Recognizing Polarity and Attitude of Words in Text

Laritza Hernández[1], A. López-Lopez[2] and José E. Medina[1],

[1] Advanced Technologies Application Center, La Habana, Cuba
{lhernandez,jmedina}@cenatav.co.cu
[2] Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México.
allopez@inaoep.mx

Abstract. In this work, we present a problem of Sentiment Classification in texts. Sentiment Classification is an Opinion Mining and Sentiment Analysis task that has opened several problems, including a significant variety of applications. We propose a strategy to distinguish words that convey evaluation of an item from the rest, as well as to classify the evaluation polarity (positive or negative). In addition, relying on Appraisal Theory, we intend to classify the evaluation words in affect, judgment and appreciation. Both, polarity and attitude are recognized using a corpus-based approach. We have the purpose of applying this task in Spanish texts; thereby, we have created a corpus of movie reviews in this language that was manually processed. In the experiments, we noticed that our strategy has a good performance, achieving 76.78% and 80.16% of accuracy classifying polarity and attitude, respectively.

Keywords: Opinion Mining, Sentiment Analysis, Sentiment Classification, Appraisal Theory.

1 Introduction

Sentiment Classification is an Opinion Mining and Sentiment Analysis task; these are novel research areas strongly related. Some initial works dates back to the late 1980's and early 1990's [19], [21]; however, today there is a lot of controversy about the boundaries between these two areas. Some authors have defined Opinion Mining as a task "in which text mining methods are used to find interesting and insightful correlations between writer's opinions" [1]. Whereas, Sentiment Analysis is conceived as Sentiment Classification, referring to the task of categorizing texts, or pieces of text, based on their subjectivity and orientation [18]. Others extend it to identify or classify appraisal targets, determining the source of an opinion in a text, and developing interactive and visual opinion mining methods [3].

In this paper, we focus on the Sentiment Classification task; we propose a strategy to determine whether a given word conveys the evaluation of an item, and recognizing the evaluation kind. There have been previous works trying to make a distinction of evaluation kind in text finer than single semantic orientation or polarity (e.g. positive, negative or neutral). Some authors attempt to discern kinds of emotions (affects) [7], [13]. Whereas others, relying on Appraisal Theory, seek out expressions of attitude (affect, judgment, and appreciation) [14], [18]. Our work is one in this latter line. Besides, we classify the polarity of word in positive, negative or no-polarity.

Appraisal Theory studies the evaluative use of language, and is divided in three subsystems. Attitude corresponds to the words that emit an evaluation or that invite to take it. Graduation considers the words that intensify, diminish, sharpen or blur the evaluation. Engagement corresponds to those words that indicate the posture that the issuer adopts with its statement. In this paper, we focus only on Attitude that is subdivided in affect (evaluation of sentiments or emotional states), judgment (evaluation of the human behavior), and appreciation (evaluation of objects, processes, or people when they are valued from an aesthetic viewpoint). Attitude, also, can be positive or negative.

According to the Appraisal Theory, there is an overlap among the affect, judgment, and appreciation categories, since affect is considered as the basic system of Attitude, whereas judgment and appreciation are derivations of this, manifesting institutionalized emotions [8].

This theory has been poorly studied in terms of Computational Linguistics and Natural Language Processing; we have only found research for English. Even from a linguistic viewpoint, this theory has not been researched enough for the Spanish language; we have only found three reported works [8], [9], [20]. Two of them are very interesting studies of Kaplan, which we rely on for this work. In Fig. 1 we show some of the categories of the Appraisal Theory and we provide some examples of Spanish words expressing Attitude.

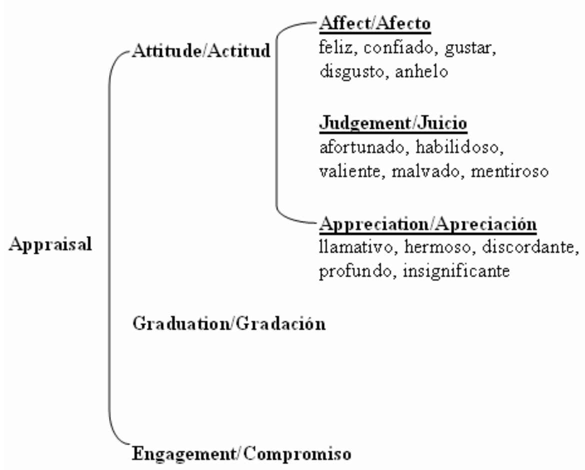

Fig. 1. Categories of Appraisal Theory and example words of Attitude system in Spanish.

