CSL: A Combined Spanish Lexicon
Resource for Polarity Classification and Sentiment Analysis

Luis G. Moreno-Sandoval[1], Paola Beltrán-Herrera[1], Jaime A. Vargas-Cruz[1], Carolina Sánchez-Barriga[1], Alexandra Pomares-Quimbaya[2], Jorge A. Alvarado-Valencia[3] and Juan C. García-Díaz[1]

[1]*Colombian Center of Excellence and Appropriation on Big Data and Data Analytics (CAOBA), Colombia*
[2]*Department of Systems Engineering, Pontificia Universidad Javeriana, Bogotá, D.C., Colombia*
[3]*Department of Industrial Engineering, Pontificia Universidad Javeriana, Bogotá, D.C., Colombia*

Keywords: Spanish Lexicon, Spanish Resources for Sentiment Analysis, Polarity's Classification, Opinion Mining, Sentiment Analysis.

Abstract: Opinion mining and sentiment analysis in texts from social networks such as Twitter has taken great importance during the last decade. Quality lexicons for the sentiment analysis task are easily found in languages such as English; however, this is not the case in Spanish. For this reason, we propose CSL, a Combined Spanish Lexicon approach for sentiment analysis that uses an ensemble of six lexicons in Spanish and a weighted bag of words strategy. In order to build CSL we used 68,019 tweets previously classified by researchers at the Spanish Society of Natural Language Processing (SEPLN) obtaining a precision of 62.05 and a recall of 60.75 in the validation set, showing improvements in both measurements. Additionally, we compare the results of CSL with a very well-known commercial software for sentiment analysis in Spanish finding an improvement of 10 points in precision and 15 points in recall.

1 INTRODUCTION

Evolution of social media through the last two decades has led to a growing number of activities by which people express their opinions about any kind of issues on the web (Feldman, 2013). As a consequence, increasing use of Internet has led to the possibility of extracting and analyzing a huge amount of structured and unstructured data. Interests of research communities have increased specially in extracting and analyzing views and moods expressed on these global platforms through sentiment analysis, also known as opinion mining (Taboada, 2016).

According to (Ravi and Ravi, 2015) there are several important sub-tasks to be performed for sentiment analysis. One of these subtasks is lexicon-based polarity determination for the sentiment classification. Lexicon-based approaches can use dictionaries, which are a collection of opinion words along with their positive (*+ve*) or negative (*-ve*) sentiment strength. Studies for determining polarity of a sentence expressed in social networks as Twitter using a lexicon-based approach have shown the necessity of high level pre-processing because of the common presence of abbreviations and stop words, as well as the increase of precision with different techniques (Ravi and Ravi, 2015). The literature review showed that the number of studies involving Spanish Lexicons is significantly fewer than in English (García-Moya et al., 2013; Molina-González et al., 2013; Montejo-Ráez et al., 2014; Ortigosa et al., 2014).

In line with (Molina-González et al., 2013) there are two main ways for addressing the problem of applying sentiment analysis to non-English languages: by generating corpora, dictionaries and lists of opinion words or by translation.

Considering the previous facts, this paper is focused on polarity determination for sentiment classification in Spanish, as more research in this language is needed and it is important to advance in tools for those interested in working with this language.

This paper presents a combined lexicon for Sentiment Analysis called CSL (Combined Spanish Lexicon). CSL effectiveness was developed and tested through an unsupervised method developed in

Moreno-Sandoval, L., Beltrán-Herrera, P., Vargas-Cruz, J., Sánchez-Barriga, C., Pomares-Quimbaya, A., Alvarado-Valencia, J. and García-Díaz, J.
CSL: A Combined Spanish Lexicon - Resource for Polarity Classification and Sentiment Analysis.
DOI: 10.5220/0006336402880295
In *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017) - Volume 1*, pages 288-295
ISBN: 978-989-758-247-9

Python that uses the Bag of Words approach for Sentiment Analysis of short texts. The principle is to enhance the recognition of important sentimental words using multiple lexicons and an ensemble function that takes into account the quality of polarity classification of the words available in each lexicon. Thus, strengths of available lexicons are enhanced and weaknesses are diminished.

