International Journal For Technological Research In Engineering
Volume 5, Issue 9, May-2018 *ISSN (Online): 2347 - 4718*

SENTIMENT ANALYSIS ON TWITTER

T.Adithya[1], Cyndiya Priyadrashini[2]
[1]M.E(CSC), [2]M.Tech., (Ph.D)
Sri Ramanujar Engineering College, Chennai -127

Abstract: Depression is a global health concern. Social networks allow the affected population to share their experiences. Social media provides limitless opportunities to share experiences with their best suggestion. In current scenarios and with available new technologies, twitter can be used effectively for gathering information rather than gathering information in traditional method. Twitter is a most popular online social networking service that enable user to share and gain knowledge. This enabled us to accurately represent user interactions by relying on the data's semantic content. Preprocessed tweets are stored in database and those tweets are identified and classified whether it is user keywords related post using Support Vector Machine classification. The user keywords can be predicted whether it is a best suggestion using polarity. To provide an interactive automatic system which predicts the sentiment of the review/tweets of the people posted in social media. This system deals with the challenges that appear in the process of Sentiment Analysis, real time tweets are considered as they are rich sources of data for opinion mining and sentiment analysis. The main objective of this system is to perform real time sentimental analysis on the tweets that are extracted from the twitter and provide time based analytics to the user.
Index Terms: Negation scope, Sentiment Analysis, Twitter, Spanish Opinion Mining, Polarity classification, Lexicon based system, Statistical analysis

I. INTRODUCTION

Sentiment analysis is computational study of opinions, sentiments, evaluations, attitudes, appraisal, affects, views, emotions, subjectivity, etc., expressed in text. Sometimes called opinion mining. Customer feedback from emails, call centers etc. Opinions in news articles and commentaries ‡ Personal experiences and opinions about anything in reviews, forums, blogs, Twitter, micro-blogs, etc Comments about articles, issues, topics, reviews, etc. Postings at social networking sites, e.g., Facebook. Businesses spend a huge amount of money to find consumer opinions using consultants, surveys and focus groups, etc Make decisions to purchase products or to use services Find public opinions about political candidates and issues. Although at the beginning, most research has been ori-ented towards analyzing the sentiments in forums or web sites like Amazon or Epinion, the use of social networks as a huge source of data for SA is becoming more and more important. Specifically, micro-bloggings such as Twitter are being used to measure voting intention, consumer opin-ions and people's moods. In the last years, the number of scientific papers combining SA and Twitter has increased exponentially. However, most of this research is oriented to documents/tweets written in English, perhaps due to the novelty of the task and the lack of resources in other languages. Nonetheless people increasingly comment on their experiences, opinions, and points of view not only in English but in many other languages. Consequently, the management and study of subjectivity and SA in languages other than English is a growing need. The work presented herein is focused on polarity classification of Spanish tweets.
On the other hand, although polarity classification is the most widely studied task in SA, several challenges still remain open and are attracting the attention of researchers. One of these is the treatment of some linguistic phenomena such as irony, metaphors or negation. In this paper we focus on the treatment of negation. Actually, our main goal is to statistically demonstrate whether the detection and integration of negation in polarity classifier of Spanish tweets can improve the accuracy of the final system. To this end we first study the different cues that work as triggers of negativity. Then we define several rules in order to detect the scope of negation. Finally, we use our unsupervised lexicon based system to classify the polarity of a tweet. We demonstrate, by carrying out a statistical significance study, that the detection of negation and the application of some heuristic rules can significantly improve the final system.
The rest of the paper is organized as follows. The next section outlines existing research that has served as the basis of our work. Then the main resources used are described. Later, we introduce a section to study the scope of nega-tion. Section 5 describes the architecture of the proposed unsupervised approach. Next, the set of experiments that we have carried out are specified and analyzed. Moreover, we include a specific section to statistically demonstrate the validity of our approach. Finally, we conclude our study and present future directions for research.

II. RELATED WORK

Our main goal is to demonstrate the usefulness of consider-ing negation in a polarity classification system over a corpus of tweets. Thus, we will present some studies that have been the basis of our work. First we talk about how Twitter is considered one of the main sources of opinions that can be exploited by the SA community, paying special attention to papers dealing with negation and Spanish. Then, we de-scribe some other unsupervised systems based on lexicons because our model uses this kind of linguistic resource. Finally, a review of different papers studying negation is presented.

