SentBuk: Sentiment analysis for e-learning environments

José M. Martín, Alvaro Ortigosa, Rosa M. Carro
Department of Computer Science, Universidad Autónoma de Madrid
Francisco Tomás y Valiente, 11
28049 Madrid, Spain
joseph.martin@estudiante.uam.es, {alvaro.ortigosa, rosa.carro}@uam.es

***Abstract*— This paper presents SentBuk, a Facebook application that extracts information about the user sentiment automatically, in a non-intrusive way. It performs sentiment analysis of user writings in Facebook walls, classifying each sentence as positive, neutral or negative. Finally, the overall user sentiment is calculated. On one hand, this information is useful to enrich user models for adaptive e-learning systems, so that these systems can adapt any of their aspects (tasks to be proposed to the student, contents, and so on) according to each student sentiment, among other criteria. On the other hand, the polarity of the emotions transmitted by the students enrolled in a course can constitute a useful feedback for the course teacher.**

User modeling; sentiment analysis; adaptive e-learning systems

I. INTRODUCTION

Adaptive methods and techniques constitute powerful tools to support personalized learning. In order to provide adaptation it is necessary to store and maintain information about the students in what is called the user model [1]. Depending on the goal of each adaptive e-learning system, it will be useful to represent and use certain information about the student: features, preferences, needs, behaviors, context, previous actions, etc. Both in face-to-face and in e-learning environments, student emotional state influence the way in which students tackle the learning process: affective and emotional factors, among other aspects, seem to affect the student motivation and, in general, the outcome of the learning process [2].

On one hand, detecting and managing the student sentiments at a certain time can help to know their potential needs at that time, making it possible to contribute to fulfill them. In this sense, students can benefit from adaptive e-learning applications that, taking into consideration information about each student sentiment, are able to adapt their behavior accordingly. On the other hand, knowing the student feelings towards an ongoing course can be used as feedback for the teachers: if too many negative messages are written in a forum related to this course, then something may be happening and the teacher should know about it. Sentiment is defined as "a personal positive or negative feeling or opinion". An example of a sentence transmitting a positive sentiment would be 'I love it!', whereas 'It is a terrible movie' transmits a negative one. A neutral sentiment does not express any feeling (e.g. 'I am commuting to work').

The most commonly used procedure to obtain information about the students consists of asking them to fill in questionnaires. However, users can find this task too time-consuming. We think that information for student models should be obtained as unobtrusively as possible, yet without compromising the reliability of the model built. In this paper, we present a method for eliciting student sentiment (positive, neutral or negative) by analyzing what they write in Facebook.

The rest of this work is organized as follows. Section II describes related work in the areas of adaptive e-learning and sentiment analysis. Section III presents the method proposed to classify messages written by users as positive/neutral/negative. Section IV describes SentBuk, the Facebook application that serves as a frontend for the sentiment classifier. Section V presents the results obtained; section VI elaborates on some examples of use in e-learning systems and, finally, section VII presents the conclusions along with current and future work.

II. RELATED WORK

The earliest applications of Adaptive Hypermedia (AH) date from the 1990s decade. A first classification of AH methods and techniques was presented by Brusilovsky in 1996 [3]. Since then, AH has been used in the context of e-learning to support personalized learning, by recommending the most suitable tasks to be accomplished by each student at each time, as well as the most appropriate multimedia contents to be presented to each of them, according to each one personal features, preferences, previous actions, context, etc. Some well-known adaptive e-learning systems are ELM-ART [4], AHA[5], TANGOW [6], WHURLE [7] and COMOLE [8]. All the student data to be considered for adaptation must be stored in the user model [1]. These data can be obtained from the user through specific tests, such as those related to learning styles [9], personality [10] or intelligence [11]. More recently, some of these data have been inferred from the student interactions, to prevent students to fill in multiple questionnaires before entering a new course. We have some previous experience on inferring learning styles [12] and personality [13].

