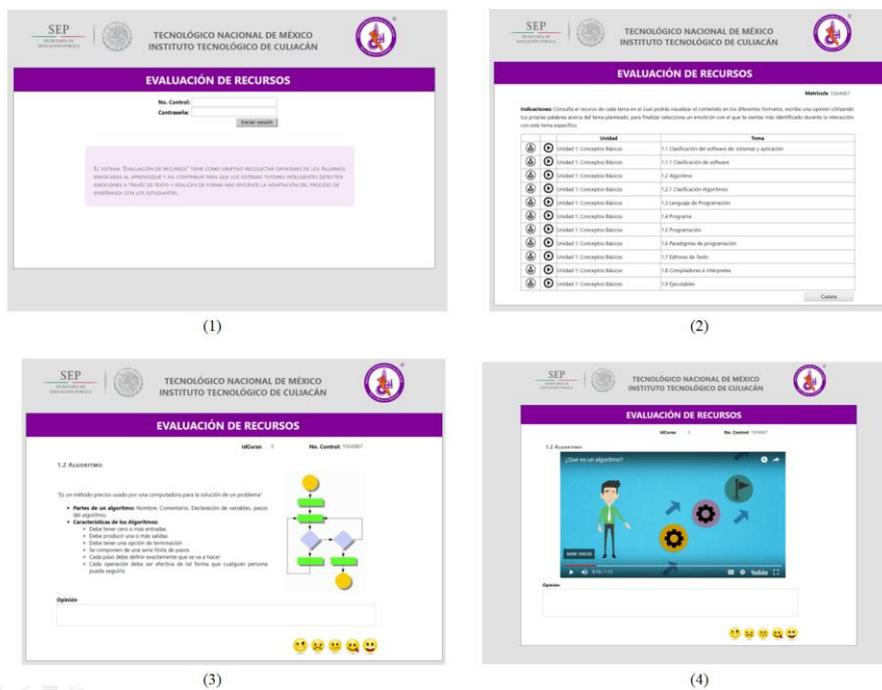


Fig. 2 System Interfaces Resource Evaluation.

3.1 Educational Resources Assessment System

The development of the Educational Resources Assessment System (ERAS) was implemented to generate a corpus of student’s opinions written in Spanish which are focused on learning. EARS is a web application that was developed for the .NET platform in Visual Studio 2013, with SQL Server 2008 R2 database to store course information, educational resources, learning objects, users, and opinions issued by students. ERAS is hosted in the URL (<http://posgradoitc.ddns.net:8000/>).

ERAS was designed to interact with students in order to allow them to express their opinions and comments freely about the educational resources or learning objects of the subjects of a course.

Figure 1 shows how the student interact with the system using different devices with Internet access such as Tablet, Smartphone, Laptop, or PC. The user login to ERAS using his/her personal account and password and then selects the course that he/she wants to study. The system displays the list of subjects and the educational resources available and then the user access a learning object and writes a textual opinion labeled with the emotion that the student expresses at that moment. All student's opinions are registered in a database which is composed of different fields as the student data, the selected course topic, the opinion and evaluation entered by the student and finally the date and time of the student's opinion.

ERAS contains several interfaces, but there are four main GUIs shown in Fig.2. The login interface shows a brief introduction of the objective of the system and allows users to login (see Fig 2(1)). In Fig. 2(2) the second interface presents the topics of the subject selected by the student; there are two different formats (Image-Text and Video) to show the subject of study. Fig. 2(3) shows the learning object selected in format text, image or both and finally in Fig. 2(4) the theme in video format is shown. In both formats, the student is requested to enter a sentence to express an opinion about the learning object and tag it with an emotion related to learning (frustrated, bored, Neutral, Excited and engagement) using an emoticon provided in the interface.

3.2 Sentiment Analysis (SA) Module

This module is responsible for determining the polarity (positive, neutral or negative) of a phrase, sentence or document.