Sentiment Analysis on Twitter

The research community of SA was one of the first to become aware of the potential of Twitter as a great source of information to extract and generate knowledge from the data that users post [1]. Perhaps the first work related to the study of user opinions in this social network was presented by [2]. The authors developed a supervised system with the aim of analyzing the most suitable lexical features to represent a tweet and the most acceptable machine-learning algorithm to identify the polarity of the tweet. After this study a wide range of methods for SA on Twitter have been published, describing systems with different features and methodologies including supervised systems [3], [4], unsupervised approaches [5] and hybrid methods [6], [7]. However, few papers explicitly focus on negation or Spanish texts and only a few studies take into account features related to these issues. In [8] the performance of the unsupervised version of the algorithm SentiStrength is evaluated. The system combines several opinion lexicons of words and idioms. The system also incorporates a straight-forward method for detecting the scope of negation cues that does not invert the polarity of a word, but makes the word neutral. The authors conclude that their lexicon-based proposal is suitable for polarity classification in Twitter. In [9] a supervised polarity classification system for English tweets is described. The system follows the same approach for defining negation as [10], that is, all subsequent words to a negative particle are considered as negated words. In [11] a novel method is shown for calculating the polarity of a given tweet following an unsupervised approach. Montejo-Raez´ makes use of the semantic resource WeFeelFine [12], which is a huge database of sentences related to feelings and emotions. The sentiment system is similar to a search engine, where each tweet is treated as a query and the system returns the hundred emotions most similar to the given tweet. The final polarity is obtained by a weighted sum of the polar score of those hundred most similar emo-tions. The same approach is applied to Spanish reviews with promising results in [13]. Finally, in [14] an experimental study exploring the lexical and syntactic information in Spanish tweets in order to improve a polarity classification system is presented. The architecture is composed of differ-ent modules including a methodology to identify the scope of negation using a few negation terms and considering only four dependency-based rules. The experimental framework used to evaluate the systems is the corpus supplied by the TASS2013 organizers [15]. Our work is very close to this approach, but we consider more rules and more particles to detect the scope of negation. In addition, in this paper we carry out a statistical study in order to demonstrate the benefits obtained from applying these rules.

Lexicon based systems

In this paper we present a system that follows a lexicon-based strategy, so in the following lines we expound some papers related to this approach. For example, in [16] a lexicon-based method is used to take advantage of WordNet in building a lexicon of opinion bearing words. The polarity lexicon is used in conjunction with a sentiment lexicon of hashtags and a module for the identification of the scope of negation in order to develop a Twitter SA system in the political domain. When a lexical-based method is selected to build a polarity classifier, the building of a new list of opinion bearing words is not mandatory, the use of an existing lexicon is possible. This is the case of the paper [17], where the authors use the opinion lexicon General Inquirer to classify both subjectivity and polarity. In order to follow a lexicon based method, a list of opinion bearing words is needed. Three Spanish opinion lexicons are the most well-known by the SA research com-munity. In 2012 [18] was published, wherein the authors describe a framework that generates sentiment lexicons in a target language by using manually and automatically anno-tated English resources. The target language in the paper is Spanish, so the authors built two Spanish opinion lexicons, one from a manually labelled English opinion lexicon and another from an automatically labelled English opinion lexicon. Despite its recent publication, the opinion lexicon of Perez-Rosas et al. is being used in some studies such as [19] and [20]. Another interesting lexicon is described in [21], where the authors present a dictionary marked with probabilities to express one of the six basic emotions. The dictionary, which is known as the Spanish Emotion Lexicon (SEL), contains 2,036 words. Due to the fact that each word of the lexicon is labelled with an emotion and not with a polar label, the lexicon is less used by the SA research community. The third opinion lexicon is iSOL, which is described in [22] and has also been used successfully in [23] and [24]. iSOL is the lexicon used for determining the polarity of the system presented in this paper, and it will be described in the next section.

Negation and Sentiment Analysis

Regarding the treatment of negation, most research has focused on opinions written in English. One of the first approaches was proposed in [25] using a simple method that adds “NOT” to the terms of the sentence that appear next to negative terms, such as “no” or “don’t”. In [10] the same approach is followed, but they assume that the negation cues (“not”, “isn’t”, “didn’t”, etc.) affect all the terms from the cue to the end of the sentence. The authors carry out different experiments with and without negation using machine learning algorithms. However, the results show no significant differences considering negation or not. In [26] not only is negation considered but they also study intensifiers and diminishers, introducing the new concept “Contextual valence shifters”. In [27] a similar methodology is used where negations are used to reverse the semantic polarity of a particular term, while intensifiers and dimin-ishers are used to increase and decrease, respectively, the degree to which a term is positive or negative. In addition, in [28] an unsupervised model is proposed based on a fixed window of 4 words to determine negation scope. Other current researchers are developing rule-based systems using syntactic dependence trees [29] or applying more complex calculations in order to obtain polarity in opinions [30]. All these studies deal with English texts. There are even some good surveys about the study of negation as a linguis-tic phenomenon [31] and concerning SA [32]. However, for Spanish SA it is very difficult to find research considering negation as a feature. In [33] the same approach as the one used for English is applied, but adapted to Spanish. Thus, using their SO-CAL tool [34] they evaluate several nega-tion cues and calculate the polarity values depending on different features related to the terms and the grammatical category. Finally, in [35] the syntactic structure of the text is considered, showing an improvement over the systems that only use lexical features. Their recent work [14] shows some interesting results over the Spanish corpus of tweets supplied in the TASS2013 workshop [15]. However, they do not make any analysis of the gain obtained using negation individually, and so it is not possible to determine which is the module responsible for the improvement obtained.

