Sentiment analysis in Facebook and its application to e-learning

Alvaro Ortigosa *, José M. Martín, Rosa M. Carro

Department of Computer Science, Universidad Autónoma de Madrid, Francisco Tomás y Valiente 11, 28049 Madrid, Spain

ARTICLE INFO

ABSTRACT

This paper presents a new method for sentiment analysis in Facebook that, starting from messages written by users, supports: (i) to extract information about the users' sentiment polarity (positive, neutral or negative), as transmitted in the messages they write; and (ii) to model the users' usual sentiment polarity and to detect significant emotional changes. We have implemented this method in SentBuk, a Facebook application also presented in this paper. SentBuk retrieves messages written by users in Facebook and classifies them according to their polarity, showing the results to the users through an interactive interface. It also supports emotional change detection, friend's emotion finding, user classification according to their messages, and statistics, among others. The classification method implemented in SentBuk follows a hybrid approach: it combines lexical-based and machine-learning techniques. The results obtained through this approach show that it is feasible to perform sentiment analysis in Facebook with high accuracy (83.27%). In the context of e-learning, it is very useful to have information about the users' sentiments available. On one hand, this information can be used by adaptive e-learning systems to support personalized learning, by considering the user's emotional state when recommending him/her the most suitable activities to be tackled at each time. On the other hand, the students' sentiments towards a course can serve as feedback for teachers, especially in the case of online learning, where face-to-face contact is less frequent. The usefulness of this work in the context of e-learning, both for teachers and for adaptive systems, is described too.

© 2013 Published by Elsevier Ltd.

1. Motivation

The use of computers in education has meant a great contribution for students and teachers. The incorporation of adaptation methods and techniques allows the development of adaptive e-learning systems, where each student receives personalized guidance during the learning process (Brusilovsky, 2001). In order to provide personalization, it is necessary to store information about each student in what is called the student model (Kobsa, 2007). The specific information to be collected and stored depends on the goals of the adaptive e-learning system (e.g., preferences, learning styles, personality, emotional state, context, previous actions, and so on).

In particular, affective and emotional factors, among other aspects, seem to affect the student motivation and, in general, the outcome of the learning process (Shen, Wang, & Shen, 2012). Therefore, in learning contexts, being able to detect and manage information about the students' emotions at a certain time can contribute to know their potential needs at that time. On one hand, adaptive e-learning environments can make use of this

information to fulfill those needs at runtime: they can provide the user with recommendations about activities to tackle or contents to interact with, adapted to his/her emotional state at that time. On the other hand, information about the student emotions towards a course can act as feedback for the teacher. This is especially useful for online courses, in which there is little (or none) face-to-face contact between students and teachers and, therefore, there are fewer opportunities for teachers to get feedback from the students.

Knowing the users' emotions is useful not only in the educational context but also in many others (e.g., marketing, politics, online shopping, and so on) (Feldman, 2013). In general, in order for a system to be able to take decisions based on information about the users, it is necessary for it to get and store information about them. One of the most traditional procedures to obtain information about users consists of asking them to fill in questionnaires. However, the users can find this task too time-consuming. Recently, non intrusive techniques are preferred (de Montjoye, Quoidbach, Robic, & Pentland, 2013). We also think that information for student models should be obtained as unobtrusively as possible, yet without compromising the reliability of the model built (Ortigosa, Carro, & Quiroga, 2013).

When reflecting about potential sources of information regarding user sentiment, we looked for digital places in which the users

* Corresponding author. Tel.: +34 91 4972271.
E-mail addresses: alvaro.ortigosa@uam.es (A. Ortigosa), joseph.martin@estudiante.uam.es (J.M. Martín), rosa.carro@uam.es (R.M. Carro).

express themselves frequently and naturally. Nowadays, the number of users interacting with others through social networks is growing exponentially. Therefore, we focused on social networks. There exist an increasing number of online social networks available through the Web. From these applications, Facebook is the more popular around the world. On October 2012, it reached 1 billion monthly active users (that is, 1 billion users accessed the network within a month) and more than 550 million daily active users (Kiss, 2012). Besides its popularity, Facebook provides a distinctive advantage for this research: it is a network of friends. That is, whilst other social networks focus on professional relationships or serve, mostly, as sources of information, people make use of Facebook mainly to share and communicate with friends; in fact, the acquaintances or relationships between users are called "friends". Messages in Facebook are spontaneous and users express their emotions more naturally. It was because of all these reasons that we chose Facebook as the development platform. In Facebook, the "wall" is the space where the users publish their own messages, contents and so on. Regarding text messages, there are several categories: status messages (each user writes them in his/her own wall), posts in others' walls, and comments to either one's or others' publications. Typically, a user's wall is visible to his/her friends, and they can make comments or express that they "like" a particular message or post.

When dealing with users and sentiments, it is useful to know the users' emotional state at a certain time (positive/neutral/negative), in order to provide each of them with personalized assistance accordingly. Moreover, it is also interesting to know whether this state corresponds to their "usual state" or, on the contrary, a noticeable variation might have taken place. Behavior variations, as detected in the messages written by a user (when sentiment histories are available), can indicate changes in the user's mood, and specific actions could be potentially needed or recommended in such cases.

With the purpose of extracting information about users' sentiments from the messages they write in Facebook and detecting changes, we have developed a new and non-intrusive method for sentiment analysis in this social network. It consists on a hybrid approach, combining lexical-based and machine learning techniques. We have implemented this method in Sent-Buk, a Facebook application that retrieves the messages written by the users and extracts information about their emotional state.

This paper is organized as follows. Section 2 presents the state of the art of the research areas related to our work. Section 3 describes the new method proposed for text-based sentiment analysis. Section 4 presents SentBuk, the Facebook application in which that method has been implemented. Section 5 includes the results obtained when making use of SentBuk, as well as the analysis of these results. Section 6 presents a discussion of the proposal and shows some applications of sentiment analysis in the context of e-learning. Finally, the conclusions of the work done, along with some lines for future work, are presented in Section 7.

2. Related work

2.1. Sentiment analysis

Sentiment analysis has been defined as the computational study of opinions, sentiments and emotions expressed in texts (Liu, 2010). For the sake of simplifying the development of an emotion recognition tool, we have tried to avoid complex and potentially controversial definitions of emotions and sentiments. In this direction, we take the simplified definition of sentiment 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"). Most of works in this research area focus on classifying texts according to their sentiment polarity, which can be positive, negative or neutral (Pang & Lee, 2008). Therefore, it can be considered a text classification problem, since its goal consists of categorizing texts within classes by means of algorithmic methods.

The earliest researches dealing with sentiment analysis consisted on classifying words or phrases according to semantic issues and date from the late 1990s (Hatzivassiloglou & McKeown, 1997). Linguistic heuristics or pre-selected sets of seed words were used. The results obtained in those works served as 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 (Turney, 2002). The appearance of WordNet (Miller, 1995) and, in general, of annotated corpora, increased the production in this research area. On one hand, WordNet is useful because it allows knowing the semantic relationships between different words. Therefore, with a reduced set of polarity words, every word could be labeled as positive, negative or neutral through its relationships. On the other hand, corpora and, in particular, the Treebanks, are very useful. They are corpora with the syntactic structure labeled, and are of great help for training the analyzers in order to label the words automatically.

One of the first works that used the term "sentiment analysis" as we currently know it was that presented in (Das & Chen, 2001), which analyzes messages written in stock boards in order to extract the market sentiment. Currently, many of the works in this area focus on document classification based on the sentiment expressed on it. One of the best known domains is that of reviews (Pang, Lee, & Vaithyanathan, 2002) (Dave, Lawrence, & Pennock, 2003). Review websites are examples of especially useful sources for sentiment analysis, such as, e.g. Epinions (Epinions, 1999). Other application areas in which sentiment analysis can be very useful are:

- Recommendation systems (Tatemura, 2000).
- Flame detection (Spertus, 1997).
- Sensitive content detection for advertising (Jin, Li, Mah, & Tong, 2007).
- Human–computer interaction (Liu, Lieberman, & Selker, 2003).
- Business Intelligence (Mishne & Glance, 2006)
- Prediction of hostile or negative sources (Abbasi, 2007).
- Classification of citizens' opinions on a law before its approval: "eRuleMaking" (Cardie, Farina, Bruce, & Wagner, 2006).
- Broadcasting based on the receiver sentiment (Rogers, 2003).
- Dynamic adaptation of daily tools, such as e-mail (Carro, Ballesteros, Ortigosa, Guardiola, & Soriano, 2012).
- Marketing or politics (Feldman, 2013).

In general, accuracy is strongly influenced by the context in which the words are used (Turney, 2002) (Aue & Gamon, 2005) (Engström, 2004). For instance, the sentence "You must read the book" is positive in a book review but is negative if the review is about films.

Additionally, the position of words in text is an interesting factor to consider, since a word at the end of a sentence can change the polarity completely (Pang et al., 2002). For example the sentence "This book is very addictive, it can be read in one sitting, but I have to admit that it is rubbish" begins with the word "addictive" and the expression "one sitting", which are positive in the context of book reviews, but it finish with the word "rubbish" that

is negative. Although the sentence contains two positive tokens against one negative, it should be marked as negative because the final word nullifies all the previous ones.

Another issue to be considered is the presence of figures of speech in the analyzed text. Some of them, such as the irony, can change the whole polarity of a text. Sometimes they are difficult to detect even for a human being, if additional information is not provided (e.g. context). Recent work in natural language processing focuses on the detection of these figures, such as (Reyes, Rosso, & Buscaldi, 2012), that build a training dataset of messages written in Twitter with the hashtag '#irony' in order to set a model with machine-learning techniques.

With respect to the techniques used for sentiment analysis, two main approaches are considered: machine-learning methods and lexicon-based approach. The survey written by Pang and Lee (Pang & Lee, 2008) covers the most popular techniques and approaches.