We describe the assembling process of six previously developed lexicons and present the results of the experiments with the obtained lexicon by comparing their effectiveness measures with two references: the classification made by the Spanish Society of Natural Language Processing and the results obtained using the IBM Solution called Bluemix. The validation yielded good results of precision (62.05), recall (60.75) and F1 Score (61.39).

The paper is organized as follows: the second part briefly describes previous related work on available Spanish lexicons and the studies regarding to lexicon-based polarity determination. In the third section we explain the methodology used to assemble the Spanish lexicon CSL. Section four presents the experiments carried out with the new lexicon with the purpose of proving its efficiency and discusses the main results obtained. Finally, we outline conclusions and further work.

2 RELATED WORK

Sentiment analysis, also known as opinion mining, is the task of detecting, extracting and classifying opinions, sentiments and attitudes concerning different topics, as expressed in textual input (Montoyo et al., 2012). The most relevant reviews of opinion, sentiment and subjectivity in text (Montoyo et al., 2012; Pang and Lee, 2009; Ravi and Ravi, 2015; Tang and Liu, 2010) highlight that a relevant task in sentiment analysis is the so-called Polarity classification or determination.

(Pang and Lee, 2009) state that the goal of a large portion of work in sentiment related to classification / regression /ranking is to classify an opinionated text unit or topic, as positive or negative, or locate its position on the continuum between these two polarities.

(Ravi and Ravi, 2015) explains that sentiment classification can be performed using machine learning, which yields better precision, or using lexicon – based approaches, which provide more generality because of their semantic orientation. Lexicon-based approaches can use dictionaries or corpora.

(Martinez-Camara et al., 2014) integrated the iSOL SWN_SP lexicons to classify opinion polarity in a Spanish review corpus. They applied each dictionary separately and combined both results through stacking meta-classification. For the stacking approach they used Support Vector Machines, Naïve Bayes and Bayesian Logistic Regression. For testing, they used the *MuchoCine* Spanish corpus. The results showed that the combination of different lexicons, with the use of meta-classifiers improve the performance of polarity classification for Spanish texts. Later, this methodology was applied on the combination of iSOL and ML-SentiCon, obtaining similar results.

(Jiménez-Zafra et al., 2016) enhanced a lexicon adapting it to a specific domain, by adding polar adjectives obtained through Term Frequency and Bootstrapping. They applied both techniques to the iSOL lexicon. The results showed that polarity classification in movie reviews was significantly improved with respect to those originally achieved with iSOL. This shows that properly augmenting a dictionary can improve polarity classification in texts.

(Taboada et al., 2011) introduced the Semantic Orientation Calculator (SO-CAL). They created separate adjective, noun, verb and adverb dictionaries, and hand ranked them using a -5 to +5 scale indicating the degree of orientation of a given word; they also incorporated numeric values for intensifiers, negations, and irrealis markers, resulting in a formula to calculate the semantic orientation of a given text. The lexicon used in the tests was manually created. Several versions of the lexicon were generated to test the calculator with four data sets. Their results indicated an advantage in creating hand-ranked dictionaries for lexicon based sentiment analysis.

As a conclusion, methods for lexicon combination and augmentation have proved to be useful in sentiment analysis. Some of these methods combine results from individual classifiers, with a meta-classifier. In contrast, in this paper, lexicon improvement was obtained directly by combining several dictionaries.