Regarding the area of sentiment analysis, the survey written by Pang and Lee [14] covers the most popular techniques and approaches. The earliest works focusing on classifying words or phrases according to semantic issues date from the late 1990s, and used linguistic heuristics or pre-selected sets of seed words [15] [16]. The results obtained in these works served as

This work has been funded by the Spanish Ministry of Science and Education, project ASIES (TIN2010-17344) and by Comunidad Autonoma de Madrid, project E-Madrid (S2009/TIC-1650).

the basis for classifying entire documents, considering that the average semantic orientation of the words in a review may be an indicator of whether the text is positive or negative [17]. With the popularization of Internet during the 2000s and other factors, such as the rise of machine learning methods and the availability of labeled datasets from websites, the area became very popular. One example of using machine learning techniques to process data retrieved from websites is presented in [18]. It compares different techniques to classify movie reviews, obtaining 82.9% of success when applying Support Vector Machines (SVM). Generally, it is difficult to obtain better results due to the characteristics of natural language. For example, irony or sarcasm are difficult to detect (even for a human being). Sentiment polarity of a text can change completely if they are not detected [14].

However, in specific domains, the use of machine learning algorithms for classifying texts according to their sentiment orientation performs well. Nevertheless, when considering different domains, the lexicon approach obtains better results, since it analyzes the text grammar [19].

In relation to the language, there are very few works focusing on languages other than English. Those works are usually applications of the methods presented in previous ones (for English) to a different language. For example, [20] applied different machine learning techniques in order to classify movie reviews, achieving an interesting 86.84 % of success with SVM in the Spanish language.

Finally, in recent years, due to the increasing ease of obtaining large collections of data through social networks, many researchers are trying to apply these techniques to them [21] [22]. However, all these works are focused on the English language and on Twitter, since it is quite easy to retrieve data from this social network [21].

III. Method Proposed

In this work we propose to obtain information about user emotional state by obtaining the messages written by him on the Facebook wall and applying sentiment analysis to them. The information obtained from this analysis constitutes an input for student models of e-learning systems such as COMOLE. The basis of the method developed consists of a sentiment classifier of messages written in Spanish and extracted from Facebook. The proposed classifier follows a lexicon-based approach, using a dictionary of words annotated with their semantic orientation (positive/negative emotional polarity) and detects additional language features such as positive interjections (i.e. laughs), negative interjections, emoticons, misspells, part of speech tagging or negation (polarity shifter).

A. Features

1) Polarity lexicon

Firstly, the dictionary of polarity words was created from the Spanish LIWC categorization (Linguistic Inquiry and Word Count) [23]. Although LIWC has formed positive and negative emotional categories, we have increased the set of words for each category adding other categories that can express emotional feelings in order to have a bigger dictionary. The resulting classes, with their respective LIWC categories, are:

- Positive: positive emotions, positive sentiment, optimistic
- Negative: negative emotions, anger, sadness, death, to swear

After removing duplicate words, the positive class contains 653 words and the negative one contains 894 words.

2) Interjection and emoticons

Secondly, the interjections that express laughs in the Spanish language (e.g. ‘jaja’, ‘jeje’) are marked as positive. Other interjections, such as ‘uff’ (tiredness) or ‘jo’ (upset) are marked as negative. In addition, the emoticons are commonly used in Facebook to express sentiment in a short way, so a list of positive and negative emoticons was also included.

3) Spelling checker

The spelling checker (GNU Aspell) was included to overcome the problem that Facebook writing is usually informal, with a large presence of misspellings. However, this feature must be used carefully, because a suggested word from the spelling checker can produce a wrong result. For example, the suggested correction for ‘Dani’ (a Spanish popular name) is ‘daño’ (damage, a negative word). To avoid these situations, a list of words that cannot be corrected is considered.