On one hand, machine-learning methods are used to classify texts. An example of the use of machine-learning techniques in order to classify movie reviews is presented in (Pang et al., 2002). It compares different techniques to classify movie reviews, obtaining 82.9% of accuracy when applying Support Vector Machines (SVM). Generally, it is difficult to obtain better results, due to characteristics of natural language. However, in specific domains, the use of machine learning algorithms for classifying texts according to their sentiment orientation performs well.

On the other hand, the lexicon-based approach consists of analyzing the text grammar and executing a function to give a sentiment score to the text, considering a predefined sentiment lexicon (Turney, 2002) (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). There exist some sentiment lexicon available, such as SentiWordNet (Esuli & Sebastiani, 2006), but it has been noticed that most of the researches build their own lexicon ad hoc, managing the semantic relationships between words with tools such as WordNet (Miller, 1995), already mentioned above.

The great advantage of the lexicon-based approach is that it is not necessary to have a labeled training set to start classifying texts. This approach tends to get worse results than machine learning approaches in specific domains, but when the domain is less bounded the results are better. This is because the lexicon approach analyzes the text grammar, whereas the machine-learning methods fit the algorithms to the training dataset particularities. As an example, in (Taboada et al., 2011) the authors use a lexicon-based method with six different corpora from different domains and obtained 75–80% accuracy. However, when using machine learning (with a preprocessing phase to summarize movie reviews), 86.4% accuracy was achieved (Pang & Lee, 2004). When comparing these two works, the machine-learning approach gets a better accuracy, but it may suffer overfitting to the training dataset, whereas the lexicon-based approach gets a lower accuracy, although is more robust when considering different domains.

In relation to language analysis, there are very few works dealing with texts in languages different from English. The works found are usually adaptations of already presented methods for English sentiment analysis. For example, in Martínez Cámara, Perea, Valdivia, and Ureña (2011), different machine learning methods are applied in order to classify movie reviews, achieving an interesting 86.84% success with SVM in Spanish language.

Finally, in recent years, due to the increasing amount of information delivered through social networks, many researches are focusing on applying sentiment analysis to these data (Go, Bhayani, & Huang, 2009) (Pak & Paroubek, 2010). However, most all these works deal with English texts and retrieve them from Twitter, since it is easier to retrieve data from this social network than from others such as Facebook.

2.2. Adaptive hypermedia and user modeling

It is a fact that not all people have the same characteristics, needs, preferences and goals. This causes a different experience when a user is interacting with a computer system. Therefore, this raises the need of considering the different user characteristics in order to adapt the system according to the user needs and other relevant aspects. This is the aim of the Adaptive Hypermedia (AH). The work (Brusilovsky, Adaptive hypermedia, 2001) presented the best known classification of AH methods and techniques, considering both adaptive presentation (content-level adaptation) and adaptive navigation support (link-level adaptation).

In particular, AH has been used in the context of e-learning to support personalized learning, by recommending the most suitable task to be accomplished by each student 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 know adaptive e-learning systems are ELM-ART (Brusilovsky, Schwarz, & Weber, 1996), AHA! (De Bra et al., 2003), TANGOW (Carro, Pulido, & Rodríguez, 1999), WHURLE (Moore, Brailsford, & Stewart, 2001) and CoMoLE (Martín & Carro, 2009).

CoMoLE (Context-based adaptive Mobile Learning Environments) is one of the adaptive e-learning systems that can take advantage of knowing the user sentiment, as it will be shown in section 6.1. CoMoLE supports the design and dynamic generation of e-learning environments able to:

– Recommend individual or collaborative activities to be carried out by each student or group according to each one's needs, previous actions, preferences or context at each time (location, available time and device).
– Generate individual or collaborative workspaces to support the realization of each activity. It generates these workspaces on the fly, according to the activity to be carried out, as well as to the user or group features. It selects the most suitable content versions and individual/collaborative tools for each case and generates the corresponding workspace.

Any adaptive system needs information about the user in order to provide adaptation. All the information about a user that will be considered with adaptation purposes must be stored in the user model and maintained up to date (Kobsa, 2007). These data can be obtained from specific tests, such as those related to learning styles (Felder, 1996), personality (Costa & McCrae, 1989) or intelligence (Thurstone, 1938). More recently, some of these data have been inferred from the user interactions with the system in order to avoid asking the user. For example, (Spada, Sanchez-Montanes, Paredes, & -@@- Carro, 2008) shows that it is possible to predict some aspects of the user's learning style by analyzing the user mouse movements. Personality can also be inferred unobtrusively, in this case by analyzing the user interactions in Facebook (Ortigosa et al., 2013).

Recent researches in the area of user modeling are trying to take advantage of the great amount of user data stored and accessed through social networks. For example, in the work presented in (Abel, Gao, Houben, & Tao, 2011), the user model contains data available on Twitter. This information is used to give news recommendations to the corresponding user. Other works have incorporated information about the user's sentiment or emotional status to the user model. In (Nasoz & Lisetti, 2007), the authors present a method to take actions in order to improve driving according to the driver's emotional status. Another interesting related work is presented in (Conati & Maclaren, 2009). In this case, adaptation is provided within an educational game, taking into consideration the emotions that arise in an evaluation situation.

In general, any system able to provide adaptation, recommendations or personalized assistance, which makes use of a user model, can benefit of the results of the work presented here, focused on extracting user emotional states and detecting sentiment changes that will be incorporated within the corresponding user models.

3. Sentiment extraction and change detection

The work presented in this article covers, on the one hand, the extraction of information about users' positive/neutral/negative sentiments from the messages they write (the way in which we obtain the messages and deal with privacy issues will be described in Section 4.1). On the other hand, we investigate the detection of sentiment changes with respect to the "usual" sentiment of each user. We present the methods developed with each purpose below.

3.1. Extracting sentiments from texts: message classification

The method developed to extract the user sentiment from texts materializes in a sentiment classifier of messages written in Spanish. Two different approaches, at least, can be used for classifying text messages according to the sentiments they show. The first one consists on a lexicon-based approach that relies on spotting words with sentiment load. The second one, the machine learning approach, bases on pattern recognition. Both approaches have advantages and shortcomings:

– Machine learning approaches obtain better results in bounded domains. However, in order to be implemented they require a training set of messages already labeled (the larger the better), so that patterns distinguishing positive, neutral and negative messages can be learnt.
– Lexicon-based approaches require a dictionary of words, each of them annotated with its semantic orientation (positive/negative emotional polarity). However, lexicons are more easily available and extensible than training sets. One of their drawbacks is that this type of approaches usually performs worse (less accurate) than machine-learning ones on bounded domains, yet they are more robust when considering cross-domain applications.

Considering the strengths and limitations of each method, we decided to design and implement a combined approach. At a first stage, a lexicon-based sentiment classifier was built and put at work, as explained in this section. Once a large number of labeled messages were available, we used them as the training set for a machine-learning based classifier, giving rise to a new hybrid classification approach, as we will explain in detail in Section 5.

Our first classifier follows a lexicon-based approach. It uses a dictionary of words, each of them annotated with its semantic orientation (positive/negative emotional polarity). It also detects other elements such as positive/negative interjections, emoticons, misspellings, part of speech tagging or negation (polarity shifter). The procedure to classify each message, step by step, is shown in Fig. 1 and is described next.

3.1.1. Preprocessing (lower-case, idiom detection)

The first step consists of preprocessing the message to convert all the words into lower-case. Afterwards, it detects idioms and joins the words involved in each of them. For example, "de mala fe" (in bad faith) is converted into "de_mala_fe", to be processed as a unique work afterwards.

3.1.2. Segmentation into sentences

Then, the message is divided into sentences. Dots are the only punctuation marks considered as separators at this step, since others such as commas or semicolons can be part of emoticons.

3.1.3. Tokenization I (partial)

In the next step, tokens are extracted from each sentence. At this time, only whitespaces are taken into consideration to separate tokens, since other separators, such as hyphens, can be part of emoticons.

3.1.4. Emoticon detection

Next, emoticons are detected. In order to detect them, the classifier searches in the text all the emoticons stored in two text files, containing positive and negative emoticons, respectively. The lists of emoticons included in these files contain those extracted from Wikipedia (Wikipedia, 2013) along with some more added when preparing this research.

3.1.5. Tokenization II (complete)

During this second tokenization phase, all the punctuation marks (including commas, semicolons and so on) are considered as separators leading to obtain the final sets of tokens for each sentence.

3.1.6. Interjection detection

The next step consists of detecting and labeling interjections. Those that express laughs, such as "jejeje" or "jajaja", are marked as positive whereas interjections such as "joopeee" (upset) or "uff" (tiredness) are marked as negative. We have implemented this detection through regular expressions, because we found out that, in most of the cases, the interjections are intensified by repeating letters or sets of letters contained in the own word. For example, "jope" represents upset whereas "jjjooooooooopeeeeee" suggests a stronger upset sentiment. Were regular expressions not used at this step, many interjections would not be detected (e.g., "jajajaja", which matches the regular expression "ja+", is detected, although it was not in the dictionary).

3.1.7. Token score assignation

The next phase consists of assigning a score to each token: 1 if it transmits a positive sentiment, 0 if it is neutral, and -1 if it is negative. To assign a score, the classifier checks if the token is a positive/negative emoticon, a positive/negative interjection, or whether it matches one of the words stored in the sentiment lexicon (L).

3.1.7.1. Building the lexicon.
The sentiment lexicon was previously built starting from the Spanish Linguistic Inquiry and Word Count (LIWC), (Ramirez-Esparza, Pennebaker, Garcia, & Suria, 2007), which considers the following categories, allowing the classification of words into one of these two classes:

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

During the preparation of this research, we found some words that clearly had a strong sentiment associated and were not in the dictionary. Therefore, we extended the dictionary by adding several words to each class (most of which were slang).