2.1 Available Spanish Lexicons

An extensive research of lexicons shows that the developments in Spanish are fewer with respect to the developments in English. However, the literature review led us to find ten (10) lexicons in Spanish classified by some polarity's category: (i). iSOL (González et al., 2015b; Martinez-Camara et al., 2014), (ii). SentiWordNet (SWN) (Baccianella et al., 2010; Esuli and Sebastiani, 2006; González et al., 2015a; Princeton University, 2015; SentiWordNet, 2010), (iii). Multilingual Central Repository (MCR) (González et al., 2015a), (iv). EuroWordNet (EWN) (EuroWordNet, 2001), (v). ElhPolar (Saralegi et al., 2013; Saralegi and San Vicente, 2013), (vi). Spanish Emotion Lexicon (SEL), (Díaz Rangel et al., 2014; Sidorov et al., 2012, 2013), (vii). Political Dictionary (PL) (Alvarado-Valencia et al., 2016), (viii). Sentiment Lexicons in Spanish (SLS) (Pérez-Rosas et al., 2012), (xi). ML-SentiCon (Fe.L. Cruz et al., 2014; Fermín L. Cruz et al., 2014) and (x). Multilingual Sentiment (MS) (Data Science Lab, 2014). Table 1 compares the features of the aforementioned lexicons.

Notice that only eight (8) of the ten (10) analyzed lexicons are available for academic use. SWN and EWN require payment and use of specialized software for their consultation. Also, MCR and EWN do not refer to polarity classifications, but rather, they refer to existing relationships between words such as a taxonomy or ontology structures. On the other hand, PL is the only lexicon containing words from the political knowledge domain which implies disjoints in the polarity's classification of existing words. This is the case of the Spanish word "*carrusel*", which in the political context has a negative connotation, but in the general context has a positive connotation. Finally, with respect to the methodology of consolidation of each available lexicon, it can be evidenced that the majority of lexicons correspond to automatic and translation processes, but the validation process mainly corresponds to a manual one.

Additionally, the polarity standardization requires pre-processing of the available data in order to get the same polarity categories in each lexicon.

3 METHODOLOGY

This section presents the main characteristics of CSL, our approach for improving the precision and recall in sentiment analysis. The first section describes the strategy applied during the assembly of independent lexicons, and the second one explains the algorithm created to identify and to extract sentiment information.

3.1 Pre-processing

Six dictionaries were selected to be used in the ensemble, using a qualitative approach where ease of access was the most important criterion. Final lexicons selected were: iSOL, Elh Polar, SEL SLS, Ml-SentiCon and MS.

Some preliminary cleaning procedures were performed on the original lexicons. Repeated words were found in some lexicons which were eliminated when they had the same qualification. Additionally, Spanish phrases like "*a la moda*", "*a pesar de*", "*acoger con agrado*", were eliminated from the lexicon.

SEL provides the probability of expressing one of the following six emotions: joy, surprise, anger, fear, repulsion and sadness. Following the methodology proposed by (Molina-González et al., 2013), joy and surprise emotions were associated with positive polarity and the other emotions were classified as negative, considering that the probability of expressing each emotion was greater or equal than

Table 1: Comparison between available Spanish Lexicons.

	iSOL	SWN	MCR	EWN	Elh Polar	SEL	PL	SLS	ML-SentiCon	MS
Tot. of words:	8,135	117,000	NA	NA	5,199	2,036	1,638	3,843	11,542	4,275
Tot. categories:	2	3	NA	NA	2	6	3	2	3	2
Polarity range:	-1 and 1	[-1, 1]	-	-	-1 y 1	-	[-1, 1]	-1 and 1	[-1, 1]	-1 and 1
Tot. positive words:	2,509	NA	-	-	3,302	-	356	1,332	955	1,555
Tot. negative words:	5,626	NA	-	-	1,897	-	260	2,511	1,300	2,720
Tot. neutral words:	-	NA	-	-	-	-	1,022	-	9,287	-
Knowledge domain:	G	G	G	G	G	G	P	G	G	G
Consolidation's Met.	Auto.	Man.	NA	NA	Auto.	Auto.	Man.	Semi Auto.	Man.	Auto.
Consolidation's Ty:	Transl.	Transl.	NA	NA	Transl.	Transl.	Own	Own	Transl.	Transl.
Validation Process:	Man.	Man	NA	NA	Man.	Man.	Man.	Man.	Man	Auto.
Measurement:	A and R	NA	NA	NA	A.	KC.	A and R.	A and R.	NA	NA
Citation/References:	29	663	103	62	NA	14	NA	21	NA	NA
Academic Availab.:	X	NA	X	NA	X	X	X	X	X	X

0.3. Words with lower emotional probability were eliminated. This threshold was decided empirically and let to discard 20% of the terms.