Many words of an attitude system, according to the Appraisal Theory, have potential to express affect, judgment and appreciation when we consider them out of context. This situation has motivated us to use a corpus-based approach. This approach allows recognizing the evaluation of words considering the context where these tend to occur.

The problem of evaluative natural language processing has been poorly explored for Spanish [2], [4]. As a result, there is a lack of tools that limits the solutions of Sentiment Classification in this language. In this work, we are taking our first steps toward that goal.

Evaluative words can be found in different kinds of documents. One of the most analyzed types is the online reviews about products, such as movies, computers, phones and others. These are documents carrying out a free style of writing, where a great variety of evaluative expressions can be observed, as well as the three attitude kinds of the Appraisal Theory. But, we should mention that theses documents are not heavily loaded of judgment expressions. However, we can find other document kinds; e.g. editorials, in which a kind of more elaborated discourse can be appreciated. In this state of research, we have manually prepared a corpus, from movie reviews in Spanish taken from the website ciao.es. This is a very useful experimentation resource[1] as we will explain later.

The Sentiment Classification has achieved and opened a significant range of applications from monitoring and automatically summarizing user opinions about commercial products, people, organizations, and so on, up to the evaluation of public relations and marketing firms. For example, in sentence (1) we can notice two subjective words, both with a positive polarity. We could infer that the polarity of the sentence is positive as well, by computing the amount of positive words.

(1) "*Viéndola, me doy cuenta de que si tanto me ha* <u>*gustado*</u>$^{+}$ *es porque la trama es* <u>*comprensible*</u>$^{+}$."

Recognizing attitude in the words (see sentence (2)) allows knowing the evaluation purpose of a sentence (sentiments, objects, or human behavior). Thus, sentences more relevant to a particular interest could be identified. For example, if the concerned item is the human behavior, you could be interested in what anyone says about the capacity, or moral integrity of a given person, more than in its physical appearance. This could be useful for tasks like opinion retrieval, opinion summarization, question answering focused to opinions, and others Opinion Mining tasks.

(2) "*Viéndola, me doy cuenta de que si tanto me ha* [**affect**: <u>*gustado*</u>$^{+}$] *es porque la trama es* [**appreciation:** <u>*comprensible*</u>$^{+}$]."

Thus, the contributions of this paper are to present a strategy for recognizing whether a given word has a polarity, positive or negative (already considered as expressing attitude) or whether it is a word with no polarity (regardless of attitude). Furthermore, according to the Appraisal Theory, if the word expresses attitude, we recognize the attitude kind; i.e., affect, judgment and appreciation. Thereby, we intend to capture the polarity of a word by other words that tend to occur in the same sentences. In a similar way, we try to capture the attitude classes of a given word, but considering the item evaluated in the sentences. On the other hand, we report our initial results toward the sentiment classification on Spanish texts.

[1]We would be glad to share this data set, if interested, contact us by e-mail.

We have structured the present work as follows. In Section 2, we briefly explain some related work to solve the *s*entiment classification problem and some of their drawbacks. In Section 3, we describe the proposed method. In the last section, we present the experimental results by using a textual corpus which we prepared from movie reviews in Spanish, selected from the website ciao.es.

2 Related Work

The detection and classification of subjective words in text using corpus-based approaches rely on syntactic patterns or co-occurrence of words. These approaches allow recognizing the polarity and attitude of the words determined by the context where they tend to appear. Nevertheless, it has the drawback of being dependent on the domain of the corpus used, and in consequence, usually tend to identify an insufficient set of words. Next, we describe some works in the context of our proposal.

Turney and Littman, in 2002 and 2003, proposed a strategy that intends to infer the "semantic orientation" or evaluative character of a word from extremely large corpora, considering its semantic association with other words, which he called "paradigms" [16], [17].