In Spanish, words are formed by adding prefixes or suffixes to a lexeme. In order to avoid a too extensive list of words, in many cases the suffix next to the lexeme was not included. Instead, the character '∗' was put next to the lexeme. For example, "bueno" ("good" in English) appears as "buen∗". This token represents the
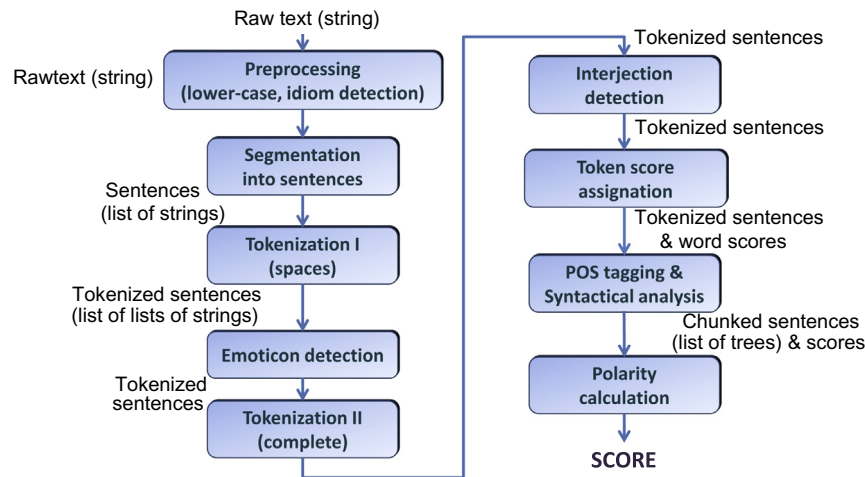
Fig. 1. Diagram representing the procedure for emotion detection.

words "bueno", "Buena" (female), "buenísimo" (superlative, male) and "buenísima" (superlative, female), among others.

Finally, after all these issues were taken into account, the positive class contains 653 tokens and the negative one contains 894 tokens. The two classes that compose the lexicon are stored in two text files, one each. The lexicon can be easily adjusted just by modifying these files.

3.1.7.2. Removing repetitive letters. The next step is, for each token, to check whether it appears in any of the two dictionary categories and, if it is the case, to tag it as either positive or negative, accordingly.

The messages written by users in Facebook usually contain very casual language. It is frequent to find words with repeated letters (e.g. "fenomenaaaaaaal" – "great" in English) or with non-alphabetic characters. Consequently, for each token, it if was not found in the dictionary in the form in which it appears in the message, letters occurring more than twice in a row are replaced by only one occurrence and the new token is looked for in the dictionary. Were this version of the token not found in the dictionary yet, then it is reduced to its lexeme.

3.1.7.3. Spelling checking. If the token does not match any word yet, then it is checked with a spelling checker. Since Facebook messages usually contain misspellings, we incorporated a spelling checker into the classifier: *GNU Aspell*, included with most Linux distributions. However, the spelling checker must be applied carefully, since some of the corrections it suggests produce bad results in the classifier. For example, for the word "Dani" (a popular Spanish nickname) the suggested correction was "daño" (damage, a negative word). With the purpose of avoiding these situations, a list of words that should not be corrected, including names and surnames, was created and incorporated within the dictionary, so that, since they are found in the dictionary, they are not checked by the spelling checker. Finally, if a token is not classified into positive/negative in any of the previous steps (even after all those considerations) the token is labeled as neutral.

3.1.8. Syntactical analysis

Once each token has received a positive/neutral/negative score, each sentence is syntactically analyzed in order to check whether any score (positive/negative) should be reversed (e.g., because of negations). Firstly, we apply part of speech (POS) tagging to discriminate words that do not reflect any sentiment (e.g. articles)

and to disambiguate words with multiple semantic meaning (e.g., words that can be both a noun and a verb).

In order to tag the messages retrieved from Facebook, a tagger was automatically built up using a training corpus, the CESS-ESP (Martí, Taulé Delor, Màrquez, & Bertran, 2007) distributed with the Natural Language Toolkit (NLTK) (Bird, Klein, & Loper, 2009). This corpus is a syntactic Treebank with constituents and functions. Tag generation was improved with *backoff* techniques. Through *backoff*, bigrams firstly and unigrams secondly are tagged. After considering bigrams and unigrams, the tokens untagged are tagged as nouns. The resulting tagger has an 81.91% accuracy when tested, providing a light improvement over the unigram tagger (80.65%). The accuracy of a bigram tagger without *backoff* is poor (10.90%).

Afterwards, suffixes are analyzed, since, in Spanish, certain suffixes denote the grammatical category of a word. For example, a word ending with 'aríamos' is surely a verb and a word ending with 'ísimo' is an adjective. The list of suffixes used in our approach is based on those lists compiled in (Vilares Ferro, 2005).

Finally, negations are detected. In Spanish, negation is generally expressed by placing a negative adverb. A negation implies the need of reversing the polarity of the words affected by that adverb. For example, "esto no me gusta" ("I do not like this") would score positive if negation were not detected, since "gusta" ("like") is a positive token. However, it should score negative, because of the negation. Therefore, the classifier looks for negative adverbs detected by the POS tagger. With the purpose of detecting which words are affected by a negation adverb, a chunking (shallow parsing) has been incorporated into the classifier. The grammar used for chunking is defined by regular expressions and is based on the grammar published in (Vilares Ferro, 2005). The inversion of polarity is applied to the words under the syntactic constituent that contains the negation adverb (Taboada et al., 2011). The chunking parser has also been developed with NLTK.

3.1.9. Polarity calculation

In order to calculate the polarity of a sentence, the number of tokens susceptible of conveying sentiments according to their grammatical category (i.e. noun, adjective, interjection or verb) is calculated. Other types of words, which appear frequently in texts (i.e. determinant, prepositions, etc.), are not considered, because they are "stopwords" for a sentiment analysis. Then, once each token has been scored, the final polarity score of a sentence is calculated as the sum of the scores divided by the sum of all the

candidates to receive a score. The score obtained is a real number from −1 to +1.

When carrying out the first proofs of concept, we noticed that the messages were classified as neutral when the numbers of positive and negative tokens were equal (we call it "sentiment neutralization"). According to (Pang et al., 2002), the words at the end of a sentence are more relevant, and the position of terms in a text can influence the overall sentiment. Therefore, following this observation, we decided for the classifier to give more importance to the tokens placed at the end of the sentence, in case of sentiment neutralization (it should be noticed that sentiment neutralization is different from a neutral sentence, in which no sentiment is detected at all).

3.2. Sentiment change detection

As it was mentioned in the motivation of this work, our goal is not only to know the sentiment of a user at a certain time, but also to capture his/her habits and usual sentiment state (the user "regular pattern"), so that those sentiments out of his/her usual state can be detected. The final goal would be to be able to help the user in the case it is needed, or to alert to whom it corresponds about someone's sentiment changes. With this purpose, we propose a method to model the user interactions and sentiment patterns.

3.2.1. Building the user regular pattern

The first decision to be taken was related to which information about the user should be collected to build the "regular pattern". Apart from considering the messages written by the user, it would be interesting to collect other data related to the users' actions that could give clues about potential emotional changes. Since the goal was to detect changes on user sentiments in social networks such as Facebook, we focused on data susceptible of changing dynamically. In this direction, we decided to consider the following information parameters:

- Mean of the sentiment showed by the messages published by the user (s).
- Number of messages written (m).
- Number of comments to messages made (c).
- Number of 'Likes' made to messages on his/her wall (l).
- Number of 'Likes' made to comments to messages on his/her wall (k).

Another critical issue to be decided is related to the minimum amount of time necessary to track a user activity, in order to be able to establish "states" for comparisons. In order words, it was necessary to decide the interval of time to which the previous parameters refer. On one hand, it would be good to be able to detect any emotional change as soon as possible. From this point of view, the shortest the reference period is the better. On the other hand, too short intervals will lead to false positives: if the analysis period is, e.g., 3 or 4 h, then the values of the parameters representing the user emotional state when he/she is interacting with the computer would be very different from those when the user is sleeping (please note that the analysis can be done at any time, since there is no online requirement to do so, and, in this example, an extremely low activity or even no activity at all – e.g., at night – would be detected as a behavioral change with respect to the rest of the day and, therefore, as an indicator of a potential emotional change). A similar reasoning can be applied for time intervals of 1 day, or even 3–4 days: during the weekends, the patterns of interaction with the social network may differ greatly from those on working days. For these reasons, after exploring different possibilities, we found that a week (seven days) would be a reasonable

time interval were a certain routine can be detected on most of the users.

3.2.2. Comparing weeks

In order to compare the user's patterns of interactions in different weeks, we use vector comparisons. Each vector, which we call profile (P), contains the data collected for each user (u) along a week (w). The data contained in each profile are those enumerated above (see Eq. (1)).

$$P(u, w) = (m, c, l, k, s) \tag{1}$$

Equation 1: Weekly user vector.

Through a set of weekly profiles, changes in the user's behavior could be detected. For example, if a user usually writes two or three messages per week and one week writes twenty messages with a lot of comments, then this may be a sign that something different is happening to him/her. Similarly, if a user interacts with Facebook daily, writing a lot of messages, commenting on other's wall and so on, and shows an extremely low activity (or even does not connect) for a while, this "silence" can represent a potential emotional change.

Emotional changes will be detected by comparing vectors. In order to set the user's regular pattern, the median of each attribute, considering a set of weekly vectors, is calculated. Prior to the median calculation, data is normalized. Values outside the range [0.1, 0.9] in the distribution data will be not considered (i.e., outliers' removal). The maximum and minimum values for each attribute are stored, since they will be needed for future calculations.