3.2 Lexicon Individual Performance Test

The individual performance of each dictionary was tested. The test consisted in qualifying as Positive, Neutral or Negative a group of tweets taken from (Villena-Román et al., 2013) which were manually classified by a group of experts. This set of tweets (68,019) were divided into two, training and validation sets, with 89% and 11% of the total tweets, respectively.

In order to use the same categories, tweets classified by TASS (Villena-Román et al., 2013) as neutral or "none" were grouped in the same category. This grouping helped to achieve balanced training and test sets. Table 2 shows the distribution of the different polarities as classified by TASS.

Table 2: Distribution of polarity.

Polarity	Training	Test
Positive:	37%	40%
Neutral:	37%	30%
Negative:	26%	30%

For testing the lexicons, a methodology based on Bag of Words with stemming and stop word elimination was used. Additionally, accent marks and special characters were also removed to increase matching. Other features as intensifiers and detection of double negatives, which are often used in NLP in Spanish (Vilares et al., 2015) were not included. During this testing task, the approach keeps in a log file the sentimental words recognized during the process and the number of documents correctly and wrongly classified where that word was found. After this process, each one of the words, for every lexicon, is weighted using a score defined as (1)

$$\text{Score} = \text{Correct Classifications} / (\text{Correct Classifications} + \text{Wrong Classifications}) \tag{1}$$

The average F1 Score, which has been used by other authors in sentiment analysis (Councill et al., 2010; Saif et al., 2012), was used to measure the performance of each dictionary, but other performance metrics as precision and recall were also calculated.

3.3 Supervised Enrichment of Polarity Lexicons

This building process is illustrated in the pseudocode shows below:

```
begin
    1. T as tweets;
    2. L as the initial lexicons;
    3. define S as a temporal array
       of lexicons;
    4. define CSL as a new lexicon;
    5. for each lexicon l_i in L do:
    6.  for each word w_j in l_i do:
    7.   counter = 0
    8.   value = 0
    9.   for each tweet t_k in T
         do:
    10.    if w_j exists in t_k then:
    11.   counter = counter + 1
    12.      if polarity(l_i[w_j]) =
             polarity(t_k)then:
    13.        value = polarity(t_k) +
               value
    14.      end if;
    15.    end if;
    16.   end for;
    17.   score = value / counter;
    18.   if score <= -0.4 or score
          >= 0.4 then:
    19.     insert (w_j,
            round(score,0)) in S[i]
    20.   end if;
    21.  end for;
    22. end for;
    23. for each word w_i in S do:
    24.   CSL(w_i) = {w_i,
                round(average
                [polarity(S_1(w_i)),
                …, polarity(S_6(w_i))]
                ,0)};
    25. end for;
    26. Return CSL;
End.
```

The new lexicon, which selected the best words from every initial lexicon, was used with the Bag of Words approach, using the test partition as corpus. Afterwards, the performance measures were calculated using the manual classification on each tweet.

Likewise, other strategies were used to create additional ensembles by changing the conditions used to assign the polarities to the words in the new lexicons (pseudocode: from line 23 to line 25). CLS_1 changed the use of average for the use of maximum value, while CLS_2 used a threshold value. The three lexicons showed in this paper were the ones with a strategy that achieved a precision and a F1 score that outperform those from the initial lexicons.

This approach presents the possibility to choose the best words of each lexicon and to discard those which have an inadequate classification or a low contribution to the identification of the tweet's polarity.

3.4 Comparison Against a Commercial Software

A comparison was also made against the Alchemy Language, a collection of Applications Program Interface (API) by IBM's Watson that offers text analyses thought NLP (IBM, 2016a). The API used was called Sentiment, which "*can calculate overall sentiment within a document*" (IBM, 2016b), among other things, as sentiment for user-specified targets, entity-level sentiment, quotation-level sentiment and directional and keyword- level sentiment.