These authors consider that a word has a positive semantic orientation if it conveys the evaluation that the item is desirable; and a negative orientation if it conveys the evaluation that the item is undesirable. They determine the positive/negative semantic orientation of a given word *w* from the strength of its semantic association with the positive/negative paradigms. Then, they compute the semantic orientation of *w* as the difference between both positive and negative strengths. The words (good, nice, excellent, positive, fortunate, correct, and superior) and (bad, nasty, poor, negative, unfortunate, wrong, and inferior) were taken as positive and negative paradigms (*pp* and *np*), respectively.

Turney and Littman in a previous work, based on Pointwise Mutual Information (PMI), proposed a measure of word semantic association using information retrieval, called PMI-IR [15]. The measure PMI-IR intends to determine which "alternative", given by a set of choices; i.e., $\{choice_1, \ldots, choice_n\}$, corresponds to a given word *w*, that they called "problem". They use Church and Hanks' PMI defined as a measure that estimates word association norms determined by word co-occurrences in a corpus as follows [6].

Let $P(w)$ and $P(choice_i)$ be the probabilities of two words, *w* and $choice_i$ respectively; then $PMI(w, choice_i)$ is the mutual information between *w* and $choice_i$ defined as:

$$PMI(w, choice_i) = \log_2\left(\frac{P(w, choice_i)}{P(w)P(choice_i)}\right) \quad \textbf{(1)}$$

$PMI(w, choice_i)$ can be interpreted as the relation existing between the probability of *w* and $choice_i$ co-occurring in the same context, and the probability of *w* and $choice_i$ when they are statistically independent. Considering that Turney and Littman were looking for the maximum score, they proposed to drop log2, (because of its monotonicity) and P(w) (because "it has the same valued for all choices, for a given problem word"). Thus, (1) can be simplified to:

$$PMI(w, choice_i) = \frac{P(w, choice_i)}{P(choice_i)} . \quad (\mathbf{2})$$

On the other hand, these authors use four ways to calculate the probabilities in (2). But, we did not consider them because these probabilities were calculated as the number of returned matching documents from AltaVista Advance Search, by means of hits (query), and using the NEAR operator.

Turney and Littman method uses an extremely large corpus that does not require a manual process for its preparation, but this method depends on the variations and availability of an online search system. Besides, the NEAR operator considers that two words are close when they are halfway at least ten words, but it does not distinguish if the words are in the same sentence, an aspect that we consider to keep in mind.

In 2009, Brooke proposed the creation of semantic orientation Spanish dictionaries, making an analogy with adjective, noun, verb, adverb, and intensifiers dictionaries in English [6]. Each adjective, noun, verb, adverb dictionary in English is automatically translated to Spanish by means of the online bilingual dictionary Spanishdict and online Google translator, maintaining the semantic orientation of words from English. Also for the bilingual dictionary and translator, the author proposed other method using a textual corpus in Spanish formed by 400 reviews about hotels, movies, music, phones, washing machines, books, cars, and computers. From this corpus, adjectives, nouns, verbs, adverbs, and intensifiers dictionaries, with the semantic orientation for each word were extracted. In the comparison of the obtained dictionaries, Brook comments that the biggest agreement was in the adjective dictionary; but its semantic orientation agreement was the worst.

That is a valid approach to cope with the problem of sentiment classification in Spanish. But, since the words "subjective sense", as well as the intensity of this subjective, can be lost in the translation, we consider that a finest study has to be done where the proper variables of this language are taken into account, avoiding loss of generality, as far as possible.

On the other hand, Taboada & Grieve work, in 2004, used a similar strategy as that applied by Turney and Littman to classify adjectives relying on the Appraisal Linguistic Theory. This classification is used to calculate the degree to which a review (opinion texts about movies, books, cars, cookware, phones, hotels, music, and computers) expresses affect, judgment, or appreciation [14].