III. RESOURCES

Increasingly, linguistic resources are becoming key players in NLP systems because they are the source of knowledge needed by NLP systems to achieve their primary objective, which is the understanding of natural language. Further-more, linguistic resources are necessary due to the fact that the performance and the quality of NLP systems have to be assessed. Therefore, two kinds of linguistic resources can be distinguished: The first ones are mainly employed as an essential element to building NLP systems, and the second ones are tools for evaluating such systems. The present paper describes a study in which the two sorts of linguistic resources are used with the aim of showing the importance of taking into consideration negation in the context of polarity classification on Twitter in Spanish. The polarity classification system developed for the study follows a lexicon-based approach, so some sets of sentiment-bearing expressions have been employed. Specif-ically, we consider a list of opinion words, a set of emoticons separated by the sentiments represented by them, and a list of hashtags that express sentiment. In addition, a corpus of Spanish tweets is necessary for the assessment. Currently, two corpora of Spanish tweets are available for the research community. The first one is the corpus used in the TASS workshop [15], and the second one is the Corpus Of Spanish Tweets COST[1] [36]. In this paper we have chosen the TASS corpus for several reasons. Firstly, the TASS corpus is broadly known by the Spanish research community, due mainly to the fact that it has been used in the previous four editions of the TASS workshop; the TASS corpus, which has about 68,000 tweets, contains considerably more tweets than the COST corpus, which is only composed of 34,634 tweets; and finally the TASS corpus was labelled following a semi-automatic process while the COST was labelled following a noisy label approach, which is similar to the one employed in [2].

iSOL lexicon

Although Spanish SA is attracting more and more re-searchers, the number of opinion lexicons is scarce com-pared to the ones available in English. For English SA we can find several resources such as the opinion lexicon compiled by Bing Liu [37], the MPQA lexicon [28], General Inquirer [38], SentiWordNet [39] and so on. However, for Spanish the number of resources is limited. In this paper, we have used the iSOL lexicon because it has been successfully applied in other studies. iSOL is a Spanish lexicon composed of 8,135 opinion words (2,509 positive words and 5,626 negative words). This resource was created taking as a basis the list of opinion words compiled by Bing Liu, which was translated into Spanish. Subsequently, the translated version of the list was manually reviewed and it was completed with more Spanish terms in order to obtain a more representative list of Spanish opinion words. All the details of the compilation process of iSOL can be found thoroughly described in [22]. The evaluation of iSOL demonstrates its validity for sentiment analysis in Spanish.

Hashtags, emoticons and laughs

The language used in Twitter has two special elements that are constantly typed by users, mentions and hashtags. A mention is the explicit reference that a user makes to another through writing the username preceded by the @ symbol. A hashtag is a string preceded by the hash key (#), and it is usually employed in order to identify the main topic, the sentiment or the semantic orientation of the tweet. Thus, taking into consideration hashtags in the process of polarity classification of tweets in Spanish could be a good idea. In [40] the effect of hashtagging emotions such as joy, sadness, anger and surprise in order to express the general emotion or sentiment in a tweet is studied. In a later paper [9], the authors describe the compilation of a lexicon of opinion using hashtags in English. To our knowledge a lexicon of Spanish opinion hashtags is not available, so the compilation of a Spanish opinion hashtag lexicon was undertaken. For this, we used a seed of positive hashtags (#bueno (#good), #bien (#well), #positivo (#positive), #fantastico (#great), #excelente (#excellent), etc.) and another of negative hashtags (#malo (#bad), #mal (#bad), #terrible (#terrible), #negativo (#negative), #horrible (#horrible), etc.) and retrieved all the tweets that had any of the seed words for three days. Then, we extracted all the hashtags present in those tweets and classified them as positive or negative depending on whether they appeared in the same tweet of a positive or negative seed. Finally, we manually reviewed these hashtags in order to obtain the final lists. In this way, the hashtags lexicon[2] was compiled and it is composed of 172 positive and 127 negative hashtags. Emoticons are other indicators of polarity that should be taken into account. In [41] it was shown that when the author of an electronic communication uses an emoticon, he/she is effectively marking up the text with an emotional state. In [2] emoticons are used to build one of the first corpus of tweets for SA. In [36] emoticons have also been used to compile a corpus of positive and negative tweets written in Spanish. According to the emotions itemized in Wikipedia[3], two lists of emoticons were generated[4]: one of them with 70 positive emoticons and another one with 46 negative emoticons. Laughs are another element frequently used in Twitter. For identifying them we have defined a regular expression with the main forms of writing laughs in Spanish and variants thereof: jajaja, jaaajajaj, jijiji, jijiij, lol, loool, etc.