In this case, the overall sentiment within a document was used, defining an individual tweet as a document. Also, for this test a total of 2,764 tweets were taken from the test partition, this number was limited by technical restrictions in the available calls to the API.

The calls were made from a subroutine in Python that uses a client library developed by IBM Watson. The parameters used were the text of the tweet and the language (Spanish).

Finally, in a similar way used to test the Bag of Words implementation with each lexicon, the results were compared against the manual classification done by the SEPLN.

4 RESULT VALIDATION

This section is divided into three parts: results of the individual pre-processing of each selected lexicon, analysis of the ensemble lexicon and results of the ensemble validation.

4.1 Individual Analysis

Lexicons once pre-processed, as described in the methodology, are compared to 89% of previously classified tweets (Villena-Román et al., 2013). Table 3 shows the percentage of identified words of each lexicon that were used to assign polarity to the set of tweets and the ratio of this words to the total words in each lexicon; this ratio seeks to establish a measure of performance with respect to lexicon size. According to the results, MLSentiCon, MS and Elh Polar present the highest percentages of identified words, but for this, only ElhPolar and MS have high ratios. MLSentiCon and iSOL, having a high percentage of identified words, show ratios of 15.8 and 12.4 identified words by each word of the lexicon respectively. This behavior is emphasized because these are the lexicons with greater number of words and have domain of general knowledge. On the other hand, two lexicons identified less than 6% of words, but show high ratios.

Table 3: Performance of Spanish lexicons.

Lexicon	% Identified words	Identified words/Total lexicon words
iSOL:	12.5	12.4
Elh Polar:	15.3	32.0
SEL:	5.7	29.8
SLS:	5.9	35.8
ML-SentiCon:	17.4	15.8
MS:	17.2	32.7

Additionally, a comparison between the words positively and negatively identified by each lexicon is established. As can be seen in Figure 1 positive ratings range from 56.5% to 73%.

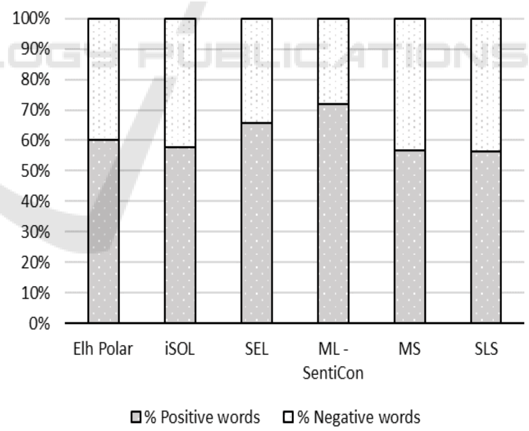

Figure 1: Percentage of positive and negative identified words by each lexicon.

Finally, precision and recall measures for each lexicon are calculated. Table 4 presents the mean of precision, recall, and F1 Score. Notice that this was a three-category classification task. The last measure will be used later in the ensemble process as described in the methodology. ElhPolar stands out with the best indicators. This shows that the methodology followed by (Saralegi et al., 2013) of eliminating conflicting words did a great work at improving the predicting power of the lexicon.

Table 4: Precision, recall and F1 Score from each lexicon.

Lexicon	Precision (Mean)	Recall (Mean)	F1 Score
iSOL:	54.87	54.83	54.85
Elh Polar:	59.64	59.76	59.70
SEL:	55.72	51.98	53.78
SLS:	51.98	49.45	50.69
ML-SentiCon:	47.09	44.92	45.98
MS:	53.33	53.96	53.64

4.2 Ensemble Lexicon

For the three ensemble exercises developed here, Table 5 presents the total of resulting positive and negative loaded words and their comparison to the individual lexicons presented in section 4.1. The number of words for the ensembles is a result of the process of selecting words that adequately classify the tweets according to those described in the methodology.

