

# Exploiting Emotive Features for the Sentiment Polarity Classification of tweets

Lucia C. Passaro, Alessandro Bondielli and Alessandro Lenci

CoLing Lab, Dipartimento di Filologia, Letteratura e Linguistica  
University of Pisa (Italy)

lucia.passaro@for.unipi.it  
alessandro.bondielli@gmail.com  
alessandro.lenci@unipi.it

## Abstract

**English.** This paper describes the CoLing Lab system for the participation in the constrained run of the EVALITA 2016 SENTiment POLarity Classification Task (Barbieri et al., 2016). The system extends the approach in (Passaro et al., 2014) with emotive features extracted from ItEM (Passaro et al., 2015; Passaro and Lenci, 2016) and FB-NEWS15 (Passaro et al., 2016).

**Italiano.** *Questo articolo descrive il sistema sviluppato all'interno del CoLing Lab per la partecipazione al task di EVALITA 2016 SENTiment POLarity Classification Task (Barbieri et al., 2016). Il sistema estende l'approccio descritto in (Passaro et al., 2014) con una serie di features emotive estratte da ItEM (Passaro et al., 2015; Passaro and Lenci, 2016) and FB-NEWS15 (Passaro et al., 2016).*

## 1 Introduction

Social media and microblogging services are extensively used for rather different purposes, from news reading to news spreading, from entertainment to marketing. As a consequence, the study of how sentiments and emotions are expressed in such platforms, and the development of methods to automatically identify them, has emerged as a great area of interest in the Natural Language Processing Community. Twitter presents many linguistic and communicative peculiarities. A tweet, in fact, is a short informal text (140 characters), in which the frequency of creative punctuation,

emoticons, slang, specific terminology, abbreviations, links and hashtags is higher than in other domains and platforms. Twitter users post messages from many different media, including their smartphones, and they “tweet” about a great variety of topics, unlike what can be observed in other sites, which appear to be tailored to a specific group of topics (Go et al., 2009).

The paper is organized as follows: Section 2 describes the architecture of the system, as well as the pre-processing and the features designed in (Passaro et al., 2014). Section 3 shows the additional features extracted from emotive VSM and from LDA. Section 4 shows the classification paradigm, and the last sections are left for results and conclusions.

## 2 Description of the system

The system extends the approach in (Passaro et al., 2014) with emotive features extracted from ItEM (Passaro et al., 2015; Passaro and Lenci, 2016) and FB-NEWS15 (Passaro et al., 2016). The main goal of the work is to evaluate the contribution of a distributional affective resource to estimate the valence of words. The CoLing Lab system for polarity classification includes the following basic steps: (i) a preprocessing phase, to separate linguistic and nonlinguistic elements in the target tweets; (ii) a feature extraction phase, in which the relevant characteristics of the tweets are identified; (iii) a classification phase, based on a Support Vector Machine (SVM) classifier with a linear kernel.

### 2.1 Preprocessing

The aim of the preprocessing phase is the identification of the linguistic and nonlinguistic elements in the tweets and their annotation.

While the preprocessing of nonlinguistic elements such as links and emoticons is limited to their identification and classification (cf. section 2.2.4), the treatment of the linguistic material required the development of a dedicated rule-based procedure, whose output is a normalized text that is subsequently feed to a pipeline of general-purpose linguistic annotation tools. The following rules have been applied in the linguistic preprocessing phase:

- **Emphasis:** tokens presenting repeated characters like *bastaaaa* “stoooooop” are replaced by their most probable standardized forms (i.e. *basta* “stop”);
- **Links and emoticons:** they are identified and removed;
- **Punctuation:** linguistically irrelevant punctuation marks are removed;
- **Username:** the users cited in a tweet are identified and normalized by removing the @ symbol and capitalizing the entity name;
- **Hashtags:** they are identified and normalized by simply removing the # symbol;

The output of this phase are linguistically-standardized tweets, that are subsequently POS tagged with the Part-Of-Speech tagger described in (Dell’Orletta, 2009) and dependency-parsed with the DeSR parser (Attardi et al., 2009).

## 2.2 Feature extraction

The inventory of features can be organized into six classes. The five classes of features described in this section have been designed in 2014, the sixth class, described in the next section is referred to the emotive and LDA features.