The TASS corpus

In order to evaluate our proposal we have used a corpus widely known by the Spanish SA research community, called General Corpus of TASS[5] [15]. It was published for the first time in 2012 and since then it has been used in all the subsequent editions of the workshop on SA at SEPLN (2013, 2014, 2015 and 2016), so up until now it is the main corpus of Spanish tweets tagged for SA. The corpus contains over 68,000 tweets gathered between November 2011 and March 2012. The tweets were written in Spanish by about 150 well-known personalities and celebrities of the world of politics, economy, communication, mass media and culture. The corpus is divided into two sets: training (10%) and test (90%), so the training set is composed of 7,219 tweets and the test one is formed by 60,017 tweets. Each tweet in both sets is tagged with its global polarity, indicating whether the text expresses a positive, negative or neutral sentiment, or no sentiment at all. Five levels have been defined: strong positive (P+), positive (P), neutral (NEU), negative (N), strong negative (N+) and one additional no sentiment tag (NONE). We consider the TASS corpus has become a benchmark for Spanish SA on Twitter. Thus, we think it is a good choice for our experiments. Because our system is completely unsupervised and does not require training data, only the test set of the TASS corpus was taken into consideration for the assessment of the proposal. In addition, we neglected the tweets tagged with NONE class and only considered Positive, Negative and Neutral classes. Thus, original strong positive (P+) and positive (P) tweets are grouped into one unique positive class (P). Alike, strong negative (N+) and negative (N) are considered as negative class (N). After all this processing, the final set of tweets used for the assess-ment is composed of 22,233 positive tweets, 1,305 tweets labelled as neutral, and 15,844 negative tweets, which is a total of 39,381 tweets.

IV. NEGATION SCOPE IDENTIFICATION

Negation is an important feature of language that requires a special treatment in the field of NLP and specifically in SA. It is considered a challenging task because it is a linguistic phenomenon that has not been studied enough, especially in Spanish. The present paper is oriented towards the study of this challenge for SA: identification of the scope of negation in Spanish texts. Our main goal is to demonstrate whether by taking into account negation we can improve the polarity classification of Spanish tweets. We think that a correct iden-tification of the negation scope could help in the polarity classification of a text because a negative opinion can be ex-pressed using positive words negated (e.g. No fue una buena idea asistir al concierto (It was not a good idea to go to the concert)) or, by contrast, a positive opinion can be expressed from the negation of negative words (e.g. “La actuacin no fue un desastre como se esperaba”/ The performance was not a disaster as expected). As a first approach to this phenomenon, we propose a set of rules based on dependency trees for identifying the scope of some negation cues. In particular, we have studied the most important according to La Real Academia Espanola˜ (Royal Spanish Academy) [42]: no (not), tampoco (neither), nadie (nobody), jam´as (never), ni (nor), sin (without), nada (nothing), nunca (never) and ninguno (none). For each nega-tion cue, a rule for determining its scope was defined. For this, we analyzed the dependency trees of diverse sentences extracted from different websites in which some of the cues considered appear. To build the dependency trees we used the dependency parser of Freeling [43], which generates the dependency tree of a sentence based on its syntactic structure. Freeling[6] [44] is an open-source language-analysis toolkit that is available for several languages, including Spanish. After the study of these trees, we realized that it is possible to generalize the treatment of these negation cues in 3 rules (Table 1). We analyzed, on average, ten dependency trees per negation cue. The dependency trees produced by Freeling were always coherent with the rules, so we decided to continue research and apply them to Spanish Twitter SA. Although we think that the use of a specialized parser is better for the processing of tweets, we also support the idea that while specialized parsers are not available, a standard parser can be used. To the best of our knowledge, there is no specific parser for Spanish tweets so the NLP tool most used for Spanish (Freeling) was chosen. Moreover, due to the fact that tweets are informal texts, we apply a spelling checker in order to keep the number of errors as low as possible and to make the dependency parser work successfully (see Section 5).