In order to create the "regular pattern" for a user for the first time, all the weekly profiles along the last year (or the time the user has been using the system, if less than one year) are built. Data from the last week are not considered for "regular pattern" creation, since, precisely, these are the ones to be compared with the "regular pattern" to look for changes. Information from one complete year (52 weeks) is more than enough to build a significant regular pattern. Within this period we can find some events that happen yearly, such as birthdays, Christmas, summer holidays, anniversaries, etc., which can provoke different emotions on users. Therefore, by extracting the sentiments transmitted by the user in all his/her messages since he/she joined Facebook (with a maximum of one year), we build the user sentiment tendency pattern. Once this regular pattern has been built (P), it is compared with last week profile (Q). During the next week, the regular pattern (P) is augmented with information from the previous week (Q), obtaining a new vector P to be compared with the one associated to the last week. The same procedure is followed for subsequent weeks.

The easiest way to compare vectors is the Euclidean Distance (Eq. (2)):

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \tag{2}$$

Equation 2: Euclidean Distance between vectors P and Q.

On one hand, before calculating the distance between both vectors, the weekly vector is normalized with respect to the minimum and maximum values collected for each attribute during the pattern construction. On the other hand, distance between two vectors is not bounded. For this reason, in order to get values for similarity within the interval [0,1], the equation shown in (Equation (3)) is applied. As it can be seen, the similarity between two vectors will be lower when the distance between them is higher and the other way round. Two vectors are equal (similarity = 1) when the distance between them is zero.
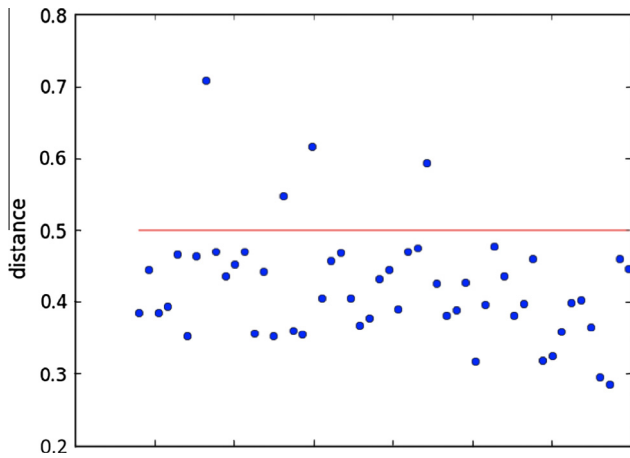
Fig. 2. Distances to weekly profiles.

$$Similarity = \frac{1}{(1 + d_E(P, Q))} \quad (3)$$

Equation 3: Distance between weekly profiles.

In Fig. 2, the X axis represents the fifty-two weeks used to build the user's regular pattern, when messages are available since then. The Y axis represents the distance between a user's sentiment and his/her regular pattern. Therefore, each dot represents, for each week, how far is the user's sentiment during this week from his/her regular pattern. In this chart it is possible to see that, in most of the cases, the distance ranges between 0.3 and 0.5, which means that the user's sentiment does not vary significantly along the time. The weeks showing the largest distances (over 0.5) are those to be considered when wishing to carry out any type of intervention based on user significant sentiment changes.

4. SentBuk: an application to support sentiment analysis in Facebook

In order to implement and test the methods described in the previous section, a Facebook application called Sentbuk was developed. The goals of SentBuk are to access the text messages written by Facebook users, to collect all the data required, to use this information for sentiment analysis and to provide information regarding the analysis done. In this context, another goal was to collect information from as many users as possible, which allows us to check the suitability of the proposal for detecting emotional states and their changes.

4.1. Obtaining data

In order to collect the user data, any Facebook application has to ask for the user's permission, following a strict policy oriented towards protecting the user's privacy. The user has to give different permissions to access to each specific piece of information, or to allow the application to use particular methods to access to this information. Therefore, the data that can be accessed by a given application depends on the permissions that the user granted to that application. For this reason, the first step once any user accesses Sentbuk for the first time is to ask him/her for the following permissions:

- *Offline_access*: Permission to access the user data even if he/she has no opened session in Facebook at that time.
- *Read_stream*: Permission to read the user wall.
- *User_about_me*: Permission to read basic user data (gender, birthday, languages, etc.).

Requesting other type of data, such as the user religious beliefs, for example, was considered risky, since the users tend to distrust on applications that request this kind of permissions, and the contribution of this information to sentiment analysis did not seem relevant enough to jeopardize the user acceptance of SentBuk.

After obtaining the corresponding permissions, SentBuk requests the user data through the Facebook API (a HTTP Web service). It is worth to remark that Facebook puts limitations to the API calls, as expressed in the API documentation. The service returns at most 50 messages when trying to get the whole message history, and messages more than 30-day old cannot be retrieved either. This limitation has been overcome by using parameters to specify searches by dates. That is, instead of trying to recover the whole message history, messages are recovered by blocks.

The Facebook API returns JSON (*JavaScript Object Notation*) streams, so they can be processed easily with an appropriate parser. Some of the data that can be obtained from a stream, corresponding to a message, are: the message author, the receiver, the creation date, the number of likes, the text itself and the comments associated to this message.

Lastly, since the retrieved data are key-value pairs, the records are stored in a NOSQL database (that is, a database able to store the information retrieved without further transformation or mapping). Having the messages stored in a database avoids querying the Facebook API every time, which is very slow process and, as explained before, suffer from query limitations.

4.2. SentBuk interface

SentBuk performs data analysis following the method explained in Section 3.2. When a user launches SentBuk, the results of sentiment analysis are shown graphically (see Fig. 3). At the top of the interface (see A in the figure), the user has the possibility to look for his/her own messages, to see his/her regular profile or to watch the evolution of his/her sentiment along the time. He/she can see the same information from his/her friends and also from public user profiles (see tabs). The user can also select a period of time to check his/her sentiment polarity during that period, as well as mark "only status messages" to focus on this type of messages.

In the same screen, a summary of his/her most active friends as well as that of the ones who write more positive/neutral/negative messages, is graphically shown (see B). Next, charts showing the user sentiment polarity (C), the percentage of positive/neutral/negative messages written by him/her (D) and a tag cloud showing his/her most frequent words is shown (E).

Below the previous elements, SentBuk presents a list showing the messages from the user wall along with the corresponding polarity scores, as calculated by the sentiment classifier. Messages classified as positive are highlighted in green; red is used for negative messages and yellow for neutral messages. Fig. 4 shows the annotated messages for a quite emotional user (her name and connection data have been removed for privacy issues). By clicking on the information icon showed next to each message, the user can see information about the procedure that leads to assigning this message the score shown. The user can give feedback to SentBuk in case that he/she does not agree with the score assigned to any message, by clicking the corresponding thumb-down icon.

When a user clicks on the button to access to the evolution of his/her sentiment along the time, SentBuk shows, graphically, the similarity between the user's weekly profile and the regular one along the time. It is shown graphically, so that the evolution of the regular profile along the time, as well as the peaks of positive/negative sentiments with respect to that profile, can be quickly seen. In the last version of the tool, we incorporated a warning system in SentBuk, so that a user is notified about
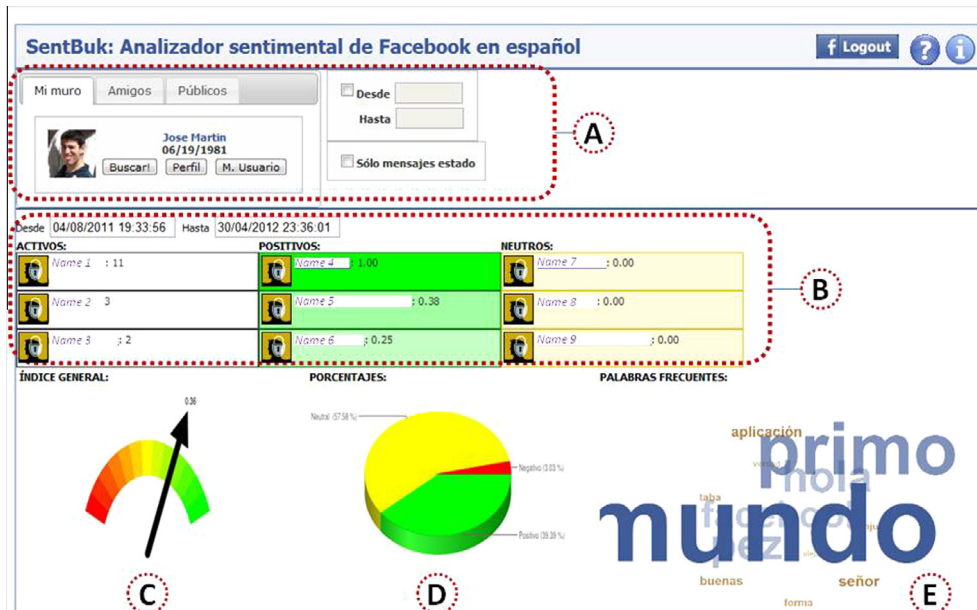
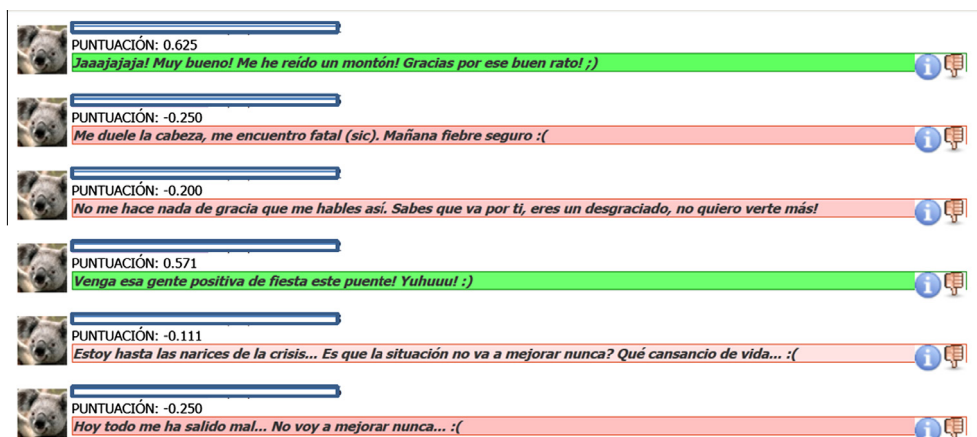**Fig. 3.** Snapshot of SentBuk main interface.