4) Part of speech tagging

The part of speech (POS) tagger implemented in the classifier is used for semantic disambiguation, e.g. ‘vino’ can be a noun (wine) or a verb (came). The POS tagger was built automatically based on a training corpus considering bigrams, the CESS-ESP [24] of the Natural Language Toolkit (NLTK) [24]. The built process has been smoothed with a backoff, labeling as a noun by default. The evaluation over a test set gave an 81.91% success. To improve the automatic tagger, the suffixes of the non-labeled words are analyzed, since, in Spanish, the suffix can indicate the kind of word. For example, almost certainly, a word ending in ‘aríamos’ is a verb or a word ending in ‘ísimo’ is an adjetive [26].

5) Negation

Finally, negations are detected. A negation implies the need of reversing the polarity of the corresponding phrase. Consequently, a chunking parser was incorporated to the classifier, using a regular expression grammar based on the Spanish grammar defined in [26]. In Spanish, negation is generally expressed by placing a negative adverb, so the classifier looks for negative adverbs detected through the POS tagger. The chunking parser has also been developed with NLTK [24].

B. Polarity computation

Firstly, the sentiment classifier receives a Facebook message. This message requires a preprocessing step before obtaining the polarity score. In this step, the message is, firstly, converted to lower case. Secondly, links are removed, since the URLs can contain polarity words. Finally, the words that, together, are detected as an adverb by the POS tagger are joined into a unique word (e.g. the words 'de mala fe' are converted to 'de_mala_fe'). After the preprocessing step, the message (M) is segmented into sentences (S), which are segmented into tokens (see (1)). Two tokenization steps take place. The first one considers only whitespaces as delimiters. After identifying the emoticons, the second tokenization step takes into account all the usual delimiters (e.g. commas, dots, whitespaces and so on).

$$\mathrm{M} = \{S_1, S_2, \dots, S_n\} \qquad \mathrm{S} = \{T_1, T_2, \dots, T_n\} \tag{1}$$

Afterwards, with the purpose of removing the ambiguity of certain words, POS tagging is applied to the tokens obtained. Then a score is assigned for each token: 1 if the polarity is positive, 0 if it is neutral or -1 if it is negative. To assign a score (δ), the classifier checks if the token is an emoticon (E) or an interjection (I), or whether it matches one word stored in the sentiment lexicon (L).

Messages in Facebook contain very casual language. It is usual to find words with repeated letters (e.g. 'hooolaaaaa'). Therefore, any letter occurring more than twice in a row is replaced with two occurrences and is analyzed again. If it fails, the repeated letters are replaced with only one occurrence. If it fails again, the words are reduced to their lemma, because the dictionary does not contain all the possible variants for each word. Lastly, if the token does not match, it is checked with the spelling checker. If all the previous steps fail, the word is labeled as neutral.

Once a score is assigned to each token, the chunking parser is applied to the sentence, to detect possible negations. When a negation is detected in one phrase, all the token scores for this phrase are inverted. Since emoticons are not affected by negation, the original score assigned to them is maintained. Then all the tokens scores are summed up (θ) and returned to the message processing function (see (2)). The number of tokens that constitute polar words (γ) according to their grammatical category (i.e. noun, adjective, interjection or verb) is calculated (see (3)). Other types of words, which are usually the most frequent words appearing in a text (i.e. determinant, prepositions, etc.), are not considered, because they are 'stopwords' for a sentiment analysis.

$$\theta(\mathrm{S}, \mathrm{L}, \mathrm{E}, \mathrm{I}) = \sum_{T_i \in L}^{i=n} \delta + \sum_{T_i \in E}^{i=n} \delta + \sum_{T_i \in I}^{i=n} \delta$$

$$\delta = \langle -1 \; if \; T_i \; has \; negative \; polarity \; | 0 \; if \; T_i \; has \; neutral \; pol. | +1 \; if \; T_i \; has \; positive \; pol. \rangle \tag{2}$$

$$\gamma(\mathrm{S}) = \sum_{T_j \in (noun, adjetive, verb, interj)}^{n} 1 \tag{3}$$

Finally, the total polarity score (X) is the sum of the scores divided by the sum of all tokens that could have been considered as polar words in each sentence (γ) (see (4)). If γ is zero, division is not applied. An example of γ=0 is a message consisting in one emoticon.