### 2.2.1 Lexical Features

Lexical features represent the occurrence of bad words or of words that are either highly emotional or highly polarized. Relevant lemmas were identified from two in-house built lexicons (cf. below), and from Sentix (Basile and Nissim, 2013), a lexicon of sentiment-annotated Italian words. Lexical features include:

**ItEM seeds:** Lexicon of 347 highly emotional Italian words built by exploiting an online feature elicitation paradigm (Passaro et al.,

2015). The features are, for each emotion, the total count of strongly emotional tokens in each tweet.

**Bad words lexicon:** By exploiting an in house built lexicon of common Italian bad words, we reported, for each tweet, the frequency of bad words belonging to a selected list, as well as the total amount of these lemmas.

**Sentix:** Sentix (Sentiment Italian Lexicon: (Basile and Nissim, 2013)) is a lexicon for Sentiment Analysis in which 59,742 lemmas are annotated for their polarity and intensity, among other information. Polarity scores range from  $-1$  (totally negative) to  $1$  (totally positive), while Intensity scores range from  $0$  (totally neutral) to  $1$  (totally polarized). Both these scores appear informative for the classification, so that we derived, for each lemma, a Combined score  $C_{score}$  calculated as follows:

$$C_{score} = Intensity * Polarity \quad (1)$$

Depending on their  $C_{score}$ , the selected lemmas have been organized into several groups:

- strongly positives:  $1 \leq C_{score} < 0.25$
- weakly positives:  $0.25 \leq C_{score} < 0.125$
- neutrals:  $0.125 \leq C_{score} \leq -0.125$
- weakly negatives:  $-0.125 < C_{score} \leq -0.25$
- highly negatives:  $-0.25 < C_{score} \leq -1$

Since Sentix relies on WordNet sense distinctions, it is not uncommon for a lemma to be associated with more than one  $\langle Intensity, Polarity \rangle$  pair, and consequently to more than one  $C_{score}$ .

In order to handle this phenomenon, the lemmas have been splitted into three different ambiguity classes: Lemmas with only one entry or whose entries are all associated with the same  $C_{score}$  value, are marked as “Unambiguous” and associated with their  $C_{score}$ .

Ambiguous cases were treated by inspecting, for each lemma, the distribution of the associated  $C_{scores}$ : Lemmas which had a Majority Vote (MV) were marked as “Inferable” and associated with the  $C_{score}$  of the MV. If there was no MV, lemmas were marked as “Ambiguous” and associated with the mean of the  $C_{scores}$ . To isolate a reliable set of polarized words, we focused only on the Unambiguous or Inferable lemmas and selected only the

250 topmost frequent according to the PAIS corpus (Lyding et al., 2014), a large collection of Italian web texts.

Other Sentix-based features in the ColingLab model are: the number of tokens for each  $C_{score}$  group, the  $C_{score}$  of the first token in the tweet, the  $C_{score}$  of the last token in the tweet and the count of lemmas that are represented in Sentix.

### 2.2.2 Negation

Negation features have been developed to encode the presence of a negation and the morphosyntactic characteristics of its scope.

The inventory of negative lemmas (e.g. “non”) and patterns (e.g. “non ... mai”) have been extracted from (Renzi et al., 2001). The occurrences of these lemmas and structures have been counted and inserted as features to feed the classifier.

In order to characterize the scope of each negation, we used the dependency parsed tweets produced by DeSR (Attardi et al., 2009). The scope of a negative element is assumed to be its syntactic head or the predicative complement of its head, in the case the latter is a copula. Although it is clearly a simplifying assumption, the preliminary experiments show that this could be a rather cost-effective strategy in the analysis of linguistically simple texts like tweets.

This information has been included in the model by counting the number of negation patterns encountered in each tweet, where a negation pattern is composed by the PoS of the negated element plus the number of negative tokens depending from it and, in case it is covered by Sentix, either its Polarity, its Intensity and its  $C_{scores}$  value.

### 2.2.3 Morphological features

The linguistic annotation produced in the preprocessing phase has been exploited also in the population of the following morphological statistics: (i) number of sentences in the tweet; (ii) number of linguistic tokens; (iii) proportion of content words (nouns, adjectives, verbs and adverbs); (iv) number of tokens for Part-of-Speech.

### 2.2.4 Shallow features

This group of features has been developed to describe distinctive characteristics of web communication. The group includes:

**Emoticons:** We used the lexicon LexEmo to mark the most common emoticons, such as :- (

and :-), marked with their polarity score: 1 (positive), -1 (negative), 0 (neutral).

LexEmo is used both to identify emoticons and to annotate their polarity.