The last column of Table 5 presents the accuracy obtained with the test data set, quantifying the correct classification of each lexicon with respect to the total. It is evident that CLS_1, CLS_2 and CLS_3 ensemble exercises have the highest values. This table shows that the accuracy achieved by all of the lexicons was better than the expected from a naïve classification, where the accuracy would be close to a 33% in a perfectly balanced dataset.

Table 5: Positive and negative loaded words in each lexicon.

Lexicon	# Positive words	# Negative words	Accuracy
iSOL:	2,509	5,624	55.26
Elh Polar:	1,379	2,502	59.95
SEL:	631	931	54.33
SLS:	477	870	50.83
ML-SentiCon:	4,453	4,482	46.99
MS:	1,553	2,720	53.99
CLS_1:	1,901	1,910	60.66
CLS_2:	1,970	1,945	60.73
CLS_3:	11,634	3,305	62.38

4.3 Combined Lexicon Validation

For the combined lexicon validation, the test data set with 11% of the tweets was used. The exercise was performed for all lexicons individually and for the three ensembles developed to ensure comparability in the results. Table 6 presents the measures of precision, recall and F1 Score.

According to the results, the three proposed ensembles surpass the individual performance of lexicons, indicating that this procedure results in a more efficient lexicon.

Table 6: Precision, recall and F1 Score for each ensemble.

Lexicon	Precision (Mean)	Recall (Mean)	F1 Score
CLS_1:	59.09	58.59	58.84
CLS_2:	59.04	58.62	58.83
CLS_3:	62.05	60.75	61.39

The improvement in precision with respect to the CLS_3 for each lexicon individually is presented in Table 7. According to the results for every individual lexicon there is an improvement in the F1 Score. The lexicon that performs best is ElhPolar, which results in a 4.65% improvement.

Table 7. Precision, recall, F1 Score and improvement for each lexicon.

	Precision (Mean)	Recall (Mean)	F1 Score	Imp. F1 Score (%)
iSOL:	53.39	52.70	53.04	15.75
Elh Polar:	59.22	58.11	58.66	4.65
SEL:	52.95	49.24	51.03	20.31
SLS:	50.90	48.05	49.43	24.20
ML-SentiCon:	47.02	44.48	45.71	34.30
MS:	50.32	50.17	50.25	22.18

As an additional benchmark for the development of this exercise, the IBM Bluemix tool was used to make comparisons regarding the performance of individual lexicons and the ensemble lexicon. The comparison of our approach was made using 2,217 tweets selected randomly from the set of previously classified tweets (Villena-Román et al., 2013). The precision and recall obtained by the commercial software were 51.97 and 45.58 respectively, with F1 Score of 48.57. These results show that although in English this tool is frequently used, for Spanish, it is possible to improve their algorithms with the language resources proposed in this paper.

5 CONCLUSIONS AND FUTURE WORK

Creating a single model which integrates different lexicon approaches has several benefits. Principally, the predictive overall precision, recall and F1 Score of the new lexicon is significantly better than lexicons individually evaluated providing researchers with a tool that integrates the potentialities of individual lexicons. This result might be due to the methodology to select the right polarity explained in section 3.3, which improves strengths and reduces weaknesses of individual lexicons.

To achieve these results, it was necessary to pre-process qualitatively and quantitatively the lexicons available in Spanish in order to review the quality of the polarity classification previously made in each of them. Likewise, choosing as a gold standard the manually classified tweets and, from this, verifying the polarity assigned to the word, allowed to improve the quality of the available polarity classification. However, it is important to emphasize that performance depends not only on the initial lexicons, but of the way in which they are used.

The algorithm used by IBM to calculate the polarity of a text could be improved by using the methodology developed in this paper. It is important to take into account that the algorithm was only tested in short texts (tweets), but given the growing importance of analyzing tweets and other short length opinions, the authors believe that this could greatly improve the performance of the alchemy API.

Furthermore, ensembles could be tested into models that consider entity recognition, negation handling and other sophistications, which would help understand the performance of assembling with different algorithms.