$$\mathrm{X(M)} = \frac{\sum_{S_i \in M}^{i=n} \theta_i}{\sum_{S_i \in M}^{i=n} \gamma_i} \tag{4}$$

The obtained score must be a real number in a range of -1 to +1. Therefore, if the result turns to be out of the range, it is fixed following these rules:

If X > 1 then X=1.
If X < -1 then X=-1

This situation can happen with emoticons, since do not compute for the calculus of γ, because of tokenizer implementation issues. An emoticon usually defines by itself the overall polarity of a message and is used in some works as a method of distant supervision to classify the messages [21]. Finally, we noticed that a message was classified as neutral when the numbers of positive and negative tokens were equal ('sentiment neutralization'). Then we decided for the classifier to give more weight to the tokens placed at the end of the sentence, since the position of terms in a text can influence the overall sentiment, being the words at the end more relevant [18] (e.g. 'you are pretty and stupid'). A schema of the classifier pipeline is shown in figure 1.

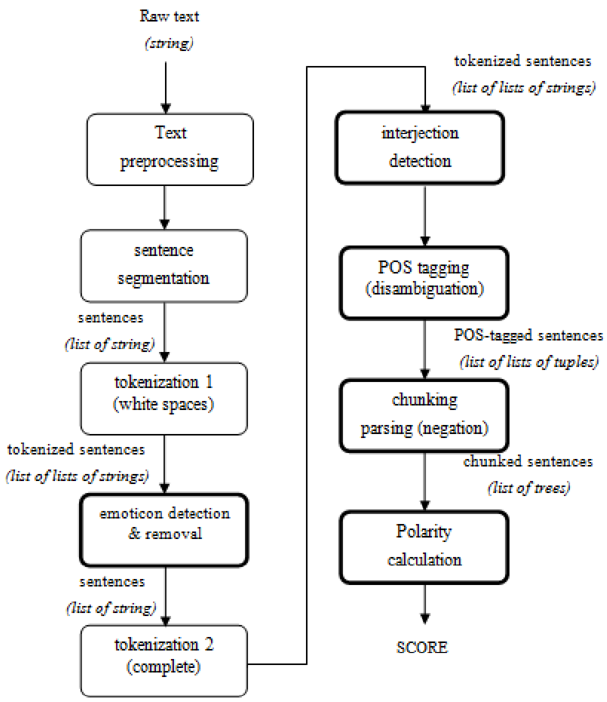

Figure 1. Classifier pipeline

IV. THE TOOL: SENTBUK

To visualize all the messages and their score, a Facebook application called SentBuk was implemented. SentBuk is an acronym of SENTiment faceBUK (Facebook forbids to register applications containing the word 'book'). It runs on Web browsers that support Javascript, and is a frontend for the sentiment classifier. When a user logs into the application, it asks the user for permission to access the data. SentBuk uses a language detector (oice.langdet) to avoid including those messages that are not written in Spanish. Then it shows the messages from the wall along with the corresponding scores as calculated by the sentiment classifier. Results are shown graphically too, including charts that summarize the results obtained (see figure 2). In addition, a tag cloud is built showing the most frequents words in the retrieved messages.

Figure 2. SentBuk screenshot

Messages can be filtered by data range or by their type (status messages versus all messages). Searching over public walls is also possible. SentBuk shows tables and charts with statistics related to the data retrieved from a user and the results obtained from the sentiment classifier through the "profile window". Finally, the tool provides feedback functionality, to let the user report wrong results.

V. RESULTS OBTAINED

From the messages stored in the database and classified by SentBuk, 1000 of each class (positive, neutral and negative) were randomly selected. These messages were delivered to a human, who manually classified them, without knowing the result of the classifier. Afterwards we compared the results in order to know the accuracy of the exposed method (Table I). From 1000 messages classified as positive by SentBuk, 920 were classified as positive by the user (95,63% of success when leaving out messages written in other languages).