TABLE 1: Rules for Identifying the Scope of Negation Cues

Cue	Rule for scope identification
no (not), tampoco (neither), nadie (nobody), jam´as (never), ninguno (none)	Parent node and the tree formed by the brother of the right, in-cluded
ni (nor), sin (without)	All children and all trees formed by them until reaching leaf nodes
nada (nothing), nunca (never)	Parent node

In order to clarify the rules that have been defined, an example of the applications of each rule is shown in Fig. 1, Fig. 3 and Fig. 2. Each figure represents the dependency tree related to a tweet in which the negation cue is represented with an ellipse and its scope is marked with a box.

The integration of these rules in a polarity classification system allows us to tag the words that are in the scope of any negation cue, with the aim of taking this into account when the polarity of a tweet is determined. For example, in the tweet Han actuado sin defensa ni garant´ıas para los usuarios. (They have acted without defense nor guarantees for the users.) (Fig. 2), the system will detect that there are two negative particles in the text, sin (without) and ni (nor), and for each one it will determine its scope using the rules defined. In this case, both particles affect all children nodes and all trees formed by them until reaching leaf nodes.

Therefore, the words defensa (defense) and garant´ıas (guarantees) will be tagged as negated words in order to modify their polarity value.

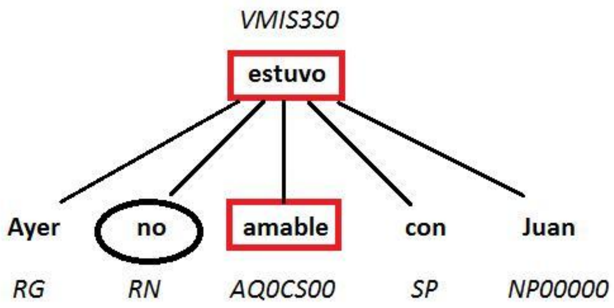

Fig. 1. Dependency tree of the negative word no (not). Tweet: Ayer no estuvo amable con... (Yesterday he was not kind to...).

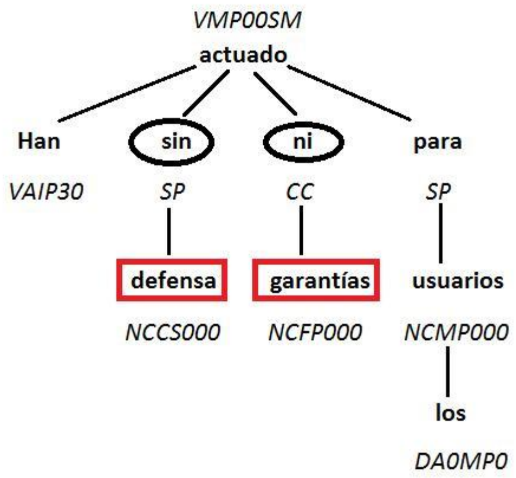

Fig. 2. Dependency tree of the negative words sin (without) and ni (nor). Tweet: Han actuado sin defensa ni garant´ıas para los usuarios (They have acted without defense nor guarantees for the users).

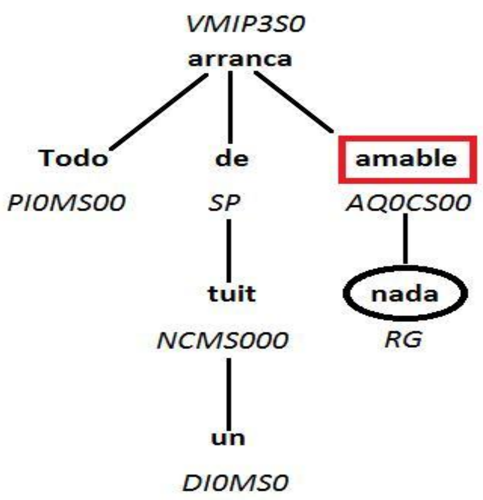

Fig. 3. Dependency tree of the negative word nada (nothing). Tweet: Todo arranca de un tuit nada amable (It all starts with an unkind tweet).

V. SYSTEM ARCHITECTURE

As we have mentioned earlier, the aim of this study is to demonstrate that taking into account negation is useful in a polarity classification system of tweets. To verify this assertion, we propose an unsupervised lexicon based system made up of different components. The main contribution of this system is the development of a normalization module that corrects misspelled words, another that detects the presence of a negation cue in a tweet and determines its scope using the rules defined, and the compilation of a Spanish opinion hashtags lexicon. The approach used for determining the polarity of a tweet is straightforward be-cause our goal is not focused on demonstrating that our system is a good polarity classifier but showing that treat-ment of negation is useful in such systems. The processing of each tweet to obtain a final polarity classification can be summarized in five steps:

- Tokenize the tweet.
- Correct misspelled words.
- Determine the part of speech of each word and the lemma of each verb.
- Detect the presence of negation cues and identify the scope of each of them using the rules defined.
- Obtain the polarity of the tweet.