Moreover, to attest the generality of ensembles in diverse contexts, the test could be done with a different gold standard, like movies or items reviews.

ACKNOWLEDGEMENTS

This research was carried out by the Center of Excellence and Appropriation in Big Data and Data Analytics (CAOBA). It is being led by the Pontificia Universidad Javeriana Colombia and it was funded by the Ministry of Information Technologies and Telecommunications of the Republic of Colombia (MinTIC) through the Colombian Administrative Department of Science, Technology and Innovation (COLCIENCIAS) within contract No. FP44842-anex46-2015.

REFERENCES

Alvarado-Valencia, J.A., Carrillo, A., Forero, J., Caicedo, L., Urueña, J.C., 2016. Análisis de sentimiento político en twitter para las elecciones de la Alcaldía de Bogotá 2.015. Presented at the XXVI Simposio Internacional de Estadística 2016, Sincelejo, Sucre, Colombia.

Baccianella, S., Esuli, A., Sebastiani, F., 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining., in: LREC. pp. 2200–2204.

Councill, I.G., McDonald, R., Velikovich, L., 2010. What's Great and What's Not: Learning to Classify the Scope of Negation for Improved Sentiment Analysis, in: Proceedings of the Workshop on Negation and Speculation in Natural Language Processing, NeSp-NLP '10. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 51–59.

Cruz, F.L., Troyano, J. a., Pontes, B., Ortega, F. j., 2014. ML-SentiCon: A multilingual, lemma-level sentiment lexicon. Proces. Leng. Nat. 53, 113–120.

Cruz, F.L., Troyano, J.A., Pontes, B., Ortega, F.J., 2014. Building layered, multilingual sentiment lexicons at synset and lemma levels.

Data Science Lab, 2014. Multilingualsentiment [WWW Document]. URL https://sites.google.com/site/datascienceslab/projects/multilingualsentiment (accessed 6.27.16).

Díaz Rangel, I., Sidorov, G., Suárez Guerra, S., 2014. Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. Onomázein Rev. Lingüíst. Filol. Trad. 29, 31–46. doi:10.7764/onomazein.29.5

Esuli, A., Sebastiani, F., 2006. Sentiwordnet: A publicly available lexical resource for opinion mining, in: Proceedings of LREC. Citeseer, pp. 417–422.

EuroWordNet, 2001. EuroWordNet:Building a multilingual database with wordnets for several European languages. [WWW Document]. URL http://www.illc.uva.nl/EuroWordNet/ (accessed 5.23.16).

Feldman, R., 2013. Techniques and Applications for Sentiment Analysis. Commun. ACM 56, 82–89. doi:10.1145/2436256.2436274

García-Moya, I., Moreno, C., Rivera, F., 2013. Sense of Coherence and Biopsychosocial Health in Spanish Adolescents. Span. J. Psychol. 16. doi:10.1017/sjp.2013.90

González, M.D.M., Cámara, E.M., Martín-Valdivia, M.T., López, L.A.U., 2015a. A Spanish semantic