TABLE I. MESSAGE CLASSIFICATION (1[ST] SET)

Predicted class	*User classification*				
	Positive	*Neutral*	*Negative*	*Other language*	*% Success (Spanish)*
Positive	920	23	19	38	95.63%
Neutral	74	704	65	157	83.51%
Negative	89	109	760	42	79.33%

It should be borne in mind that the classifier returns a score, and several messages from this set had scores near to zero. These ones do not express a strong sentiment. We took a second set of 500 messages scored as showing a highly positive emotion plus 500 messages showing negative sentiments (scores near to 1 or -1 respectively), and the results obtained were better, mainly for negative messages (see Table II): 481 messages out of 500 classified by SentBuk as positive were also classified as positive by the user (96,78% success for messages written in Spanish), while 455 out of 500 were certainly negative according to the user (94,59% success).

TABLE II. MESSAGES CLASSIFICATION (2[ND] SET)

Predicted class	*User classification*				
	Positive	*Neutral*	*Negative*	*Other language*	*% Success (Spanish)*
Positive	481	13	3	3	96.78%
Negative	6	20	455	19	94.59%

This high accuracy in the classification surprised us. We analyzed the messages used in this second set and it turned out to be that most of them were related to greetings. We have planned to separate the messages to be analyzed according to their type (greetings in other's wall, status, and so on) to try to find out whether some of them reflect the user emotion better than others (we will comment more about this when describing current and future work).

In any case, the goal of this classifier is to be able to detect positive/negative emotions to adapt learning activities and contents to be proposed/shown to each student accordingly. With this purpose in mind, what we really need is to be able to detect students with highly positive or highly negative emotions, since sentiment-based adaptation will be devoted to them (none sentiment-based adaptation will be done when the student emotion detected is neutral). With this goal, a classifier with high accuracy for identifying these extremes is useful.

VI. EXAMPLE OF USE IN COMOLE

As it has been stated, the information about the user's positive/negative sentiment obtained through this approach can be useful in the context of e-learning systems. Specifically, it can enrich the adaptation processes carried out by adaptive e-learning systems such as CoMoLE (*Context-based adaptive Mobile Learning Environments*). CoMoLE supports: i) the recommendation of Web-based individual and collaborative activities to be accomplished by students, according to their needs, preferences, previous actions or context (location, available time and device); ii) the realization of these activities through individual/collaborative workspaces, dynamically generated according to the user features along with the activity to be tackled, by selecting the most suitable content versions and tools. The recommendation mechanism is based, firstly, on different filters and recommendation criteria (related to user features, actions and contexts), either available within the environment or established by teachers; secondly, it considers information about previous actions made by similar students in analogous situations. CoMoLE stores all the information about the users in the user model, with recommendation purposes. More details about CoMoLE can be found at [8].

Knowing the user positive/negative emotion of a student at a certain time can be used by CoMoLE to recommend activities adapted to him at that time, considering not only his personal features, actions, preferences, location, device or available time, but also his emotional state at that time. This is especially interesting in the framework of dynamic context-based adaptation. For example, when a student with positive feelings connects to CoMoLE to do some work, from the set of activities available for this student according to all those features mentioned above, it would be nice to propose him/her collaborative activities so that he can spread their enthusiasm and good spirit to the others. In fact, this student is expected to engage more easily in any task suitable for him.

However, when a student with negative emotions connects to CoMoLE, it would make sense to propose him, firstly, the most motivating/engaging tasks so that his emotional state can rise up and, hopefully, he enjoys the learning experience more. These tasks could be selected from the set of available tasks, if labeled as "motivating". Apart from including the possibility of labeling tasks in CoMoLE (independently from establishing the task type), our recent works related to emotions and motivational tasks led us to include a new type of activity in CoMoLE: the motivating task. A motivating task is considered by the system as any other type of activity (theory, example, solved-problem, self-evaluation exercise, collaborative task, messaging and so on). As the rest of the activities, it can include contents to be delivered through the Web (images, videos, simulations, games, and so on). A "motivating task" can be proposed to a student when considered convenient (according to his/her features, actions, context and, specially, emotional state, among others). The recommendation criteria to decide when a motivational task should be proposed to the students must be specified in the corresponding filters and adaptation rules.