These authors improved Turney classification because it does not only determine whether an adjective is positive or negative, but the "adjective' overall evaluative potential", defined as the probability of using an adjective in evaluative discourses to express affect (AfP), judgment (JP), or appreciation (ApP). These are calculated as follows.

$$AfP = \frac{MI(I\ was, w)}{MI(I\ was, w) + MI(he\ was, w) + MI(it\ was, w)}, \quad (\mathbf{3})$$

$$JP = \frac{MI(he\,was, w)}{MI(I\,was, w) + MI(he\,was, w) + MI(it\,was, w)}, \qquad (4)$$

$$ApP = \frac{MI(It\,was, w)}{MI(I\,was, w) + MI(he\,was, w) + MI(it\,was, w)}, \qquad (5)$$

$$MI(PRO\,was,\, A) = \log_2\left(\frac{hits(PRO\,was\,w)}{hits(PRO\,was)hits(w)}\right), \qquad (6)$$

where MI is the mutual information between a word (adjective) w and a "pronoun-copula" pair (PRO) estimated from hits(query) on AltaVista.

This is a first interesting approach to classify attitude using context; but, there are several examples that show that the three proposed combinations (I was (affect), He was (judgment), It was (appreciation)) are not enough and they even fail.

Whitelaw, Garg and Argamon, in 2005, presented a method for sentiment classification that extracts and analyzes "adjectival appraisal groups" from texts, relying on the Appraisal Theory. Although they do not use a corpus-based approach, we consider important to comment their work [18].

These authors consider that an adjectival appraisal group, in English, is a coherent group of words that expresses together a particular attitude. It is formed by a head appraising adjective with an optional preceding list of appraisal modifiers (*very*, *sort of*, *not*, and other), each denoting a transformation of one or more appraisal attributes of the head. They also take into account the English language, easing the extraction of adjectival appraisal groups, i.e., they consider the word-order, inherent to this language, to remove all pre-modifiers of an appraising adjective. This word-order is not the same in Spanish, for instance.

They used semi-automated methods to build a lexicon of appraising adjectives and modifiers. They obtained, from word and phrase seeds taken from Martin [10] and Matthiessen [5] works and supported on WordNet and other thesaurus, expanded lists of "candidate" terms for each appraisal category that they considered (only the related word was included in the lists; i.e., synonyms, members of each synset and others, but with same part-of-speech as seed term). Then, the terms of each category obtained by this process were ranked by its occurrence frequency in the candidate lists. Later, a manual inspection was carried out to obtain the list of final terms, by removing less frequent ones from each category in order to reach a final set.

3 Proposed Strategy

As we known, many words in the human language are ambiguous (they do not convey a single message) when they are studied out of context; i.e., the context strongly determines the word sense. The evaluative language is not an exception (e.g. it is difficult to know if *big* or *much* conveys a negative or positive evaluation). On the other hand, according to proponents of the Appraisal Theory, some words out of context can be ambiguous according to their attitude class (e.g. *aburrido* (boring), *cómodo* (pleasant), or *agradable* (nice)). For this reason, we believe that we have to consider the contextual relations among the words in our proposal. This work, based on a corpus approach, tries to discriminate among attitude (words that convey an item evaluation, positive or negative) and no-attitude (words that do not convey an evaluation). Besides, relying on the Appraisal Theory, we intend to recognize the attitude of words. We describe a supervised strategy to learn sentiment classifiers of words. This is not a pioneer work in the machine learning using to solve sentiment classification problems; for example, Pang and Vaithyanathan (2002) and Mullen, N. Collier (2004) employed this technique for movies reviews classification [11], [12], and Wilson, Wiebe, and Hoffmann (2005) use it to classify sentiment expressions [22].

Next, we present the word classification divided in two parts. First, we present how to classify the polarity of words in positive, negative, and no-polarity. Second, we proposed how to classify attitude of words in affect, judgment, and appreciation. Then, we describe the corpus and the lexicon that we prepared as validation tools.

First, we assume that the sentences are the atomic units of coherent messages in texts. Therefore, we assume that the words that tend to co-occur in the same sentences are used with the intention to express similar or identical messages.

If one sentence conveys some appraisal, the following can happen:

a) There are words that indicate the appraisal in an explicit way.
b) There are no words that indicate the appraisal in an explicit way, but implicitly.
c) The polarity of appraisal is indicated with words of the opposed polarity. (Irony)

In this stage of our research, we are only interest in the first case; i.e., we only study the appraisal that is indicated in an explicit and direct way.