orientation approach to domain adaptation for polarity classification. ResearchGate 51, 520–531. doi:10.1016/j.ipm.2014.10.002
González, M.D.M., Cámara, E.M., Valdivia, M.T.M., 2015b. CRiSOL: Base de Conocimiento de Opiniones para el Español. Proces. Leng. Nat. 55, 143–150.
IBM, 2016a. AlchemyLanguage | IBM Watson Developer Cloud [WWW Document]. URL https://www.ibm.com/watson/developercloud/alchemy-language.html (accessed 12.19.16).
IBM, 2016b. AlchemyLanguage Service Documentation | Watson Developer Cloud [WWW Document]. URL https://www.ibm.com/watson/developercloud/doc/alchemylanguage/ (accessed 12.19.16).
Jiménez-Zafra, S.M., Martin, M., González, M.D.M., Lopez, L.A.U., 2016. Domain Adaptation of Polarity Lexicon combining Term Frequency and Bootstrapping, in: ResearchGate. Presented at the Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pp. 137–146. doi:10.18653/v1/W16-0422
Martinez-Camara, E., Martin-Valdivia, M., Molina-Gonzalez, M., Perea-Ortega, J., 2014. Integrating Spanish lexical resources by meta-classifiers for polarity classification. J. Inf. Sci. 40, 538–554.
Molina-González, M.D., Martínez-Cámara, E., Martín-Valdivia, M.-T., Perea-Ortega, J.M., 2013. Semantic orientation for polarity classification in Spanish reviews. Expert Syst. Appl. 40, 7250–7257. doi:10.1016/j.eswa.2013.06.076
Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M.T., Ureña-López, L.A., 2014. Ranked WordNet graph for Sentiment Polarity Classification in Twitter. Comput. Speech Lang. 28, 93–107. doi:10.1016/j.csl.2013.04.001
Montoyo, A., Martinez-Barco, P., Balahur, A., 2012. Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. Decis. SUPPORT Syst. 53, 675–679.
Ortigosa, A., Martín, J.M., Carro, R.M., 2014. Sentiment analysis in Facebook and its application to e-learning. Comput. Hum. Behav. 31, 527–541. doi:10.1016/j.chb.2013.05.024
Pang, B., Lee, L., 2009. Opinion mining and sentiment analysis. Comput. Linguist. 35, 311–312. doi:10.1162/coli.2009.35.2.311
Pérez-Rosas, V., Banea, C., Mihalcea, R., 2012. Learning Sentiment Lexicons in Spanish 5.
Princeton University, 2015. About WordNet - WordNet - About WordNet [WWW Document]. URL https://wordnet.princeton.edu/ (accessed 5.23.16).
Ravi, K., Ravi, V., 2015. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowl.-BASED Syst. 89, 14–46.
Saif, H., He, Y., Alani, H., 2012. Semantic sentiment analysis of twitter, Lecture Notes in Computer Science.
Saralegi, X., San Vicente, I., 2013. Elhuyar at TASS 2013, in: Workshop on Sentiment Analysis at SEPLN (TASS2013). Presented at the XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural, Madrid, pp. 143–150.
Saralegi, X., San Vicente, I., Ugarteburu, I., 2013. Cross-lingual projections vs. corpora extracted subjectivity lexicons for less-resourced languages, Lecture Notes in Computer Science.
SentiWordNet, 2010. SentiWordNet [WWW Document]. URL http://sentiwordnet.isti.cnr.it/ (accessed 5.23.16).
Sidorov, G.(1), Miranda-Jiménez, S.(1), Viveros-Jiménez, F.(1), Gelbukh, A.(1), Castro-Sánchez, N.(1), Velásquez, F.(1), Díaz-Rangel, I.(1), Suárez-Guerra, S.(1), Treviño, A.(2), Gordon, J.(2), 2013. Empirical study of machine learning based approach for opinion mining in tweets, Lecture Notes in Computer Science.
Sidorov, G., Miranda-Jiménez, S., Viveros-Jiménez, F., Gelbukh, A., Castro-Sánchez, N., Velásquez, F., Díaz-Rangel, I., Suárez-Guerra, S., Treviño, A., Gordon, J., 2012. Empirical Study of Opinion Mining in Spanish Tweets 1–14.
Taboada, M., 2016. Sentiment Analysis: An Overview from Linguistics. Annu. Rev. Linguist. 2, 325.
Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M., 2011. Lexicon-Based Methods for Sentiment Analysis. Comput. Linguist. 37, 267–307.
Tang, L., Liu, H., 2010. Community Detection and Mining in Social Media.
Vilares, D., Alonso, M.A., Gómez-Rodríguez, C., 2015. A syntactic approach for opinion mining on Spanish reviews. Nat. Lang. Eng. 21, 139–163. doi:10.1017/S1351324913000181
Villena-Román, J.(1), Lana-Serrano, S.(2), Martínez-Cámara, E.(3), González-Cristóbal, J. c. (4), 2013. TASS - Workshop on sentiment analysis at SEPLN. Proces. Leng. Nat. 50, 37–44.