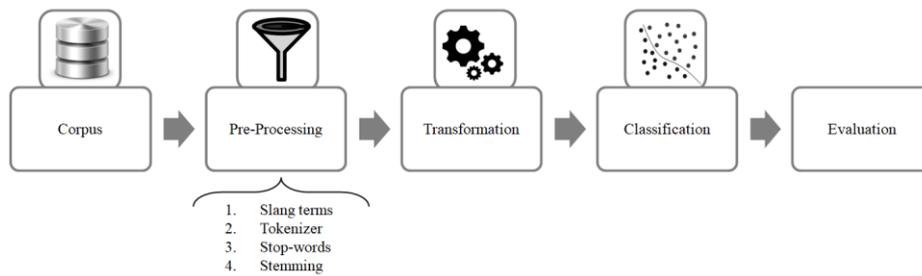
SA module [15], was developed using a Bernoulli Naive Bayes classifier after doing several tests with different classifiers: Multinomial Naive Bayes, Support Vector Machine, Linear Support Vector Machine, Stochastic Gradient Descent Classifier, and K-Nearest Neighbors.

There are a series of metrics that are used to estimate the quality of the classification algorithm that is part of our learning model. The simplest method to calculate the effectiveness of a classifier is the accuracy measure that calculates the percentage of text document correctly classified over the total of documents to be classified. To obtain this value we applied a cross-validation technique with a dataset of 8174 texts in Spanish with 90% for the training data and 10% for the test data.

Table 1. Accuracy values obtained using different classifiers.

Classifier	Accuracy
Bernoulli NB	76.77%
Multinomial NB	75.31%
SVC	75.79%
Linear SVC	74.69%
SGDC Classifier	76.69%
KNN	68.46%

The classifier algorithms we used for the evaluation were: Bernoulli Naive Bayes, Multinomial Naive Bayes, Support Vector Machine, Linear Support Vector Machine,



María Lucia Barrón-Estrada, Fig 3 The algorithm of the SA Module, et al.

Stochastic Gradient Descent Classifier, and K-Nearest Neighbors (KNN). Table 1 shows the accuracy values obtained according to the size of the corpus. The classifier with the highest score was Bernoulli Naive Bayes with an accuracy of 76.77%. In general, Bernoulli Naive Bayes is a robust classifier against irrelevant features and it is suitable for classifying texts.

To test the functioning of the SA module, we used a corpus of phrases labeled with their respective polarity which was collected from Twitter, this corpus is called TASS [16]. Figure 3 shows this process.

- A. *Corpus*: a corpus of textual opinions is required as input to the module.
- B. *Pre-Processing*: It normalizes the sentences performing 4 steps for each entry in the corpus. The preprocessing is executed in 4 steps:
 1. Slang terms: It moves jargon terms and emoticons to its equivalent text.
 2. Tokenizer: It converts sentences into words by removing points and signs.
 3. Stop-words: It reduces words to its root word.
 4. Stemming: It removes unnecessary words.
- C. *Transformation*: It calculates the term frequency (number of times a given term appears in a document or dataset) and reverses term frequency (number of documents in which a given term appears) for each word in the corpus generating a TF-IDF matrix. This is known as extraction of weighting characteristics.
- D. *Classification*: It defines a function to predict the tag entered as input.
- E. *Evaluation*: It evaluates the learning machine model to predict the polarity (positive, negative) of an input text; this helps us to find a reliable model.

4 Evaluation and Results

In this section, we describe the study and the results obtained about the corpus of textual opinions written in Spanish created by using the Educational Resources Assessment System (ERAS).

4.1 Participants

The study was conducted in August 2017; a total of 53 students, from the computer systems engineering career at the Instituto Tecnológico de Culiacán, participated from which 45 were male and 8 female.

4.2 Material

For this study, ERAS shows the Fundamentals of Programming course. This course contains five chapters. Chapter 1 is related with introduction and vocabulary about programming topic. It contains 10 topics; most of them are presented using learning objects in two formats (text-image or video). All LO were available for access by all students. In this study, the other chapters of the course were not available.

Table 2. Some examples of the opinions in the Corpus related to learning. *Building a Corpus of Phrases Related to Learning for Sentiment Analysis*

Key	Opinions (in Spanish)	Format	Evaluation
E-15	I liked the video very much (Me gusto bastante el video)	Video	Engagement
E-64	I didn't like that the voices were from Spain, besides that figures in the animation are kind of ugly. (No me gusto que las voces fueran de España, además los gráficos de la animación están algo feos.)	Video	Frustrated
E-70	The video is good but I think I need to go deeper (el video es bueno aunque creo que le falto profundizar mas)	Video	Neutral
E-29	Wow, it is kind of complex (Vaya, es algo complejo)	Video	Frustrated
E-67	Perhaps with an example it would be clearer. (Quizás con un ejemplo quedaría más claro.)	Image- Text	Bored
E-71	It would have been better to put a table with their differences and then compare them, and be more diverse. (Hubiese sido mejor poner una tabla con sus diferencias y así compararlas y sea más diverso.)	Image- Text	Excited

4.3 Evaluation of Educational Resources Assessment System

The system was available for 10 days in order to allow students to login and access the learning resources.