With respect to multimedia content adaptation, adaptive e-learning systems such as CoMoLE support the selection of different content versions adapted to each student, from a pool of content fragments and diverse multimedia materials. This technique has been used, for example, for teaching Mathematics to children through adaptive educational games, by adapting stories and scenarios to children according to their cultural environment, trendy topics or educational purposes [18]. In this direction, knowing the emotions transmitted by the students and analysing the contents of the messages that produce more positive emotions on each of them would facilitate content adaptation: configurable examples, tests or problem statements can be personalized according to the subjects that produce positive emotions on each student. For instance, examples to show how the "rule of three" works could deal with fruits, number of LPs produced by a well-known artist, or size of a giant chessboard, among others.

In order to incorporate this functionality into CoMoLE, its content adaptation mechanisms need almost no change. The only need is for content versions to include information about "secondary subject" to which the content is related (e.g., fruits, singers, chessboard). Regarding SentBuk, it already gets the most frequently words for each user, which, related to the score (positive/negative) of the sentences in which these words appear, can also give a clue about his interests.

E-learning collaborative systems can also benefit from the information provided by SentBuk. Collaborative activities can be proposed to workgroups according to the group members' sentiment: the time at which the activity is proposed, or the messages to be sent to students at each time (regarding the collaborative task) can vary according to the group emotional state. The information about whether the emotion that a subject provokes in a student is positive or negative can also be useful during dynamic group formation [27]. It would be useful, for example, to minimize the number of groups in which all the students have negative emotions related to the subject.

In order to extract information about the student feeling towards a subject (positive/negative emotion), apart from extracting information from Facebook, if available, we propose to analyze the messages written by the students in the forums associated to the subject, whose length can be similar to those in Facebook. Free-text answers, as well as self-written essays (all of them related to the subject) could be analyzed too. We have some recent experience on extracting specific emotions (joy, anger, fear and sadness) from long essays [28].

VII. Conclusions and Future Work

The work presented in this paper contributes to the development of advanced e-learning systems, since it brings the possibility of adapting different aspects of e-learning processes and environments according to the student emotional state, without asking each of them to fill in questionnaires to capture their state. SentBuk contributes to enrich existing student models by getting information through a sentiment analysis of the messages written by students on Facebook. This possibility is quite attractive because it is not intrusive: it only requires accessing to a Facebook account. Nowadays there are more than 800 million active users on Facebook and more than half of them use it every day.

In the sentiment analysis presented in this paper, the accuracy of the classifier when working with the second set (messages classified as being the most positive and those being the most negative) was surprisingly high, with a 96.78% and 94.59% of success, respectively. As a comparison, the success when classifying messages from Twitter as positive, neutral or negative in [21] was 83.0%.

We analysed the messages processed and found out that the most frequent words in the case of positive emotions were related to greetings and, therefore, quite easy to classify. Words like 'feliz' (happy), 'abrazo' (hug), 'felicidades' (congratulations) or 'beso' (kiss) were among the most frequent words in the set analysed. It is a fact that if one is not in a good mood he will probably neither greet others nor send kisses. However, it could be the case, for instance, that one greets other because it is the other's birthday and he feels that he has to do it, even when his feeling are negative.

With these worries in mind, we have been recently working on the separation of messages according to their type. One can transmit emotions when writing in others' wall, and, in the case of Facebook, we have the feeling that "status" messages might reflect more precisely how a person feels, since, normally, users describe their state in a quite spontaneous way, with no obligation at all. Currently we are replicating the analysis described in this paper considering only status messages.