Fig. 4. Snapshot of annotated messages for a given user.

significant changes in a friend's sentiment (see Fig. 5). Warnings are showed when the similarity between a user weekly profile and his/her regular one is lower than 0.5 (a preset value). In the case shown in Fig. 5, a positive change was detected. It turned out that the user's birthday had been in that week and it seems that he expressed his positive sentiments towards this fact through Facebook indeed.

5. Data analysis and results

In order to evaluate the accuracy of the message-based sentiment classifier developed through our lexicon approach, the decisions taken by that classifier were compared with those of a human judge. This analysis was carried out in two phases: firstly, all the messages and comments contained in the users' walls were classified; in a second stage, only the status messages (and their corresponding comments) where classified. The results of both approaches were compared. Sections 5.1 and 5.2 explain each of these analyses, respectively.

Afterwards, the use of machine-learning techniques was explored. Moreover, lexicon-based and machine-learning techniques

were combined. The results of the lexicon-based approach, those when using machine-learning, and the ones obtained when combining both techniques, are contrasted and described in Section 5.3.

5.1. Evaluating our lexicon-based approach with all the messages

In this approach, firstly, SentBuk classified all the messages retrieved from Facebook. Afterwards, 1000 messages of each class (positive, neutral or negative) were randomly selected. They were delivered altogether to a human, who manually classified them, without knowing the result obtained using the classifier. We compared the two classifications to get information about the accuracy of the lexicon-based approach. When considering all the messages together, this approach gets an accuracy of 86.28%. When separating messages according to the class predicted by SentBuk, it gets an accuracy of 95.63% for the positive class, 83.51% for the neutral, and 79.33% for the negative one. It should be borne in mind that messages whose score is near zero (from -1 to $+1$) do not express a strong sentiment. Therefore, if we focus only on the messages expressing strong sentiments, and consider those 500 messages

Fig. 5. Warning of possible emotional change.

Table 1
Confusion matrix for the complete message set.

Predicted class	Actual class				
	Positive	Neutral	Negative	Another language	Accuracy (Spanish) (%)
Positive	920	23	19	38	95.63
Neutral	74	704	65	157	83.51
Negative	89	109	760	42	79.33

Table 2
Confusion matrix considering only strong sentiment messages.

Predicted class	Actual class				
	Positive	Neutral	Negative	Another language	Accuracy (Spanish) (%)
Positive	481	13	3	3	96.78
Negative	6	20	455	19	94.59

Table 3
Total number of messages per class.

	Total	%
Positive	94,019	66.89
Neutral	35,432	25.21
Negative	11,117	7.90
Total	140,568	100.00

Table 4
Total number of comments per class.

	Total	%
Positive	103,811	51.50
Neutral	73,355	36.40
Negative	24,402	12.10
Total	201,568	100.00

Table 5
Percentages of messages per class and day.

	% Pos	% Neu	% Neg	% Day	Total
Monday	**67.64**	24.40	7.96	15.46	21,731
Tuesday	66.35	25.31	8.34	15.19	21,352
Wednesday	66.83	25.23	7.94	15.85	22,275
Thursday	66.21	25.38	**8.41**	15.65	21,993
Friday	67.60	24.91	7.49	14.80	20,808
Saturday	66.40	**26.20**	7.40	11.03	15,508
Sunday	67.11	25.31	7.58	12.02	16,901

with lowest score (negative) and those 500 messages with highest score (positive), the accuracy is very good: 96.78% for the positive class and 94.59% for the negative one. Tables 1 and 2 summarize these results.

When analyzing all the retrieved messages, we observe the following. Most of the messages are positive: 66.89% when considering all the messages (see Table 3), and 51.5% in the case of comments to other users' messages (see Table 4). At this point, it would be interesting to remind that (Boucher & Ossgod, 1969) stated that natural language tends to be positive in the human being.

We looked for the influence of the day of the week on the user mood, expecting to find, e.g., more positive messages during the weekend. The data showed that there is no great difference between different days (see Table 5, where higher and lower percentages are highlighted with bold font). Monday turned out to be the day in which most positive messages were sent (67.64% out of the messages sent this day), while Thursday is the day when

more proportion of negative messages were sent with respect to other days (26.20%). Anyway, the differences between percentages are not significant.

When looking at message contents, at first sight, it was observed that some messages classified as negative included irony or tease among friends. These messages should not be considered negative, since they do not convey negative sentiments actually. This may be the reason why the accuracy for the negative class is lower than that for the positive one.

We also noticed that many messages classified as positive where related to greetings. Moreover, a lot of them were short and easy to classify, such as "Felicidades!!!" ("*Congratulations!!!*"), "Feliz Navidad!!!" ("Merry Christmas!!!", ":-)" or "Jajajaja" (*laughs*). In order to verify this observation, we built a ranking of words based on their frequency of appearance. Apart from prepositions, articles and conjunctions, the words with positive sentiment polarity occupy the top of the ranking, confirming the greater presence of positive messages. The most frequent words did relate to greetings, confirming what we suspected. Words like "felicidades" ("congratulations"), "cumple" (short form of "birthday"), "besos" ("kisses") and "abrazos" ("hugs") occupied the top positions of the ranking.

Since greetings showed to have a high influence on the results obtained, to avoid this effect, we focused on analyzing, from all the messages available in Facebook, only those written by the user on his/her own wall (Facebook "status messages"). In Facebook, status messages usually relate to feelings, opinions or happenings related with the user's life. The users write these messages spontaneously, in contrast with greetings, which are usually written near the day an event happens, and, in some cases, they may even be written for commitment. According to (Viswanath, Mislove, Cha, & Gymmandi, 2009), the 54% of the interactions among a pair of Facebook users that barely keep in contact is due to birthday greetings. Therefore, by focusing on status messages, we expect to get more accurate information about the user own sentiment, since greetings are usually sent in the form of messages or comments to others' messages.

5.2. Evaluating our lexicon-based approach with status messages

In order to minimize the influence of greeting messages on the results obtained previously, we did another analysis that considered only status messages. Before proceeding with the classification, we built a rank of words appearing in status messages according to their frequency, to check the presence/absence of greetings. This time, the top of the ranking was occupied by words such as: "vida" ("life"), "gracias" ("thanks"), "nuevo" ("new"), "mejor" ("better"), "amigos" ("friends"), or "amor" ("love"). Moreover, it is worth to commenting that the set of most frequent words found is similar to that in the Spanish reference corpus list of frequencies (Real Academia Española, 2012).

When classifying status messages, it turns out that the percentage of positive messages decreases (from 66.89% to 41.19%, see Table 6), while the percentages of neutral and negative messages increase (from 25.21% to 44.93% in the case of neutral messages and from 7.9% to 13.88% for negative ones).

Table 6
Status messages per class.

	Total	%
Positive	17,440	41.19
Neutral	19,021	44.93
Negative	17,440	13.88
Total	42,337	100.00

Table 7
Messages per class and day.

	% Pos	% Neu	% Neg	% Day	Total
Monday	40.74	45.03	14.23	0.14	6029
Tuesday	40.73	44.60	**14.67**	0.15	6226
Wednesday	41.68	44.61	13.71	0.15	6557
Thursday	41.76	43.81	14.43	0.16	6619
Friday	**41.90**	45.08	13.02	0.15	6368
Saturday	40.40	**46.47**	13.13	0.12	5242
Sunday	40.88	45.26	13.86	0.13	5296

We grouped the messages according to the day of the week in which the users wrote them, obtaining the results shown in Table 7. Once again, the differences between the days were minimal, so that no generalization could be done.

Then, the results of the lexical-based wall message classification were compared with those from a human classification. As expected, the percentages of accuracies decrease with respect to those obtained when including greeting messages (which are much easier to classify). However, the results obtained are reasonably good too: the accuracy ranges between 78.91% for neutral messages and 81.56% for negative ones. It can be said that the accuracy of this method is around 80% (see Table 8).

5.3. Evaluating machine learning approaches and hybrid solutions

Since focusing on status messages seemed to be more reasonable because of the reasons explained in the previous subsection, we tested our hybrid approach with them. A dataset of 3000 status messages correctly labeled was available (1000 for each class: positive, neutral and negative), as a result of applying the lexicon-based classifier to status messages and contrasting its output with the opinion of a human judge. This dataset was used as the training set for machine-learning.

In this dataset, the number of messages per user was limited to ten, in order to avoid the influence of specific writing styles. We used Weka (Witten, Frank, & Hall, 2011), a set of tools for data analysis and predictive modeling. Once the dataset was defined, we applied a Weka filter to convert messages into word vectors, which are the input attributes for the classification method. The resulting vector for a message contains all the words appearing in the message along with the labeled class (Eq. (4)):

$$V = [w_1, w_2, w_3, \cdots w_n, class] \tag{4}$$

Equation 4: Word vector.

Some usual filtering was made: letters were converted to lowercase; and only words with two or more occurrences were considered. After applying the filters, more than 8000 tokens were obtained. That is, each item of the training dataset will have more than 8000 attributes. This is a huge number of dimensions. Therefore, attribute selection was required. After applying CFS (Correlation-based Feature Selection) (Hall, 1999) over the dataset, the number of attributes were reduced to 94 tokens.

Then we applied several machine-learning techniques with the goal of obtaining an accurate classifier (Witten et al., 2011). The algorithms used to build the classifiers were:

- J48 implementation of C4.5 decision-trees (Quilan, 1993).
- Naïve-Bayes.
- Support Vector Machines (SVM).

On one hand, we selected Naïve-Bayes and SVM because they have already been used for sentiment analysis (Pang et al., 2002) (Martínez Cámara, Perea, Valdivia, & Ureña, 2011). On the other

Table 8
Status message classification.