After accessing each learning object of a specific topic in any format (text-images or video), the system ask the students to write their opinion with a sentence (15 to 255 characters long) with respect to the content of the studied topic, they also have to label the opinion with an emoticon representing their emotional state: frustrated, bored, neutral, excited and engagement. The system recorded a total of 851 student´s opinions labeled with an emotion. Table 2 shows some examples of the student´s opinions that are in the corpus.

With the information obtained, a group of teachers validated that the opinions of the students had a relationship with the emotion captured with the emoticon which was associated. A 5% of inconsistencies were found and rectifications were made. It was also detected that 20% written opinions were definitions of basic terms such as algorithm, computer, among others so that these opinions were considered with the neutral emotion. Subsequently, we used the Feeling Analyzer Module to validate the coincidences with the emotional content of the texts, considering frustrated, bored, as a negative and neutral emotion, excited, and engaged as a positive emotion, since the Feeling Analyzer Module only recognizes polarity positive or negative. We found 458

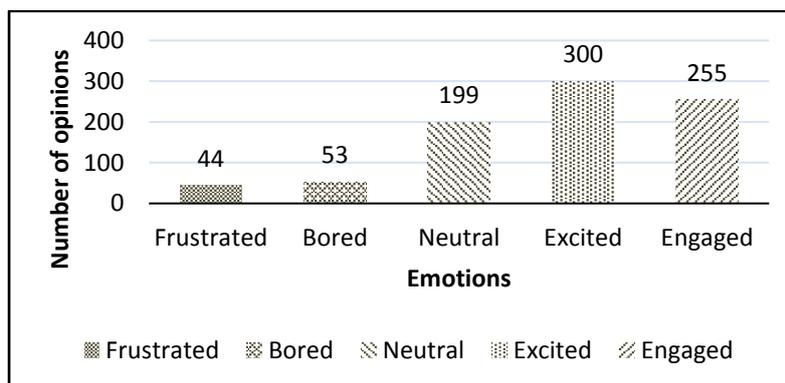


Fig. 4. Distribution of opinions in the corpus.

Table 3. Accuracy values obtained using different classifiers.

Classifier	Accuracy
Bernoulli NB	40.70%
Multinomial NB	40.70%
SVC	33.72%
Linear SVC	30.23%
SGDC Classifier	23.26%
KNN	36.05%

matches which are 53.81% of the total 851 opinions. Finally, the 458 emotions were incorporated into the corpus of the Feeling Analyzer Module.

4.4 Corpus of Reviews Focused on Learning

One of the main objectives of the ERAS was to collect opinions from students to generate a new corpus of opinions in Spanish to recognize emotions focused on learning (frustrated, bored, neutral, excited and engagement) with the aim of SA module which only recognizes two emotional states: positive and negative.

Of the 851 opinions (754 Positives and 97 Negatives), there was a neutral, excited and engaged emotional tendency which indicates that we must include resources that induce students to write negative comments. Figure 4 shows the distribution of the opinions collected in the first experiment.

With the new corpus based on emotions (frustrated, bored, neutral, excited, engaged); obtained with the ERAS application, the following classifiers were trained: Bernoulli Naïve Bayes (BNB), Multinomial Naïve Bayes (MNB), Support Vector Machine (SVM), Linear Support Vector Machine(LSVM), Stochastic Gradient Descent (SGD), and K-Nearest Neighbors (KNN); using cross-validation of 90% for training data and 10% for test data.

Table 3 shows the accuracy values obtained according to the size of the corpus. The classifiers with the highest score were BNB and MNB, both with an accuracy of 40.70%. The level of prediction is low since the corpus size is small and it is not balanced, as shown in figure 4.