At present, we are tuning up the classifier according to the users' feedback and increasing the sentiment lexicon. As future word, we plan to specify a range of values to identify not only positive/negative emotions, but also how good or how bad the emotion is. Some studies have built the sentiment lexicon through 'seed' words for which the polarity is already known, and their related words and synonyms are extracted from lexical databases such as WordNet [29]. Other works have taken into account intensifiers to increase (e.g. very) or to decrease (e.g. slightly) the polarity [19].

Regarding learning systems, our current goal is to look for and analyse messages written by students in specific Facebook pages or forums related to courses, to extract information about how a course is going on. The sentiment analysis as implemented in SentBuk makes it possible to know the polarity of the emotions transmitted by the students enrolled in a course, which can give feedback about the course progress: if too many negative messages are delivered in a forum, then something may be happening in this course, and the teacher should know it.

Moreover, we are currently designing a way to find patterns related to emotional changes. Given a "usual" emotional state for the students of a course (as scored and represented by SentBuk starting from the messages posted by the students enrolled in that course), when detecting, e.g., a more negative sentiment than the usual state, the teacher can be informed, so that he can take the appropriate actions to correct these situations that may affect student learning, if any.

Furthermore, in the same context, we plan to extract the topics involved in highly positive/negative messages, to give feedback to the teacher about which issues are the main subject of student praise/complaint. The main goals from an educational point of view are, on one hand, to detect potential problems within a course dynamically, so that they can be solved as soon as possible, and on the other hand, to give positive feedback to the teacher, so that he can know which issues work better.

Finally, we would like to incorporate sentiment-based adaptation in some of our already developed courses in CoMoLE, to make this type of adaptation available to our students at EPS-UAM.

ACKNOWLEDGMENT

We are very grateful to all the users that joined SentBuk.

REFERENCES

[1] A. Kobsa, "Generic User Modeling Systems. The Adaptive Web, LNCS, Vol. 4321, pp. 136-154, Springer, 2007.

[2] L. Shen, M. Wang, and R. Shen, "Affective e-Learning: Using 'Emotional' Data to Improve Learning in Pervasive Learning Environment". Educational Technology&Society, 12(2), 176–189. 2009.

[3] P. Brusilovsky, "Methods and Techniques of Adaptive Hypermedia" User Modeling and User-Adapted Interaction, 6, 2/3, pp. 87-129, 1996.

[4] P. Brusilovsky, E.W. Schwarz, and G. Weber, "ELM-ART: An Intelligent Tutoring System on World Wide Web," Intelligent Tutoring Systems, pp. 261-269, Springer, 1996.

[5] P. De Bra, A. Aerts, B. Berden, B. de Lange, B. Rousseau, T. Santic, D. Smits, and N. Stash, "AHA! The Adaptive Hypermedia Architecture," Proc. 14th ACM Conf. Hypertext and Hypermedia, pp. 81-84, 2003.

[6] R.M. Carro, E. Pulido, and P. Rodríguez, "Dynamic Generation of Adaptive Internet Based Courses," Journal of Network and Computer Applications, vol. 22, pp. 249-257, 1999.

[7] A. Moore, T.J. Brailsford, and C.D. Stewart, "Personally Tailored Teaching in WHURLE Using Conditional Transclusion," Proc. 12th ACM Conf. Hypertext and Hypermedia, pp. 163-164, 2001.

[8] E. Martín, R.M. Carro: Supporting the Development of Mobile Adaptive Learning Environments: A case study. IEEE Transactions on Learning Technologies, vol.2, no.1, pp. 23-36, Jan-March 2009.

[9] R.M. Felder, "Matters of Style," ASEE Prism, vol. 6, 4, pp. 18-23, 1996.

[10] P.T. Costa and R.R. McCrae, "The NEO-PI/NEO-FFI Manual Supplement," Odessa, FL Psychological Assessment Resources, 1989.

[11] L.L. Thurstone, Primary Mental Abilities. Chicago Univ. Press, 1938.