Then, considering the assumptions above, we also assume that words with a given polarity probably tend to occur in sentences of same polarity. That probably does not happen in sentences with different polarity or without polarity. Therefore, if we start forming a set of seed words and its polarity (positive, negative, and no-polarity), and if we represent them by a vector of words that occur in their sentence; then we assume that is possible to learn the context (words) of each polarity class and increasing our lexicon with new words of same polarity.

Subsequently, the attitude class of words is related to its sense and to the item (sentiment, human behavior, or object) target of evaluation. In this first stage of the research, we have assumed that in a single sentence, the evaluation of a single item prevails. Therefore, we start from the hypothesis that the words that tend to co-occur in the same sentence are being used with the intention of expressing the same kind of attitude.

According to our hypothesis, words are represented by a vector of dimension n, where n is the corpus size (sentences), assuming that sentences can be adequate to discriminate the attitude class of words. Given a word, the n-entry of the associated vector is 1 if the word is in the n-th sentence, and 0, otherwise.

Finally, taking into account the overlap of the classes inherent to Appraisal Theory, which we also found in polarity as well (see next section), we consider that some words might potentially be in more than one of these classes. Therefore, we do not treat the polarity classification neither attitude as a multi-classification problem. But, we provide a binary classifier for each polarity and attitude class. For example, for no-polarity, we take as positive examples all word vectors labeled with no-polarity, and as negative examples the remaining vectors labeled as positive and negative. Similarly, we proceed with positive and negative class and with the three attitude class as well.

3.1 The Corpus

We use as corpus a set of sentences selected from movie reviews in Spanish, gathered from the website ciao.es. This corpus was manually prepared, selecting from each review the sentences that were considered as containing words expressing some attitude class, denoted as "attitudinal sentences". Furthermore, we added some sentences that did not contain that kind of words, denoted as "non attitudinal sentences". To identify the words being used to express attitude, the Appraisal Theory for Spanish was taken into account [8], [9], [20]. The manual process relied on a set of tutorials and examples of this theory, as well as an annotation scheme that we prepared. Thus, we created a corpus of 1408 sentences, composed by attitudinal and non attitudinal sentences.

In addition, we compile all the words annotated manually with attitude in a list of 1247 terms. Then, taking into account the overlap of the classes, inherent to Appraisal Theory, we compile a list of words per class, being the appreciation the larger class and judgment the smallest one (788 and 287 words, respectively), and we calculated their percentages of overlap (see Fig. 2).

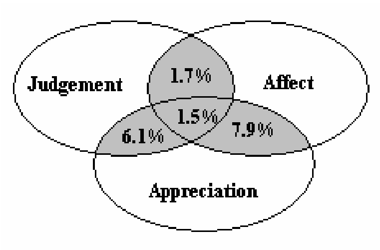

Fig. 2. Overlap percentages among attitude classes.

Table 1. Data collection statistics.

List	**Size**
Affect	352
Judgment	287
Appreciation	788
Positive	573
Negative	389
Non-attitude	178
Corpus	
No. sentences	1408
No. words	32, 920

Moreover, besides affect, judgment, and appreciation, some words from the 1247 terms manually annotated, were annotated in context as positive and negative, and we compiled two lists with 573 and 389 words, respectively, which had a 4% of overlap. Finally, we prepare a list with words that were not annotated, i.e. no-polarity words. We should remark that these words do not correspond with what has been defined as neutral polarity, which refers to words on which we can not decide about their polarity. In the Table 1, the size of the prepared collection is summarized.

4 Evaluation

To validate the proposed strategy in this work, we use as reference of "*good-classification*" the five lists of words manually classified, described in previous section. Thus, the obtained results are compared against the human judgment. Since the number of examples in the classes is unbalance, we used an over-sampling method called Smote (Synthetic minority over-sampling technique), that increases the proportion of the instances from the minority class.

Regarding the classifiers, we use the suite of Data Mining algorithms that Weka system[2] provides; namely, K*, Support Vector Machine (SVM), and Naive Bayes, NB. We selected these algorithms because we need to estimate the membership degree of the words in each class. We maintained default parameter values for each classifier, except for SVM that we set true the "buildLogisticModels" parameter to obtain the probabilistic version of this algorithm. To measure the performance of classification, we used 10-folds cross validation test method, and Precision, Recall, F-Measure and Accuracy percentage (see Tables 2 and 3). We can note that SVM and K* algorithms show the better results.