Emoticon-related features are the total amount of emoticons in the tweet, the polarity of each emoticon in sequential order and the polarity of each emoticon in reversed order. For instance, in the tweet :- (*quando ci vediamo? mi manchi anche tu!* :\*:\* “:- (when are we going to meet up? I miss you, too :\*:\*” there are three emoticons, the first of which (:-) is negative while the others are positive (:\*;\* :\*).

Accordingly, the classifier has been fed with the information that the polarity of the first emoticon is -1, that of the second emoticon is 1 and the same goes for the third emoticon. At the same way, another group of feature specifies that the polarity of the last emoticon is 1, as it goes for that of the last but one emoticon, while the last but two has a polarity score of -1.

**Links:** These features contain a shallow classification of links performed using simple regular expressions applied to URLs, to classify them as following: *video*, *images*, *social* and *other*. We also use as feature the absolute number of links for each tweet.

**Emphasis:** The features report on the number of emphasized tokens presenting repeated characters like *bastaaaa*, the average number of repeated characters in the tweet, and the cumulative number of repeated characters in the tweet.

**Creative Punctuation:** Sequences of contiguous punctuation characters, like !!!, !?!?!?!?!?!?!?!?! or . . . . ., are identified and classified as a sequence of dots, exclamations marks, question marks or mixed. For each tweet, the features correspond to the number of sequences belonging to each group and their average length in characters.

**Quotes:** The number of quotations in the tweet.

### 2.2.5 Twitter features

This group of features describes some Twitter-specific characteristics of the target tweets.

**Topic:** This information marks if a tweet has been retrieved via a specific political hashtag or keywords. It is provided by organizers as an attribute of the tweet;

**Usernames:** The number of @username in the tweet;

**Hashtags:** Hashtags play the role of organizing the tweets around a single topic, so that they are useful to be considered in determining their polarity (i.e. a tweet containing hashtags like *#amore* “#love” and *#felice* “#happy” is expected to be positive and a tweet containing hashtags like *#ansia* “#anxiety” and *#stressato* “#stressedout” is expected to be negative. This group of features registers the presence of an hashtag belonging to the list of the hashtags with a frequency higher than 1 in the training corpus.

### 3 Introducing emotive and LDA features

In order to add emotive features to the CoLing Lab model, we created an emotive lexicon from the corpus FB-NEWS15 (Passaro et al., 2016) following the strategy illustrated in (Passaro et al., 2015; Passaro and Lenci, 2016). The starting point is a set of seeds strongly associated to one or more emotions of a given taxonomy, that are used to build centroid distributional vectors representing the various emotions.

In order to build the distributional profiles of the words, we extracted the list  $T$  of the 30,000 most frequent nouns, verbs and adjectives from FB-NEWS15. The lemmas in  $T$  were subsequently used as target and contexts in a square matrix of co-occurrences extracted within a five word window ( $\pm 2$  words, centered on the target lemma). In addition, we extended the matrix to the nouns, adjectives and verbs in the corpus of tweets (i. e. lemmas not belonging to  $T$ ).

For each  $\langle$ emotion, PoS $\rangle$  pair we built a centroid vector from the vectors of the seeds belonging to that emotion and PoS, obtaining in total 24 centroids<sup>1</sup>. Starting from these spaces, several groups

<sup>1</sup>Following the configuration in (Passaro et al., 2015; Passaro and Lenci, 2016), the co-occurrence matrix has been re-weighted using the Pointwise Mutual Information (Church and Hanks, 1990), and in particular the Positive PMI (PPMI), in which negative scores are changed to zero (Niwa and Nitta, 1994). We constructed different word spaces according to PoS because the context that best captures the meaning of a word differs depending on the word to be represented (Rothenhusler and Schtze, 2007).

of features have been extracted. The simplest ones include general statistics such as the number of emotive words and the emotive score of a tweet. More sophisticated features are aimed at inferring the degree of distinctivity of a word as well as its polarity from their own emotive profile.