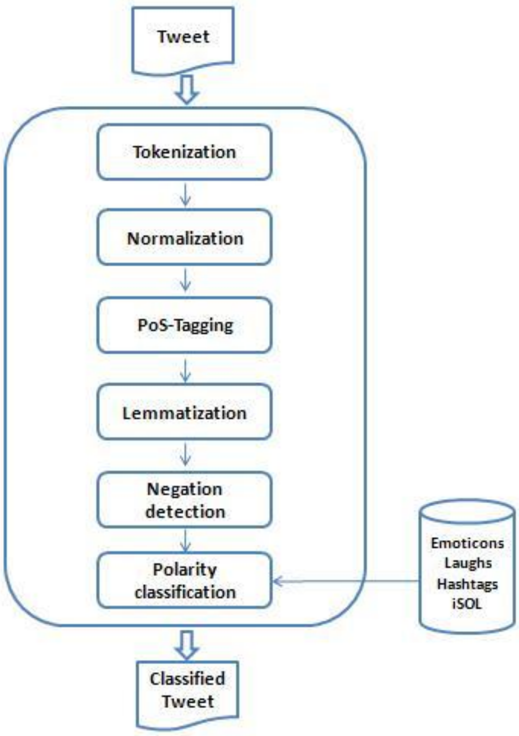

Fig. 4. Architecture of the polarity classification system.

The process outlined is shown in Figure 4. Below, a de-tailed explanation of all elements is shown with the sample tweet Todo arranca de un tweet nada amaaable. #maldad =((It all starts with an unkind tweet. #wickedness = ().

Tokenization: In order to process the text in the tweet, sentence splitting and word tokenization have to be performed. For this, the Freeling splitter and an adapted version to the Spanish language of the Christopher Potts'

tokenizer[7] were used. The tokenizer developed takes into account all special features of the language used in Spanish tweets: emoticons, urls, mentions, hashtags, dates, multi-words, etc. Below, the tokens that the system would identify in the sample tweet are shown in square brackets:
[Todo] [arranca] [de] [un] [tweet] [nada] [amaaable] [.] [#maldad] [=(]

Normalization: After the identification of the to-kens, the next step is to perform a normalization process in order to correct all misspelled words and to mark the tokens that have repeated letters. We mark the tokens that have repeated letters to con-sider their intensity when we calculate the overall sentiment of the tweet. The two reasons for per-forming normalization to correct spelling errors are, firstly, that our system needs to build the syntactic tree of each tweet, so if there are fewer misspellings in the text the dependency parser will be more likely
to be successful. The second reason is that the sys-tem is based on the use of the lexical resource iSOL which is a list of words, most of them well written. The spelling corrector of Peter Norvig[8] has been modified with the aim of correcting misspellings in Spanish texts. This spelling corrector only needs to work a large corpus in the target language. In our case, the target language is Spanish, so we have to compile a representative corpus of Spanish. This large corpus is composed of a list of Spanish lemmas, a list of Spanish verb conjugations and a list of Spanish names and surnames. All the lists were compiled by Ismael Olea[9]. The initial lists were complemented by the list of words of the corpus CREA[10], which was compiled by La Real Academia Espanola˜ (Royal Spanish Academy). Nor-malization of the sample tweet would correct the token [amaaable] and would also mark it as a token with repeated letters:
[Todo] [arranca] [de] [un] [tweet] [nada] [amable] [.] [#maldad] [=(]
Repeated letters

PoS-Tagging and Lemmatization: The third step is to learn the PoS-tag of each token in order to obtain the lemma of each verb, because iSOL does not have all the verbal forms of polar verbs, it only has the lemma of each one. Therefore, we used the Part-of-Speech tagger module of Freeling. This resource has two different modules for performing PoS tagging [44]. The first one is the hmm tagger which is a classical trigram Markovian tagger [45] and the second one, named the relax tagger, is a hybrid system capable of integrating statistical and hand-coded knowledge [46]. We used the hmm tagger because it is faster than the relax tagger. In the case of the sample tweet, the system would tag each token with its pertinent part of speech and would obtain the lemma of the token [arranca] because it is a verbal form.
[Todo] [arrancar] [de] [un] [tweet] [nada] [amable] [.] [#maldad] [=(]
Repeated letters

Negation detection: This module, in the first place, detects whether the tweet has any negation cue and if so, it determines the scope of each cue with the set of syntactic rules that has been defined (Table 1). In this way, if a tweet has a negation cue the system will generate its dependency parser and will mark each word affected by negation as "negated" by "name of the cue", in order to take this into account when the semantic orientation of the tweet is calculated. In the sample tweet there is a negation cue, the token [nada]. In this case, the system would generate the dependency tree of the tweet (Figure 5) and would mark as negated by [nada] the token [amable] that is in its scope according to the rule defined.