In the comparison of our results against human judgment, a good performance of attitude classification can be observed, resulting appreciation and positive the classes with worst results. Besides, we compare our proposal against the Turney & Litman results that we commented in the related work section. The other works in that section have not available results for the comparison, for this reason we do not refer to them here. We intend to use the same lists of words (657 positives and 679 negatives words from General Inquirer lexicon), but not all the words were in the English prepared corpus from The SFU Review Corpus[3] (only 114 positive and 53 negative words were recognized). The results are displayed in Table 4. We did not find reported results to compare the attitude classification.

Although our results are not directly comparable with the results of the selected method, since the corpora used in the evaluation are different, we can observe a good performance of the proposed method; even though these are not conclusive, as it is commented later on, in the future work.

[2] http://www.cs.waikato.ac.nz/ml/weka/

[3] http://www.sfu.ca/~mtaboada/research/SFU_Review_Corpus.html.

Table 2. Precision, Recall, F-Measure of SVM, K* and NB classifiers for polarity and attitude classification.

Polarity				Attitude		
Positive				**Affect**		
	Precision	**Recall**	**F-Measure**	**Precision**	**Recall**	**F-Measure**
SVM	***0.51***	***0.615***	***0.558***	0.84	***0.832***	***0.836***
K*	0.505	0.575	0.538	***0.863***	0.808	0.835
NB	0.52	0.188	0.276	0.962	0.539	0.691
Negative				**Judgment**		
	Precision	**Recall**	**F-Measure**	**Precision**	**Recall**	**F-Measure**
SVM	0.59	***0.779***	0.673	0.874	***0.866***	0.87
K*	***0.597***	***0.779***	***0.676***	***0.889***	0.856	***0.872***
NB	0.564	0.776	0.653	0.734	0.595	0.734
Non-attitude				**Appreciation**		
	Precision	**Recall**	**F-Measure**	**Precision**	**Recall**	**F-Measure**
SVM	0.864	***0.837***	***0.85***	***0.572***	0.632	0.601
K*	***0.913***	0.72	***0.85***	0.553	***0.679***	***0.609***
NB	0.839	0.652	0.743	0.539	0.939	0.685

Table 3. Accuracy of SVM, K* and NB classifiers for polarity and attitude classification.

	Accuracy (%)		
Class	**SVM**	**K***	**NB**
Polarity	***76.78***	70.45	58.22
Attitude	***80.16***	79.88	64.87

Table 4. Accuracy of Hatzivassiloglou & McKeown, Turney & Littman -PMI and LSA, and our method for positive, negative, and no-attitude class.

Methods	Corpus size (words)	Accuracy (%)
Proposed method (SVM)	94 905	87.21
Proposed method (K*)	94 905	86.76
Turney & Littman -PMI	One-hundred-billion	82.84
Turney & Littman -PMI	Two billion	76.06
Turney & Littman -PMI	Ten million	61.26

5 Conclusions

In this paper, we showed our initial steps toward sentiment classification on Spanish texts, first recognizing word polarity (positive, negative, and no-polarity), and attitude (affect, judgment, and appreciation).

The results show a good performance of the proposed classification strategy when is compared against human judgments and early proposals, achieving 76.78% and 80.16% accuracy in polarity and attitude classification in Spanish (only considering

SVM algorithm). Also, we can note that when we compare our proposal with earlier results, promising results are observed. But, to reach a final conclusion, we have to test the reported method in the same or similar corpus, as far as possible.

In further works, we plan to prepare a new version of Spanish corpus, by increasing the size of sentences, and extending their domain. Besides, we will consider that more than one item could be evaluated inside a single sentence. We will work in classification of expressions (word sequences) rather than individual words, also tackling the problem with a multiclass approach, considering the overlap inherent to Appraisal Theory.