5 Conclusions

In this paper, we describe the development of a system to generate a corpus of textual opinions in Spanish, labeled with emotions based on learning. The corpus generated with the ERAS system contains 851 textual opinions and each of them is tagged with an emotion related to learning.

At this time, the ERAS system remains available for more participants to express their opinions on educational resources and we expect to expand the number of opinions in the corpus in a few months balancing the occurrences of each of the emotions related to learning in order to improve the accuracy of the classifier.

On the other hand, the corpus will be used later for the administrator of an Intelligent Learning Environment to make decisions about the relevance of the educational resources contained in the system and to propose changes or improvements. This will help Intelligent Tutoring Systems detect emotions through text and make the teaching process more efficient for students, adjusting the content to the particular needs of each of them.

For future work we will take into consideration the inclusion of exercises for the development of Java programs in the ERAS system, this will also create a bank of problems for the learning of computer programming as well as will expand the corpus of opinions in this area of knowledge.

References

1. D’Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Graesser, A.: AutoTutor detects and responds to learners affective and cognitive states. In: Workshop on emotional and cognitive issues at the international conference on intelligent tutoring systems, pp. 306–308 (2008)
2. Ekman, P.: An argument for basic emotions. *Cognition and Emotion*, 6, pp. 169–200 (1992).
3. Baker, R. S., D’Mello, S. K., Rodrigo, M. M. T., Graesser, A. C.: Better to be frustrated than bored: The incidence, persistence, and impact of learners’ cognitive affective states during interactions with three different computer-based learning environments. *Int. J. Hum-Comput. Stud.* 68(4), pp. 223–241 (2010)
4. D’Mello, S., Graesser, A.: Dynamics of affective states during complex learning. *Learning and Instruction* 22(2), pp. 145–157 (2012)
5. Pekrun, R.: The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology* 41(4), 359–376 (1992)
6. Pekrun, R., Goetz, T., Titz, W., Perry, R. P.: Academic emotions in students’ self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist* 37(2), pp 91–105 (2002)
7. Feldman, R.: Techniques and applications for sentiment analysis. *Communications of the ACM* 56(4), pp 82–89 (2013)
8. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal* 5(4), pp.1093–1113 (2014)
9. Pang, B., Lee, L., Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), pp 1–135 (2008)
10. Gaber, M. M., Altrabsheh, N., Cocea, M.: SA-E: Sentiment Analysis for Education. *Frontiers in Artificial Intelligence and Applications*, vol. 255, (2013)
11. Ortigosa, A., Martín, J. M., Carro, R. M.: Sentiment analysis in Facebook and its application to e-learning. *Comput. Hum. Behav.* vol. 31, no. 1, pp. 527–541 (2014)
12. Rowe, A. D., Carless, D., Bridges, S. M., Chan, C. K. Y., Glofcheski, R.: Feelings about Feedback: The Role of Emotions in Assessment for Learning. In: *Scaling up Assessment for Learning in Higher Education*, Springer Singapore, pp. 159–172 (2017)
13. Altrabsheh, N., Cocea, M., Fallahkhair, S.: Learning Sentiment from Students’ Feedback for Real-Time Interventions in Classrooms. In: *Adaptive and Intelligent Systems: Third International Conference, ICAIS 2014, Bournemouth, UK, September 8-10*, Bouchachia, A. (Ed.), Springer International Publishing, pp. 40–49 (2014)

María Lucia Barrón-Estrada, Ramón Zatarain-Cabada, Raúl Oramas-Bustillos, et al.

14. Munezero, M., Montero, C. S., Mozgovoy, M., Sutinen, E.: Exploiting sentiment analysis to track emotions in students' learning diaries. In: Proceedings of the 13th Koli Calling International Conference on Computing Education Research, pp. 145–152 (2013)
15. Barrón-Estrada, M.L., Zatarain-Cabada, R., Oramas-Bustillos, R., González-Hernández, F.: Sentiment Analysis in an Affective Intelligent Tutoring System. In: 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, pp. 394–397 (2017)
16. Villena-Román, J., Martínez-Cámara, E., Lana-Serrano, S., González-Cristóbal, J. C.: TASS - Workshop on Sentiment Analysis at SEPLN. In: TASS - Taller de Análisis de Sentimientos en la SEPLN, pp. 37–44 (2013)