Predicted class	Current class				
	Positive	Neutral	Negative	Another language	Accuracy (Spanish) (%)
Positive	761	125	71	43	79.52
Neutral	84	715	107	94	78.91
Negative	96	83	792	29	81.56
Total	–	–	–	–	**80.02**

hand, decision-trees are easier to understand and interpret, and they make it easier to explain the rationale behind a given classification.

The domain of Facebook messages is boundless. The users can express themselves regarding any subject they want, which makes any text analysis difficult. On one hand, it is well known that using large sets of words as features in machine learning methods is not a good approach. On the other hand, it is very difficult to select appropriate features for classification from a huge domain.

We did a first attempt and used all the words that appeared in the messages collected from Facebook (without pre-processing them). The results obtained for each classification algorithm are shown in Table 9. As it can be seen, in this case the percentage of success is much lower that the percentages of accuracy obtained with the lexicon-based approach.

In a second attempt, we preprocessed the messages before feeding the classifiers with them, following some of the steps of the lexicon-based approach. Firstly, we used the spelling checker to correct misspellings. Secondly, emoticons were joined in two groups: positive and negative. Since the J48 decision-tree algorithm had obtained the highest accuracy previously (see Table 9), we started using this method. The resulting classifier showed a slight improvement (58.30% success) with respect to the previous one (53.97%). When analyzing the resulting tree, it was observed that the classification tree obtained was smaller this time. The most discriminative tokens were, precisely, the grouped emoticons. Following a similar approach, tokens representing laughs were grouped too. After this grouping, the accuracy increased to 61.07%, and the group of laughs turned out to be a very discriminative token too.

The results obtained when grouping emoticons and laughs suggested us that grouping the words by their semantic meaning could be a good approach. Therefore, the words were tagged (using POS tagging, as in the lexicon-based approach) and those words different from names, adjectives, interjections and verbs were removed, since they do not show sentiment. The dimensionality was reduced considerably at this step. Lastly, the remaining words were grouped according to their polarity scores, those previously assigned in the based-lexicon approach. That is, all the words different from emoticons or interjections were tagged as positive, neutral or negative. In that way, the attributes of each instance were: number of positive, neutral and negative words, number of positive and negative emoticons, number of positive and negative interjections, and number of laughs. All these actions did reduce

Table 9
Initial results using machine-learning methods.

Algorithm	Parameters	Accuracy (%)
J48	confidenceFactor = 0.25	53.97
Naive-Bayes		53.77
SVM	Kernel: radial $\exp(-gamma * |u - v|^2)$	49.77
SVM	Kernel: sigmoide $\tan h(-gamma * u' * v + \text{coef0})$	43.90

the dimensionality, leading to a finer attribute selection and, consequently, to better classification results.

In summary, the dataset was preprocessed in order to reduce its huge dimensionality, using lexicon-based techniques. Groups were formed with positive/negative interjections, positive/negative emoticons, and positive/neutral/negative words. After applying the J48 algorithm, the classifier got 83.17% accuracy, which is a good result according to the sentiment analysis literature.

The classification tree obtained was composed by 9 leaves and 8 internal nodes (see Fig. 6). This tree allows efficiency in computational terms and is robust analyzing different datasets since it does not suffer from overfitting with respect to the training set. For example, this tree predicts that a sentence without negative words ([neg] ≤ 0) will be negative if it contains negative emoticons ([ei_neg]); if it does not, then it will be positive is it contains positive words ([pos]), positive emoticons ([ei_pos]) or laughs ([ri_pos]).

The accuracies of the classifiers obtained when using different machine-learning algorithms are shown in Table 10. As it can be seen, all the classifiers lead to good classification results: the accuracy of all of them is over 83%.

6. Discussion

We have explored the use of lexical-based and machine-learning techniques to extract sentiments from messages written by users. The accuracy of the lexical-based approach when classifying status messages is higher than 80.02% when compared with a human judge. The success of the classifiers obtained from the application of machine learning techniques without lexical preprocessing was lower. However, when combining lexical-based and machine-learning techniques, the classification accuracy is higher: over 83%. These results are similar to others obtained in noticeable works of this research area. For example, when analyzing movie reviews written in English by following a lexical-based approach, 76.37% of accuracy was obtained in (Taboada et al., 2011). In (Pang et al., 2002), an accuracy of 82.90% was obtained when analyzing English movie reviews by using SVM. Regarding Spanish texts, (Martínez Cámara, Perea, Valdivia, & Ureña, 2011) obtained an accuracy of 86.84% when analyzing movie reviews using SVM. With respect to social network message analysis, we have found an approach to analyze texts retrieved from Twitter (Go et al., 2009). The researchers use SVM, and the accuracy achieved is 82.2%.

We think that, in the current context, the results of our approach are good, considering that extracting sentiments from Facebook messages is a great challenge, since:

- The domain is boundless; messages can refer to any subject (in contrast with, e.g., movie reviews).
- Facebook messages are informal, sent between friends, and, therefore, sometimes they include slang, irony, or words that are not written in Spanish. They may also contain spelling or grammatical mistakes.

These limitations make sentiment analysis harder and require the use of additional tools, such a language detectors or spelling checkers. Most of the previous works found in this area focus on formal text analysis. In most of them the domain is bounded (e.g., film reviews), which constitutes an advantage with respect to boundless domains. The only work found involving social networks is the one cited above, in which the accuracy achieved when classifying Twitter messages was 82.2%, slightly below the best result we obtained (83.27% when using SVM).

Besides the intrinsic characteristics of Facebook messages, the analysis is further complicated by the constraints imposed by the Facebook API regarding the data collecting procedure (already
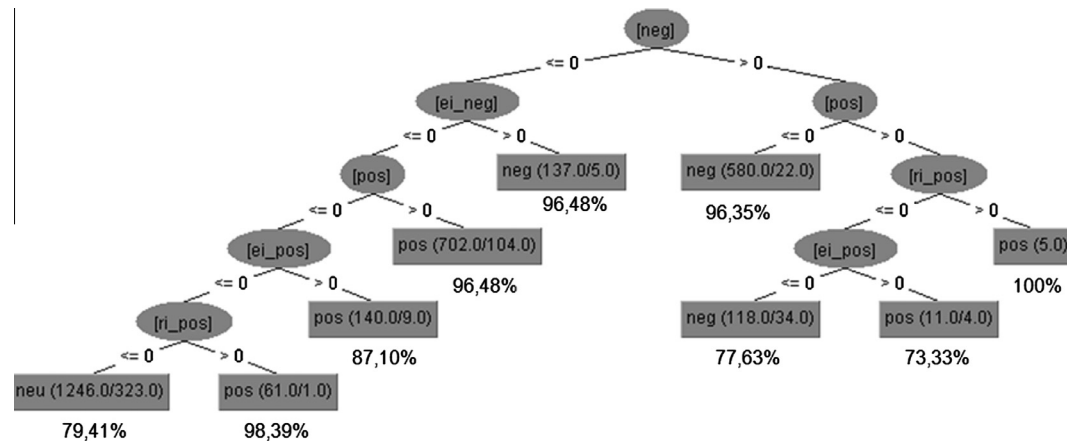
Fig. 6. Decision tree.

Table 10
Final results applying machine learning methods.

Algorithm	Parameter	Accuracy (%)
J48	confidenceFactor = 0.25	83.17
Naive-Bayes		83.13
SVM	Kernel: radial $\exp(-gamma * |u - v|^2)$	83.07
SVM	Kernel: sigmoide $\tanh(-gamma * u' * v + coef0)$	**83.27**

mentioned in Section 4.1, related to the maximum number of messages that can be retrieved with the same query from the same history, and the available time intervals). Furthermore, there is a non-documented constraint limiting the number of queries that a given application can make to the Facebook database, most probably to prevent their servers to be collapsed.

Considering these limitations, the question would be why choosing Facebook as a source of information regarding user's mood. The answer is that Facebook provide yet greater advantages: (i) there is no need to ask the user to carry out any specific activity or fill in any sentiment questionnaire, since most of them use Facebook (even daily, many of them); (ii) the nature of the messages exchanged through Facebook is spontaneous and informal, mainly between friends, and with good chances to have emotional load. These are the reasons why it was worth to design a method able to extract sentiment information from Facebook messages, in spite of the hard conditions.

Regarding the method employed for sentiment analysis, at the beginning of the research, as it was explained before, it was not possible to apply machine-learning approaches, since there were no examples of already classified messages to be used as training datasets. At that time, the lexicon-based approach was the only feasible option. However, once a large enough set of labeled messages was available, a combination of machine-learning and lexicon-based techniques was applied. We implemented them within a hybrid classifier, currently working on SentBuk. The best result (above 83% of accuracy) constitutes a small improvement of the results obtained by the lexicon-based approach (80%), also increasing the results obtained with machine-learning (lower than 54% in all the techniques used). Were a larger set of labeled examples available, it would be possible to improve the classifier, at least in theory. In practice, the classification tree shown in Fig. 6 represents a rather simple structure, which would possibly be hard to improve without suffering from overfitting. Furthermore, the lexicon-based approach cannot be completely replaced, since the preprocessing steps in the hybrid approach involves lexicon-based techniques: labeling the words as neutral, negative or positives is needed for reducing the dimensionality of the vectors used by the machine-learning algorithms.

6.1. Applications for e-learning

In the educational context, the capacity of knowing the student sentiment broadens the possibilities for e-learning. That information is especially useful for adaptive e-learning systems, which are able to guide each student through the learning process according to his/her particular needs and preferences at each time. In order to do so, they need to obtain and store information about each user in what is called the student model. We propose to incorporate information about student sentiments in student models so that this information can be used by them with adaptation purposes.