**Number of emotive words:** Words belonging to the emotive Facebook spaces;

**Emotive/words ratio:** The ratio between the number of emotive words and the total number of words in the tweet;

**Strongly emotive words:** Number of words having a high (greater than 0.4) emotive score for at least one emotion;

**Tweet emotive score:** Score calculated as the ratio between the number of strongly polarized words and the number of the content words in the tweet (Eq. 2). The feature assumes values in the interval [0, 1]. In absence of strongly emotive words, the default value is 0.

$$E(\textit{Tweet}) = \frac{\textit{Count}(\textit{Strongly emotive words})}{\textit{Count}(\textit{Content words})} \quad (2)$$

**Maximum values:** The maximum emotive value for each emotion (8 features);

**Quartiles:** The features take into account the distribution of the emotive words in the tweet. For each emotion, the list of the emotive words has been ordered according to the emotive scores and divided into quartiles (e.g. the fourth quartile contains the most emotive words and the first quartile the less emotive ones.). Each feature registers the count of the words belonging to the pair  $\langle$ emotion, quartile $\rangle$  (32 features in total);

**ItEM seeds:** Boolean features registering the presence of words belonging to the words used as seeds to build the vector space models. In particular, the features include the top 4 frequent words for each emotion (32 boolean features in total);

**Distinctive words:** 32 features corresponding to the top 4 distinctive words for each emotion. The degree of distinctivity of a word for a given emotion is calculated starting from the VSM normalized using Z-scores. In particular, the feature corresponds to the proportion

of the emotion  $\langle emotion_i \rangle$  against the sum of total emotion score  $[e_1, \dots, e_8]$ :

**Polarity (count):** The number of positive and negative words. The polarity of a word is calculated by applying Eq. 3, in which positive emotions are assumed to be JOY and TRUST, and negative emotions are assumed to be DISGUST, FEAR, ANGER and SADNESS.

$$Polarity(w) = \frac{JOY+TRUST}{2} - \frac{DISGUST+FEAR+ANGER+SADNESS}{4} \quad (3)$$

**Polarity (values):** The polarity (calculated using Eq. 3) of the emotive words in the tweet. The maximum number of emotive words is assumed to be 20;

**LDA features:** This group of features includes 50 features referred to the topic distribution of the tweet. The LDA model has been built on the FB-NEWS15 corpus (Passaro et al., 2016) which is organized into 50 clusters of thematically related news created with LDA (Blei et al., 2003) (Mallet implementation (McCallum, 2002)). Each feature refers to the association between the text of the tweet and a topic extracted from FB-NEWS15.

## 4 Classification

We used the same paradigm used in (Passaro et al., 2014). In particular, we chose to base the CoLing Lab system for polarity classification on the SVM classifier with a linear kernel implementation available in Weka (Witten and Frank, 2011), trained with the Sequential Minimal Optimization (SMO) algorithm introduced by Platt (Platt, 1999).

The classification task proposed by the organizers could be approached either by building two separate binary classifiers relying of two different models (one judging the positiveness of the tweet, the other judging its negativeness), or by developing a single multiclass classifier where the possible outcomes are Positive Polarity (Task POS:1, Task NEG:0), Negative Polarity (Task POS:0, Task NEG:1), Mixed Polarity (Task POS:1, Task NEG:1) and No Polarity (Task POS:0, Task NEG:0). In Evalita 2014 (Passaro et al., 2014) we tried both approaches in our development phase, and found no significant difference,

so that we opted for the more economical setting, i.e. the multiclass one.

## 5 Results

Although this model is not optimal according to the global ranking, if we focus on the recognition of the negative tweets (i.e. the NEG task), it ranks fifth (F1-score), and first if we consider the class 1 of the NEG task (i.e. NEG, F.sc. 1). Such trend is reversed if we consider the POS task, which is the worst performing class of this system.

Task	Class	Precision	Recall	F-score
POS	0	0,8548	0,7682	0,8092
POS	1	0,264	0,3892	0,3146
POS task		0,5594	0,5787	0,5619
NEG	0	0,7688	0,6488	0,7037
NEG	1	0,5509	0,6883	0,612
NEG task		0,65985	0,66855	0,6579
GLOBAL		0,609625	0,623625	0,6099

Table 1: System results.

Due to the great difference in terms of performance between the results obtained by performing a 10 fold cross validation, we suspected that the system was overfitting the training data, so that we performed different feature ablation experiments, in which we included only the lexical information derived from ItEM and FB-NEWS (i.e. we removed the features relying to Sentix, Negation and Hashtags (cf. table 2). The results demonstrate on one hand that significant improvements can be obtained by using lexical information, especially to recognize negative texts. On the other hand the results highlight the overfitting of the submitted model, probably due to the overlapping between Sentix and the emotive features.

Task	Class	Precision	Recall	F-score
POS	0	0,8518	0,8999	0,8752
POS	1	0,3629	0,267	0,3077
POS task		0,60735	0,58345	0,59145
NEG	0	0,8082	0,6065	0,693
NEG	1	0,5506	0,7701	0,6421
NEG task		0,6794	0,6883	0,66755
GLOBAL		0,643375	0,635875	0,6295

Table 2: System results for a filtered model.