References

1. Argamon, S., Bloom, K., Esuli, A. and Sebastiani, F.: Automatically Determining Attitude Type and Force for Sentiment Analysis. *Procs. of the 3rd Language and Technology Conference* (LTC'07), pp. 369--373 (2007)
2. Balahur, A., Montoyo, A: Multilingual Feature–Driven Opinion Extraction and Summarization from Customer Reviews. *Procs. of the 13th international conference on Natural Language and Information Systems: Applications of Natural Language to Information Systems*, Lecture Notes In Computer Science, vol. 5039, pp. 345--346 (2008)
3. Bloom, K., Garg, N.and Argamon, S.: Extracting Appraisal Expressions. *Procs. Of North American Chapter of the Association for Computational Linguistics - Human Language Technologies conference*, pp. 308--315 (2007)
4. Brooke, J.: *A semantic approach to automated text sentiment analysis*. Master Thesis. Simon Fraser University (2009)
5. Matthiessen, C.: *Lexico-grammatical cartography: English systems*. International Language Sciences Publishers (1995)
6. Church, K.W., & Hanks, P.: Word Association Norms, Mutual Information and Lexicography. Procs. of the 27th Annual Conference of the Association of Computational Linguistics New Brunswick, NJ: *Association for Computational Linguistics*, pp. 76--83 (1989)
7. Genereux, M., Santini, M.: Exploring the use of linguistic features in sentiment analysis. *Corpus Linguistics* 2007, Birmingham, UK, pp. 27-30
8. Kaplan, N.: Nuevos Desarrollos en el Estudio de la Evaluación en el Lenguaje: La Teoría de la Valoración. *Boletín de Lingüística*, Universidad Central de Venezuela, pp. 52--78 (2004)
9. Kaplan, N.: *La Construcción Discursiva del Evento Conflictivo en las Noticias por Televisión*. PhD Thesis Universidad Central de Venezuela, Facultad de Humanidades y Educacion (2007)
10. Martin, J. R., White, P.R.R.: *The Language of Evaluation: Appraisal in English*. Palgrave, London (2005)
11. Mullen, T., Collier, N.: Sentiment analysis using Support Vector Machines with diverse information sources. *Procs. of 2004 conf. on Empirical Methods in Natural Language* (2004)
12. Pang, B., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. *Procs. of 2002 conf. on Empirical Methods in Natural Language* (2002)
13. Rubin, V.L., Stanton, J.M., Liddy, E.D.: Discerning Emotions in Texts. *Procs. of AAAI Spring Symposium on Exploring Attitude and Affect in Text* (2004)
14. Taboada, M., Grieve, J.: Analyzing appraisal automatically. In *Procs. of AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pp. 158--161 (2004)

15. Turney, P.D.: Mining the Web for synonyms: PMI-IR versus LSA on TOEFL. *Procs. of the Twelfth European Conference on Machine Learning*, pp. 491--502 (2001)
16. Turney, P.D., Littman, M.: Unsupervised learning of semantic orientation from a hundred-billion-word corpus. Technical Report ERC-1094 (NRC 44929), National Research Council of Canada (2002)
17. Turney, P.D., Littman, M.L.: Measuring praise and criticism: Inference of semantic orientation from association, ACM Transactions on Information Systems (TOIS), vol. 21, no. 4, pp. 315--346 (2003)
18. Whitelaw, C., Navendu, G., Argamon, S.: Using Appraisal Groups for Sentiment Analysis. Procs. of the 14th ACM International Conference on Information and Knowledge Management, pp. 625--631 (2005)
19. Wiebe, J., Rapaport, W.: A Computational Theory of Perspective and Reference in Narrative. Procs. of the Association for Computational Linguistics, pp. 131--138 (1988)
20. White, P.R.R: Un recorrido por la teoría de la valoración. (Translated by Elsa Ghio). In The Appraisal Website: Homepage. http://www.grammati cs.com/appraisal /. [consulted:30/02/09] (up-to-date 2005)
21. Wiebe, J.: Recognizing Subjective Sentences: A Computational Investigation of Narrative Text. PhD Thesis. State Univ. of New York at Buffalo (1990)
22. Wilson, T., Wiebe, J., Hoffmann, P: Recognizing contextual polarity in phrase-level sentiment analysis". In Procs. of 200 conf. on Empirical Methods in Natural Language (2005)