In particular, one adaptive e-learning system that can make use of sentiment information about students is CoMoLE. Its main adaptation capabilities have been presented in Section 2. There are different ways to provide sentiment-based support to students when interacting with any course through CoMoLE. One of them deals with the possibility of proposing motivational tasks to students whose emotional state is detected as negative, when guiding each of them. With this purpose, we have incorporated a new type of activities in CoMoLE: the motivational activities. They are activities whose goal can be no other than motivating the students (e.g., games, simulations, etc.). In such a way, when a student connects to the system to receive personalized advice about which educational activity to be carried out next (according not only to his/her progress but also to his/her needs and context at that time), if the system detects that he/she has a negative sentiment, then it can propose him/her a motivational activity to try to engage him/her first. The only requirements for adaptive systems to incorporate emotion-based adaptation are to include the user emotion as a new attribute in the user model, to feed this model with the results of the sentiment analysis, and to specify the adaptation/recommendation criteria based on this attribute, according to the desired support to be given to the user emotions.

Another application of sentiment analysis in adaptive e-learning systems is related to collaborative learning. In CoMoLE, the activities to be proposed to each group of students can vary depending on the group features, actions or context. Therefore, knowing the group sentiment makes it possible, for example, that when a workgroup show rather negative emotions towards a course, the system carries out specific motivational actions intended to encourage the group. Finally, in the case of collaborative e-learning systems that support automatic group formation on the

fly (Paredes, Ortigosa, & Rodríguez, 2010), information about the student sentiments can serve as input for group formation mechanisms. It would be useful, for example, to minimize the number of groups in which none of the students have positive sentiments towards the subject to work on; in other words, it will be helpful to minimize the number of groups in which all the students have negative feelings towards the subject.

A different useful application of sentiment analysis in e-learning consists of being able to detect positive/negative emotions towards an ongoing course, and using this information as feedback for the teachers or persons in charge of the course.

The way of getting information about the student feelings towards a specific course can be done in different ways. One of them consists of following the hybrid approach as presented in this paper in combination with student selection and topic detection: (i) from all the messages accessible to SentBuk, only those written by the students enrolled in the course are analyzed; (ii) from these messages, only the ones mentioning topics related to the course are considered (as well as the comments and likes associated to them). In order to follow this approach, a list of the students' usernames in Facebook, as well as a dictionary of keywords associated to the course, need to be built. In such a way, sentiment analysis would refer to the students' sentiments towards this specific course. This approach has the advantage that there is no need to create specific Facebook groups or pages associated to the courses. However, it has the disadvantage that the students have to give SentBuk permission to access to their whole Facebook profile, which they can dislike due to privacy issues. Another drawback deals with the keyword dictionary creation and, more specifically, with its use to filter on the messages retrieved. Depending on the subject, the dictionary can contain either very technical words (e.g., Botanic or Chemistry), which is good for message filtering, or words used daily (e.g., Social Science or Humanistic). In the latter case, it would be much more complex to determine whether the messages relate to the subjects taught.

An alternative way to reach the same goal consists of analyzing all the messages written by the students within Facebook pages/groups created for specific courses, when available. Nowadays, some teachers create Facebook pages or groups and make them available for their students to be used within the context of a given course. Students write in these pages at their own convenience. They normally use them to communicate with their partners, to ask for (or give) advice/help, to share information related to the topics involved in the course and, in general, to make comments about the course. In this case, the application of our sentiment analysis approach is straightforward: the only requirement is for the teacher to create the page and to grant SentBuk the permission to access to it. The messages, comments and links to be analyzed by SentBuk are all those appearing in the course page, and the history of the course starts when the page is created.

In any case, SentBuk analyses, weekly, the messages written by the students during the last week. Each student's sentiment towards the course, as obtained through our hybrid approach, is stored. Moreover, it is compared with the student "regular pattern" regarding the course, in order to detect changes, if any. Finally, it is used to update this pattern, by adding the sentiment detected in the last week to those from the previous weeks, as described in Section 3.2. Significant negative emotion changes may suggest the course responsible the need of taking certain steps or modifying recommendation criteria in case of adaptive e-courses. In the same direction, it is possible to detect the impact of new methodologies or activities on the students by knowing the positive/negative emotional reactions that they produce on them.

Although SentBuk performs the sentiment analysis weekly, it is possible for the teacher to ask SentBuk to analyze the messages written by the students at any time. In addition, specific time periods can also be selected through the SentBuk interface, so that the analysis considers only those messages, comments and likes within this period to get the emotion state of the user within this timeframe.

In summary, the analysis of the student writings in the pages associated to specific courses throws light about the student feelings towards these courses. This would make it possible to detect potential problems within courses dynamically, starting from the sentiments transmitted by the students enrolled in them. Moreover, the capacity of SentBuk to detect sentiment changes with respect to the usual sentiment state of a student makes it possible to detect peaks of positive/negative emotions, which can be associated to course contents (good/bad comprehension of them), specific activities proposed in class, or even teacher methodologies. Both individual and group sentiment changes can be considered. This information can serve as feedback for the teachers or the persons in charge of online courses, so that they can receive clues about which issues can be working better (or worse) in the context of a given course. This is especially useful in the context of distance learning, since, in this context, there is little (or none) face-to-face contact between students and teachers and, therefore, there are fewer opportunities for teachers to get feedback from the students.

7. Conclusions and future work

The work described in this paper demonstrates that it is feasible to extract information about the student's sentiments from the messages they write in Facebook with high accuracy. We have presented a new method for sentiment analysis in Facebook. It supports, on one hand, to get information about the users' sentiment polarity (positive, neutral or negative) according to the messages they write, and, on the other hand, to model the users' regular sentiment polarity and to detect significant emotional changes. With the aim of determining the feasibility of our approach, we have developed SentBuk, a Facebook application that retrieves the messages, comments and likes on the user's profiles, classifies the messages according to their polarity, and builds/updates the user sentiment profile. The latter version of SentBuk focuses on "status messages", along with those comments and likes associated to these messages. Other types of messages are discarded, since the large number of greetings that users write on others' walls gave rise to misleading results, as explained in Section 3. SentBuk shows the results through an interactive interface, which also supports the representation of emotional change detection, friend's emotion finding, user classification according to their messages, and statistics, among others.

The classification method implemented in SentBuk follows a hybrid approach: it combines lexical-based and machine-learning techniques. Even if the general idea of using lexicon-based sentiment analyzers is not new, on one hand, we have created our own lexicon-based approach, combining well-known techniques with our own enhancements to improve the results of the analysis of Facebook messages. On the other hand, we have created different hybrid classifiers by combining our own lexicon-based approach with several machine-learning techniques, to improve the performance of Facebook message sentiment analysis. Table 11 summarizes the percentages of accuracy obtained by each technique. The accuracy obtained with the combination of lexicon-based techniques (for preprocessing) and Support Vector Machines (for classifying) was the highest one (83.27%). In fact, this is the combination supported by Sentbuk currently.

Adaptive and recommendation systems in general can take advantage of knowing the users' sentiments at a certain time, as well as significant emotional changes with respect to their usual

Table 11
Accuracy of proposed classifiers.

Method	Accuracy (%)
Lexicon-based approach	80.02
Tree decision (J48-C4.5) + Lexicon-based tagging	83.17
Naïve-Bayes + Lexicon-based tagging	83.13
SVM + Lexicon-based tagging	83.27

state. However, asking the users about their sentiments is intrusive, can be bothering and, furthermore, in some contexts there is a high probability that they do not admit negative sentiments, because of, e.g., their cultural background or the need of social approval. In the educational context, in particular, it is not highly probable that a student directly transmits the teacher his/her feelings towards a subject or methodology when they are negative. The process to get this information is usually harder (e.g., anonymous/distinctive surveys at the end of the semester, or teachers receiving comments through class delegates during the course, at most). The work presented in this paper demonstrates that it is possible to extract it from the messages the students write in Facebook. This can be considered as input for adaptive e-learning systems to provide sentiment-based adaptation, making it possible, e.g., to recommend the most suitable activities to be performed by each student at a certain time according to his/her sentiment at that time; another application of this work in the learning context deals with extracting feedback for teachers about the sentiments of their students towards their courses or teaching methodologies. More details about the applications of sentiment analysis in the context of e-learning have been described in the previous section.

The main advantages and difficulties of using Facebook as a source of sentiment information have been explained in Section 3.1. In Facebook, the discourse domain is unbounded. Therefore, since the classifier developed makes use of a lexicon-based approach, this approach could be useful in other contexts in which sentiment analysis of Spanish texts is required (e.g., marketing or politics). The work developed has proved to be good enough to provide significant results and to be useful in the context of e-learning. However, some aspects could still be further evaluated or extended.

Regarding the SentBuk interface, although sentiment analysis can be done at any time (the only need is for the user to click on the corresponding button), the results of sentiment changes are shown, graphically, per week. In the educational context, it would be interesting to provide feedback to teachers either weekly or daily, or even at the level of hours or minutes. Therefore, a simple update should be done to the interface, so that teachers (or users in general) can decide the granularity of the sentiment representation (that is, the X axis unit).

Regarding the limitations of the sentiment analysis itself, currently, all the words tagged as positive get the same score, that is, all of them add 1 to the final sentiment computation. The same holds for negative words. However, it would be possible to assign a different weight to different words, depending on the strength of the sentiment they transmit. For example, the term "pésimo" (very bad) would have a stronger influence than "malo" (bad). In the same direction, currently, all the words grouped because of being considered the same, receive equal scores. For example, the word "jjjooooooooopeeeeee" is transformed to "jope" (interjection suggesting upset) and the repetition of letters is discarded. However, this repetition indicates an extra emphasis that could be taken into account if different weights were assigned to each word. The same happens with suffixes that correspond to intensifiers: "malísimo" (very bad) is more negative than "malo" (bad). The incorporation

of these variations to the method could be explored in order to check whether it is possible to achieve a finer-grained classification.

Another potential improvement would be related to the detection of sentiment changes. Currently, the temporal window to detect changes is one week, and the threshold to distinguish between small changes on user sentiment and significant changes that should be considered (by teachers or adaptive e-learning systems) is set to 0.5. We would like to carry out further tests to determine whether these values are the best ones in each case.