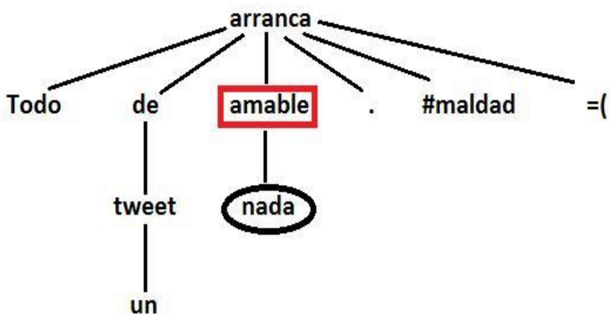

Fig. 5. Dependency tree of the tweet: Todo arranca de un tweet nada amable. #maldad =((It all starts with an unkind tweet. #wickedness = (.)
[Todo] [arrancar] [de] [un] [tweet] [nada] [amable] [.] [#maldad] [=(]
Repeated letters Negated by nada

Polarity classification: The last step is to determine the polarity of the tweet. For this purpose, a polarity classifier that takes into account the presence of emoticons, hashtags, expressions of laughing and negation was developed. This component uses the resources described in Section 3: the bag of words of emoticons tagged as positives and negatives, the bag of hashtags and iSOL lexicon. For each tweet the classifier determines its positivity and negativity value. Thus, if a token is in the bag of positive/negative emoticons, a polarity value of 2 is added to its positivity/negativity value. If it detects that a token is an expression of laughing, the positivity value is increased by 2. In the other case, if the token is in the bag of positive/negative hashtags, the counter of positivity/negativity is in-creased by 2. Finally, if the token is in the iSOL positive/negative list, a polarity value of 1 is added to the positivity/negativity counter and if it also has repeated letters the value is increased by 1. If the token is negated its polarity is reversed (positive ! negative, negative ! positive). Using these values, the system is able to classify the tweet in one of the 3 defined classes following the equation:

$$\text{Macro } F1 = \frac{1}{|c|} \sum_{j=1}^{|c|} \frac{2P_iR_i}{P_i + R_i} \qquad (6)$$

In the same way, we can obtain the Macro-Recall and Macro-Precision as follows:

$$\text{Macro } R = \frac{1}{|C|} \sum_{i=1}^{|C|} \frac{TP_i}{TP_i + FN_i} \quad (7)$$

$$\text{Macro } P = \frac{1}{|C|} \sum_{i=1}^{|C|} \frac{TP_i}{TP_i + FP_i} \quad (8)$$

Results

After these clarifications, the results achieved in the experiments with the Total set are shown in Table 3.

TABLE 3: Results Total Set

	Macro-P	Macro-R	Macro-F1	Accuracy	Improvement Accuracy
BS	0.5764	0.5235	0.5486	0.6258	-
BSN	0.5705	0.5190	0.5435	0.6205	-0.885%
NR	0.5810	0.5296	0.5541	0.6308	0.80%

Note: The improvement in the Accuracy is measured over the BS method.

It can been seen that the integration of the most common approach to detect the scope of negation in English tweets (BSN), [48]) does not work well in the system that we use for the polarity classification of Spanish tweets. On the other hand, when the rules-based approach that we propose is included (NR), there is an improvement, but perhaps it seems that it is not so significant. However, if we observe the confusion matrix of the experiments (Table 4 and Table 5) we can see that there is a difference of about 200 tweets which have been correctly classified with the NR experiment.

TABLE 4: Confusion Matrix BS Experiment with Total Set

	Predicted P	Predicted NEU	Predicted N	Recall
Real P	16,476	4,768	989	0.7410
Real NEU	511	446	348	0.3418
Real N	2,758	5,360	7,726	0.4876
Precision	0.8344	0.0422	0.8525	

TABLE 5: Confusion Matrix NR Experiment with Total Set

	Predicted P	Predicted NEU	Predicted N	Recall
Real P	16,566	4,746	921	0.7451
Real NEU	511	457	337	0.3502
Real N	2,685	5,341	7,818	0.4934
Precision	0.8383	0.0433	0.8614	

As we can see, RuleAffect is a subset of NegCue and NegCue is a subset of Total. The reason why the dataset has been reduced to carry out experiments with two subsets is because most of the tweets in the corpus do not have negation cues (Table 2), and so in order to determine whether negation improves the polarity classification we need to compare the subsets with and without negation cues.