[12] D. Spada, M. Sanchez-Montanes, P. Paredes, and R.M. Carro, Towards Inferring Sequential-Global Dimension of Learning Styles from Mouse Movement Patterns, vol. 5149, pp. 337-340. Springer, 2008.

[13] A. Ortigosa, Quiroga, J.I. and Carro, R.M., "Inferring user personality in social networks: A case study in Facebook. " Proc of the 11th International Conference ISDA 2011, pp. 563 - 568 , Cordoba (Spain), nov 2011. ISSN: 2164-7143. doi: 10.1109/ISDA.2011.6121715. 2011.

[14] B. Pang and L. Lee. "Opinion Mining and Sentiment Analysis" Foundation and Trends in Information Retrieval 2 (1-2), pp. 1-135, 2008

[15] V. Hatzivassiloglou and K.R. McKeown. "Predicting the semantic orientation of adjetives". Procs. of the 35th Annual Metting of the ACL, and 8th Conf. of the European Chapter of the ACL (pp. 174-181). 1997

[16] P.D. Turney and M.L. Littman. "Unsupervised learning of semantic orientation from a hundred-billion-word corpus". Technical Report EGB-1094. National Research Council Canada. 2002.

[17] P.D. Turney. "Thumbs Up dor Thumbs Down? Semantic Orientaton applied to unsupervised classification of reviews". Procs. of the 40th Annual Meeting of the ACL, Philadelphia, pp. 417-424, 2002.

[18] B. Pang, L. Lee and V. Shivakumar. "Thumbs up? Sentiment Classification using Machine Learning Techniques". Procs. of the Conf. on Empirical Methods in NLP, pp. 79-86, Philadelphia, PA. 2002.

[19] M. Taboada, J. Brooke, M. Tofiloski, K. Voll and M. Stede. "Lexicon-Based Methods for Sentiment Analysis". Computational Linguistics 37 (2): 267-307. 2011.

[20] E Martínez Cámara, J.M. Perea, M.T. Valdivia and L.A. Ureña. "Técnicas de clasificación de opiniones aplicadas a un corpus en español". Procesamiento del Lenguaje Natural (47), 163-170. 2011

[21] A. Go, R. Bhayani and L. Huang. "Twitter Sentiment Classification using Distant Supervision". Technical report, Stanford Digital Library Technologies Project. 2009

[22] A. Pak. and P. Paroubek. "Twitter as a corpus for sentiment analysis and opinion mining". Procs. of LREC. 2010.

[23] N. Ramirez-Esparza, J. Pennebaker, F. Garcia and R. Suria. "La psicología del uso de las palabras: Un programa de computadora que analiza textos en Español (The psychology of word use: A computer program that analyzes texts in Spanish)". Revista Mexicana de Psicología (24), 85-99. 2007

[24] M. A. Martí and M. Taulé, "CESS-ECE: corpus anotados del español y catalán", Arena Romanistica. A new Nordic Journal of Romance Studies, num. 1. Monographic about Corpus and Text Linguistics in Romance Languages. 2007.

[25] Bird, S., Klein, E. and Loper, E. "Natural Language Processing with Python". O'Reilly Media. 2009.

[26] Vilares Ferro, J. "Aplicaciones del procesamiento del lenguaje natural en la recuperación de información en español". Thesis. U. da Coruña. 2005.

[27] P. Paredes, A. Ortigosa and P. Rodriguez, "A Method for Supporting Heterogeneous-Group Formation through Heuristics and Visualization", JUCS, vol 16, n 19, pp. 2882-2901, 2010.

[28] P. Rodriguez, A. Ortigosa and Rosa M. Carro, "Extracting Emotions from Texts in E-learning Environments". 2nd Int. Workshop on Adaptive Learning via Interactive, Collaborative and Emotional approaches. Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems (CISIS'2012). IEEE Computer Society, pp. 887-892, 2012.

[29] A. Esuli and F. Sebastiani, "SentiWordNet:A publicly available lexical resource for opinion mining". Procs. of Language Resources and Evaluation (LREC), 2006.