Moreover, we have observed, at first sight, that changes in a user sentiment, as detected from the messages he/she writes, could be related to changes in that user activity. In other words, a user that is happier than usually transmits more positive emotions than normally. Therefore, an emotional change is detected. It turns out to be that this user's activity is also higher during this week (number of likes, number of comments, and so on). We have looked for another user showing a significant change towards negative sentiment. We have found that this user's activity was also higher during the week the change was detected. We looked for similar cases and we found that, for users with positive sentiment changes, higher activity is detected. However, for users with negative sentiment changes, the activity is either rather higher or rather lower, but not similar to the regular activity. One hypothesis that arises when inspecting these users' profiles is that angry users might tend to share and express their angriness more actively, while sad users might tend to interact less with the network. However, these first observations must be tested adequately.

Acknowledgements

References

Abbasi, A. (2007). Affect intensity analysis of dark web forums. In *Proceedings of intelligence and security informatics (ISI)* (pp. 282–288).

Abel, F., Gao, Q., Houben, G. -J., & Tao, K. (2011). Analyzing user modeling on twitter for personalized news recommendations. In *19th International conference, UMAP 2011* (pp. 1–12). Girona, Spain: Springer Berlin/Heidelberg.

Aue, A., & Gamon, M. (2005). Customizing sentiment classifiers to new domains: A case study. In *Proceedings of Recent Advances in Natural Language Processing (RANLP)*, Borovets, Bulgaria.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python.* O'Reilly Media.

Boucher, J., & Ossgod, C. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behaviour, 8*(1), 1–8.

Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction, 11*(1/2), 87–110.

Brusilovsky, P., Schwarz, E., & Weber, G. (1996). ELM-ART: An intelligent tutoring system on World Wide Web. In *Proc. of 3rd international conference on intelligent tutoring systems, ITS-96* (pp. 261–269). Montreal: Springer.

Cardie, C., Farina, C., Bruce, T., & Wagner, E. (2006). Using natural language processing to improve eRulemaking. In *Proceedings of the 7th annual international conference on digital government research* (pp. 177–178), San Diego.

Carro, R., Ballesteros, F. J., Ortigosa, A., Guardiola, G., & Soriano, E. (2012) HYPERLINK. <http://www.bibsonomy.org/bibtex/6c910cf2aa9b3c48690deb3cabbdde09> AngryEmail? Emotion-based E-mail tool adaptation. *Lecture notes in computer science*, 7657 (pp. 399–406). Springer.

Carro, R., Pulido, E., & Rodríguez, P. (1999). Dynamic generation of adaptive based courses. *Journal of Network and Computer Applications, 22*, 249–257.

Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Model User-Adapted Interaction* (19), 267–303.

Costa, P., & McCrae, R. (1989). *The NEO-PI/NEO-FFI manual supplement.* Odessa, Florida: Psychological Assessment Resources.

Das, S., & Chen, M. (2001). Yahoo! for Amazon: Extracting market sentiment from stock message boards. In *Proceedings of the Asia Pacific Finance Association annual conference (APFA)*.

Dave, K., Lawrence, S., & Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of World Wide Web* (pp. 519–528), Budapest, Hungary.

De Bra, P., Aerts, A., Berden, B., de Lange, B., Rousseau, B., Santic, T., et al. (2003). AHA! The adaptive hypermedia architecture. In *Proceedings of 14th ACM conference on hypertext and hypermedia* (pp. 81–84).

de Montjoye, Y.-A., Quoidbach, J., Robic, F., & Pentland, A. (2013). Predicting people personality using novel mobile phone based metrics. *Social Computing, Behavioral–Cultural Modeling and Prediction. Lecture Notes in Computer Science, 7812*, 48–55.

Engström, C. (2004). *Topic dependence in sentiment classification*. Master thesis. University of Cambridge.

Esuli, A., & Sebastiani, F. (2006). SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of Language Resources and Evaluation (LREC)* (pp. 417–422).

Felder, R. (1996). Matters of style. *American Society for Engineering Education (ASEE) Prism, 6*(4), 18–23.

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM, 56*(4), 82–89.

Go, A., Bhayani, R., & Huang, L. (2009). *Twitter sentiment classification using distant supervision*. Technical report. Stanford University. Stanford Digital Library Technologies Project.

Hall, M. A. (1999). *Correlation-based feature selection for machine learning*. PhD thesis. University of Waikato.

Hatzivassiloglou, V., & McKeown, K. (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the eight conference of the European chapter of the Association for Computational Linguistics (ACL/EACL)* (pp. 174–181).

Jin, X., Li, Y., Mah, T., & Tong, J. (2007). Sensitive webpage classification for content advertising. In *Proceedings of the 1st international workshop on data mining and audience intelligence for advertising (ACM)* (pp. 28–33), San Jose, CA, USA.

Kiss, J. (10.04.2012). *Facebook hits 1 billion users a month* (Retrieved February 2013). The Guardian. <http://www.guardian.co.uk/technology/2012/oct/04/facebook-hits-billion-users-a-month>.

Kobsa, A. (2007). Generic user modeling systems. *The adaptive web: Methods and strategies of web personalization. Lecture notes in computer science* (Vol. 4321, pp. 136–154). Springer.

Liu, H., Lieberman, H., & Selker, T. (2003). *A model of textual affect sensing using real-world knowledge* (pp. 125–132). Proceedings of Intelligent User Interfaces (IUI).

Liu, B. (2010). Sentiment analysis and subjectivity. In N. Indurkhya & F. J. Damerau (Eds.), *Handbook of natural language processing* (pp. 627–666). Chapman and Hall: CRC Press.

Martí, M. A., Taulé Delor, M., Màrquez, L., & Bertran, M. (2007). *CESS-ECE: A multilingual and multilevel annotated corpus*. Universitat de Barcelona (CLiC-UB); Universidad de Alicante (UA; Universitat Politècnica de Catalunya (UPC); Euskal Herriko Unibertsitatea (EHU-UPV).

Martín, E., & Carro, R. M. (2009). Supporting the development of mobile adaptive learning environments: A case study. *IEEE Transactions on Learning Technologies, 2*(1), 23–36.

Martínez Cámara, E., Perea, J. M., Valdivia, M. T., & Ureña, L. A. (2011). Técnicas de clasificación de opiniones aplicadas a un corpus en español. *Procesamiento del Lenguaje Natural* (47), 163–170.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*, 39–41.

Mishne, G., & Glance, N. (2006). Predicting movie sales from blogger sentiment. In *AAAI symposium on computational approaches to analysing weblogs (AAAI-CAAW)* (pp. 155–158).

Moore, A., Brailsford, T., & Stewart, C. (2001). Personally tailored teaching in WHURLE using conditional transclusion. In *Proc. 12th ACM conf. hypertext and hypermedia* (pp. 163–164).

Nasoz, F., & Lisetti, C. L. (2007). Affective user modeling for adaptive intelligent user interfaces. *Human–Computer Interaction, III*(4552), 421–430.

Ortigosa, A., Carro, R., & Quiroga, J. (2013). Predicting user personality by mining social interactions in Facebook. *Journal of Computer and Systems Sciences. Special issue on Intelligent Data Analysis*. http://dx.doi.org/10.1016/j.jcss.2013.03.008 (ISSN 0022-0000).

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the seventh conference on international language resources and evaluation* (pp. 1320–1326).

Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the Association for Computational Linguistics (ACL)* (pp. 271–278).

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP)* (pp. 79–86).

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval, 2*(1–2), 1–135.

Paredes, P., Ortigosa, Á., & Rodríguez, P. (2010). A method for supporting heterogeneous-group formation through heuristics and visualization. *Journal of Universal Computer Science, 16*(19), 2882–2901.

Quilan, J. (1993). *C4.5: Programs for machine learning*. California, USA: Morgan Kaufmann Publishers.

Ramirez-Esparza, N., Pennebaker, J., Garcia, F., & Suria, R. (2007). La psicología del uso de las palabras: Un programa de comutadora 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.

Real Academia Española (2012). *Corpus de Referencia del Español Actual (CREA) – Word frequency listing*. <http://corpus.rae.es/lfrecuencias.html> (Retrieved 03.04.12).

Reyes, A., Rosso, P., & Buscaldi, D. (2012). From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering, 74*, 1–12.

Rogers, E. (2003). *Diffusion of innovations*. New York: Free Press.

Shen, L., Wang, M., & Shen, R. (2012). Affective e-learning: Using "Emotional" data to improve learning in pervasive learning environment. *Educational Technology & Society, 2*, 176–189.

Spada, D., Sanchez-Montanes, M., Paredes, P., & -@@- Carro, R. M (2008). Towards inferring sequential-global dimension of learning styles from mouse movement patterns. *Adaptive hypermedia and adaptive web-based systems. Lecture notes in computer science* (Vol. 5149, pp. 337–340). Springer.

Spertus, E. (1997). Smokey: Automatic recognition of hostile messages. In *Proceedings of Innovative Applications of Artificial Intelligence (IAAI)*.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics, 37*(2), 267–307.

Tatemura, J. (2000). Virtual reviewers for collaborative exploration of movie reviews. *Proceedings of Intelligent User Interfaces (IUI)*, 272–275.

Thurstone, L. (1938). *Primary mental abilities*. Chicago Univ. Press.

Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the Association for Computational Linguistics (ACL)* (pp. 417–424).

Vilares Ferro, J. (2005). *Aplicaciones del procesamiento del lenguaje natural en la recuperación de información en español*. PhD thesis. Universidade da Coruña.

Viswanath, B., Mislove, A., Cha, M., & Gymmandi, K. (2009). On the evolution of user interaction in Facebook. In *Proceedings of the 2nd ACM workshop on online social networks* (pp. 37–42).

Wikipedia (2013, February). *List of emoticons* (Retrieved February 2013). Wikipedia. <http://en.wikipedia.org/wiki/List_of_emoticons>.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann Publishers.