TABLE 2: Tweets Used in the Experiments

	Tweets	Percentage
Total	39,382	100%
NegCue	8,604	22%
RuleAffect	2,326	6%

As we have mentioned earlier, in order to evaluate the rules defined we should pay attention to the tweets with negative particles (NegCue) and mainly to the tweets with polar tokens in the scope of negation (RuleAffected). Table 6 and Table 7 show the results obtained using these subsets.

TABLE 6: Results Neg Cue Set

	Macro-P	Macro-R	Macro-F1	Accuracy	Improvement Accuracy
BS	0.4861	0.4702	0.4780	0.4866	-
BSN	0.4621	0.4514	0.4567	0.4622	-5.01%
NR	0.5060	0.4936	0.4997	0.5092	4.64%

Note: The improvement in the Accuracy is measured over the BS method.

Results Rule Affect Set

	Macro-P	Macro-R	Macro-F1	Accuracy	Improvement Accuracy
BS	0.3971	0.3949	0.3960	0.4463	-
BSN	0.4431	0.4545	0.4487	0.5026	12.61%
NR	0.4660	0.4792	0.4725	0.5292	18.57%

Note: The improvement in the Accuracy is measured over the BS method.

As was to be expected, the values of the evaluation measures are lower than using the Total set (Table 3) because these subsets contain the most problematic tweets, i.e. the tweets with negation cues that are the most difficult to classify. In addition to these tweets, the Total set has other tweets without negation cues that are easier to classify, meaning that precision and recall increase. However, the improvement obtained with the rule-based approach is more evident. Furthermore, according to the results, it is reasserted the fact that the method most used to determine the scope of negation in English tweets (BSN) does not classify better in our system than the method that we propose for Spanish tweets (NR). The results achieved with the RuleAffect subset show the evaluation of the rules that we have presented in this paper. Of course, the rules are not perfect and can be improved in order to increase accuracy. However, there is apparently a significant differ-ence between BS and NR because as we can see there is an improvement of 18.57% in the accuracy and 19.19% in the F1 measure. Therefore, to avoid wrong conclusions we will perform a statistical analysis to check whether the rules defined for the treatment of negation really do improve the classification.

As was to be expected, the values of the evaluation measures are lower than using the Total set (Table 3) because these subsets contain the most problematic tweets, i.e. the tweets with negation cues that are the most difficult to classify. In addition to these tweets, the Total set has other tweets without negation cues that are easier to classify, meaning that precision and recall increase. However, the improvement obtained with the rule-based approach is more evident. Furthermore, according to the results, it is reasserted the fact that the method most used to determine the scope of negation in English tweets (BSN) does not classify better in our system than the method that we propose for Spanish tweets

(NR). The results achieved with the RuleAffect subset show the evaluation of the rules that we have presented in this paper. Of course, the rules are not perfect and can be improved in order to increase accuracy. However, there is apparently a significant differ-ence between BS and NR because as we can see there is an improvement of 18.57% in the accuracy and 19.19% in the F1 measure. Therefore, to avoid wrong conclusions we will perform a statistical analysis to check whether the rules defined for the treatment of negation really do improve the classification.

VI. CONCLUSIONS AND FURTHER WORKS

Negation is a linguistic phenomenon that can change the meaning of a sentence, so its treatment can influence posi-tively in the performance of NLP tasks like SA. In this study, we have presented a set of syntactic rules for determining the scope of negation in Spanish. We have integrated these rules into a polarity classification system of Spanish tweets and it has been demonstrated that the results obtained with them are significantly greater than those without taking into account negation. This rule-based approach has also been compared with the method most used to determine the scope of negation in English tweets, and it has been proved that the classification with our approach is better. Moreover, we have analyzed the rules defined showing the performance of them with each negation cue. The results obtained encourage us to follow in the study of the correct treatment of negation in the context of SA. However, one of the main problems in this area is the lack of resources. For example, there is no labeled corpus including negation for Spanish. Thus, we are currently working on the annotation of negation cues and their scope [53] in the Spanish version of the SFU corpus [33] in order to evaluate the rules with the aim of checking whether the system correctly determines the scope of the negation cues studied or if some of the errors are caused by the polarity classifier used. Moreover, we will also study in which cases the polarity of a word that is within the scope of negation should be swapped, considered neutral or if its value should be increased or decreased or, by contrast, whether it should not be changed.

REFERANCE

[1] http://sinai.ujaen.es/cost
[2] http://sinai.ujaen.es/hashtags-sp
[3] http://es.wikipedia.org/wiki/Anexo:Emoticonos
[4] http://sinai.ujaen.es/emoti-sp
[5] http://www.daedalus.es/TASS2014/tass2014.phpn#corpus
[6] http://nlp.lsi.upc.edu/freeling/