The advantage of using only the lexical features derived from ItEM are the following: i) the emotional values of the words can be easily updated; ii) the VSM can be extended to increase the lexical coverage of the resource; iii) the system is “lean” (it can do more with less).

## 6 Conclusions

The Coling Lab system presented in 2014 (Passaro et al., 2014) has been enriched with emotive features derived from a distributional, corpus-based resource built from the social media corpus FB-NEWS15 (Passaro et al., 2016). In addition, the system exploits LDA features extracted from the same corpus. Additional experiments demonstrated that removing most of the non-distributional lexical features derived from Sentix, the performance can be improved. As a consequence, with a relatively low number of features the system reaches satisfactory performance, with top-scores in recognizing negative tweets.

## References

- Giuseppe Attardi, Felice Dell’Orletta, Maria Simi, and Joseph Turian. 2009. Accurate dependency parsing with a stacked multilayer perceptron. In *Proceedings of EVALITA 2009 Evaluation of NLP and Speech Tools for Italian 2009*, Reggio Emilia (Italy). Springer.
- Francesco Barbieri, Valerio Basile, Danilo Croce, Malvina Nissim, Nicole Novielli, and Viviana Patti. 2016. Overview of the EVALITA 2016 SENTiment POLarity Classification Task. In *Proceedings of EVALITA 2016 Evaluation of NLP and Speech Tools for Italian*, Napoli (Italy). Academia University Press.
- Valerio Basile and Malvina Nissim. 2013. Sentiment analysis on italian tweets. In *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 100–107, Atlanta.
- David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022.
- Kenneth W. Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16:22–29.
- Felice Dell’Orletta. 2009. Ensemble system for part-of-speech tagging. In *Proceedings of EVALITA 2009 Evaluation of NLP and Speech Tools for Italian 2009*, Reggio Emilia (Italy). Springer.
- Alec Go, Richa Bhayani, and Lei Huang. 2009. Twitter sentiment classification using distant supervision. *Processing*, pages 1–6.
- Verena Lyding, Egon Stemle, Claudia Borghetti, Marco Brunello, Sara Castagnoli, Felice Dell’Orletta, Henrik Dittmann, Alessandro Lenci, and Vito Pirrelli. 2014. The PAISÀ Corpus of Italian Web Texts. In *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, pages 36–43, Gothenburg (Sweden). Association for Computational Linguistics.
- Andrew K. McCallum. 2002. Mallet: A machine learning for language toolkit. <http://mallet.cs.umass.edu>.
- Yoshiki Niwa and Yoshihiko Nitta. 1994. Co-occurrence vectors from corpora vs. distance vectors from dictionaries. In *Proceedings of the 15th International Conference On Computational Linguistics*, pages 304–309, Kyoto (Japan).
- Lucia C. Passaro and Alessandro Lenci. 2016. Evaluating context selection strategies to build emotive vector space models. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Portoro (Slovenia). European Language Resources Association (ELRA).
- Lucia C. Passaro, Gianluca E. Leboni, Emmanuele Chersoni, and Alessandro Lenci. 2014. The coling lab system for sentiment polarity classification of tweets. In *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & and of the Fourth International Workshop EVALITA 2014*, pages 87–92, Pisa (Italy).
- Lucia C. Passaro, Laura Pollacci, and Alessandro Lenci. 2015. Item: A vector space model to bootstrap an italian emotive lexicon. In *Proceedings of the second Italian Conference on Computational Linguistics CLiC-it 2015*, pages 215–220, Trento (Italy).
- Lucia C. Passaro, Alessandro Bondielli, and Alessandro Lenci. 2016. Fb-news15: A topic-annotated facebook corpus for emotion detection and sentiment analysis. In *Proceedings of the Third Italian Conference on Computational Linguistics CLiC-it 2016.*, Napoli (Italy). To appear.
- John C. Platt, 1999. *Advances in Kernel Methods*, chapter Fast Training of Support Vector Machines Using Sequential Minimal Optimization, pages 185–208. MIT Press, Cambridge, MA, USA.
- Lorenzo Renzi, Giampaolo Salvi, and Anna Cardinaletti. 2001. *Grande grammatica italiana di consultazione*. Number v. 1. Il Mulino.
- Klaus Rothenhusler and Hinrich Schtze. 2007. Part of speech filtered word spaces. In *Sixth International and Interdisciplinary Conference on Modeling and Using Context*.
- Ian H. Witten and Eibe Frank. 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition.