Consumer Concern Extraction in Social Web Reviews

Joseph Lark, Sebastián Peña Saldarriaga, Emmanuel Morin, Fabien Poulard, Sylvain Ornetti

Consumer Concern Extraction in Social Web Reviews

J. Lark[1,2], S. Peña Saldarriaga[2], E. Morin[1], F. Poulard[2], and S. Ornetti[2]

[1] LINA, Université de Nantes, Nantes, France
joseph.lark@etu.univ-nantes.fr, emmanuel.morin@univ-nantes.fr
[2] Dictanova, Nantes, France
{sebastian,sornetti,fpoulard}@dictanova.com

Abstract. The main goal we adress in this work is text comprehension in social web through Natural Language Processing. The type of content we are interested in includes noisy sentences for the most part and thus standard topic extraction is generally not applicable. We present here a semi-automatic method based on machine learning and human defined linguistic features that performs points of interest extraction in consumer reviews. We first run a classification of subjective adjectives with a high precision and then use this extracted lexicon to identify co-occurring common nouns. While there is still room for improvement on several levels of our process, the results we obtain with this first method are very promising.

Keywords: domain-specific extraction, consumer concerns, social web reviews, opinion mining, subjectivity, nlp, social web research

1 Introduction

With the fast expanding use of social networks and in particular the popular tendency of consumers to share opinions about products on the social web, there is a growing interest for brands to keep track of the concerns users may have towards their business. According to a 2013 Weber Shandwick & KRC Research survey[3], consumer reviews appear as the third trigger in purchase decision of consumer electronics, behind price and product features and in front of professional critic reviews. In this paper we introduce concern extraction which is the process of ranking the key nouns that reflects consumer concerns in product reviews. This work is based on the assumption that such nouns frequently co-occur with subjective markers *e.g.* subjective adjectives, nouns or adverbs. In order to restrain our first results in a coherent frame of work, the only type of markers we consider in this work are adjectives. In addition to the subjectivity hurdle, the difficulty we try to overcome with our method is the quality of language. The challenges induced by this type of noisy content have led us to put our trust in predefined lexicons, and so the first step of this work is to semi-automatically expand a lexicon of subjective adjectives, in order to identify the key nouns that are qualified by these adjectives.

[3] http://www.webershandwick.com/uploads/news/files/ReviewsSurveyReportFINAL.pdf

2 Related Work

To the best of our knowledge, concern extraction from social web corpora has not been addressed, more specifically in the case of French language. However, related work on similar subjects gave us some leads as to what type of methods we should consider. In particular, in [2] authors describe a subjectivity detection method in noisy data using grammatical patterns, in [3] named entity recognition (NER) is also performed in a noisy context, using both grammatical and morphological patterns. An automatic expansion of domain-specific lexicons is presented as a supervised learning task in [1]. Finally, work in [4] focuses on the choice of relevant initial terms when running a bootstrapping method.

3 Method

This work can be seen as a two-step process, as we first need to identify subjective adjectives, not only for the extraction but also as a way to expand a subjective terms lexicon which can be a useful tool on its own. This is done by a supervised learning task using a predefined lexicon of subjective adjectives. We then apply a scoring function based on the colocation distance to sort the key nouns that are the most frequently qualified by subjective markers.

3.1 Subjective Adjectives Detection

A Machine Learning Task To achieve the subjective adjectives detection task we first applied the method described in [3] using a prefix tree of grammatical patterns to perform NER. It turned out that the context of our experiment was too different to provide acceptable results. Switching to a machine learning task with grammatical patterns as features greatly improved the detection.

Choice of Features The very nature of the text we process implies strong variations, both in morphology and syntax, and thus the best set of features for adjective classification can vary. It would be possible to define several sets of features to best fit each type of corpus, but we considered this for future work and instead we only kept the features that provided the best overall results:

- The part-of-speech and lemma of the two preceding words
- The suffix of the candidate
- A flag indicating if the preceding word is an adverb from a list of amplifiers

Where the suffix is either a known subjective adjective suffix from a manually defined list (*"eux", "ard", "ible"...*) or the last 4 letters of the candidate, and the list of amplifiers a set of 32 recurring adverbs preceding subjective adjectives, such as *"très"*("very"), *"trop"*("so") or *"assez"*("quite"). The features are extracted for each candidate in the text, that is to say each token that our postagger annotated as an adjective. In order to improve the precision of our results we filter the candidates list through a French morphological adjective

lexicon which is a combinaison of both Morphalou[4] and Lefff[5] lexicons. We used LIBLINEAR [6] to run the supervised classification with a linear regression model. Many experiments showed the importance as well as the ambiguity of the choice of the train corpus, due once again to the linguistic variations that can be found in social web corpora.

3.2 Target Concepts Extraction

In the second step of our method, we rank nouns that appear in the test corpus, in order to highlight positive and negative topics that the consumers talk about. The ranking is currently done by a scoring function that rely on the adjective - noun distance within a sentence. This function cannot provide a perfect match between the term and the adjective in French sentences but the results are promising enough to show that the overall process could be of great use; improvements on this point will be discussed in the Conclusion.

The score for each term t in a window of size n around it is such that:

$$subjectivity\_score(t) = \sum_{occ_t} \sum_{i=1}^{n} \frac{p(i)}{i} \times \frac{1}{\sum_{i=1}^{n} \frac{1}{i}} \quad (1)$$

where $p(i) \in \{0, 1\}$ is defined by the presence of a subjective adjective at the i^{th} position. Final score is calculated by adding scores of individual occurrences.

4 Experiments

As a practical case of concern extraction in the context of opinion mining, we present here our results when analysing a corpus of social web posts about several French hamburger restaurants. This corpus is composed of 8,000 texts, which contain one to five sentences at most. Many of these sentences are either grammatically incorrect and/or contain spelling mistakes. We manually annotated each adjective of the corpus as either subjective or not subjective.

Subjective Adjectives Detection Using the first 4,000 texts of the corpus we build the classification model with the linear regression module of LIBLINEAR. Each feature set is extracted as described earlier using our first bootstrapping subjective adjectives lexicon. We then run the classification on the last 4,000 texts of the corpus. Without any post processing operations the evaluation of the classification gives the following results:

Table 1: Evaluation of the subjective adjective classification (without threshold).

#Adj. extracted	Precision	Recall	F-measure
515	50.09	49.90	49.99

[4] http://www.cnrtl.fr/lexiques/morphalou/
[5] http://alpage.inria.fr/~sagot/lefff.html
[6] http://www.csie.ntu.edu.tw/~cjlin/liblinear/

In order to improve precision and thus extract an acceptable subjective adjectives lexicon, we set a threshold on the classification score given by LIBLINEAR, below which a candidate would be rejected. This classification score range is in the interval [-44,61] so we set the threshold to 0. The precision we obtain is much better while keeping a reasonable recall for our use case.

Table 2: Evaluation of the subjective adjective classification ($threshold = 0$).

#Adj. extracted	Precision	Recall	F-measure
264	97.72	33.59	49.99

Concept Extraction From the 264 subjective adjective we detected with a high precision, we then process the 4,000 test documents to highlight each common noun that share a strong cooccurrence with the subjective markers. Here is an sample of the ranked list we obtained when applying (1) as described in the Method section. It should be noted that the score for each key noun is strongly related to its frequency, because of the consolidation that sums all scores for each occurrence. However a comparison with the list of the most frequent nouns in the corpus shows great dissimilarities and thus reinforce the idea that this concern extraction can help better understand customer's key concerns regarding its experience with the brand or the service.

Table 3: Extracted key nouns with associated adjectives

Rank	Concept	Score	Examples of cooccurring adjectives	#Adj.
1	frites (chips)	42.04	*maison(homemade), délicieux(delicious), petites(small)*	30
2	burgers	34.71	*meilleurs(better), vrais(real), élaborés(sophisticated)*	26
3	accueil (reception)	19.54	*chaleureux(warming), sympa(nice), jeune(young)*	19
10	prix (price)	10.45	*élevé(high), raisonnable(reasonable), correct*	22
20	personnel (staff)	6.67	*sympa(nice), dynamique(dynamic), agréable(pleasant)*	10
42	goût (taste)	3.39	*vrai(real), unique, subtile(subtle)*	13

This ranked list highlights facts of high value about the points of interest that the subject raises. First of all, it confirms that *frites*(french fries) and *burgers* are not only key topics but also key customer's concerns on their hamburgers restaurant's experience. Furthermore, the fact that *accueil*(reception) is almost as important as the later two, or that *goût*(taste) and *recettes*(recipes) do not appear in the top 30 extracted subjects can be key points for the brand to know and are not trivial to a human reader. This work brings high value information on qualitative market researchers: it helps them focus more quickly on critical topics and brings objectivity to their study of customer experience's analysis.

5 Conclusion and Future Work

This paper presents a two-step process to achieve concern extraction in social web corpora. First we showed that a lexicon of subjective adjectives can be

extracted from an initial seed lexicon using linear regression classification on linguistic features. This extraction greatly improved the results when compared to a method based on grammatical patterns. Secondly we showed that a simple scoring function based on the distance between subjective adjectives and key nouns can highlight with reasonnable accuracy the users needs. This accuracy could be improved but already shows encouraging results for this type of noisy sentences. In particular, the use of both grammatical patterns and morphological features proved to be very effective. Some improvements are already in experiment and some further features that could be added to this extraction include:

Human Defined Patterns Some natural and intuitive patterns seem to occur frequently in our corpora, as {*être(to be), candidate*} or {*être(to be), adverb, candidate*}. In fact recurring patterns are very few and this lack of consistency can explain the results of the prefix tree method. With that said, human defined patterns in our case could ensure precision and therefore reduce the bootstrapping drift.

Polarity Classification Currently, classification aims at separating subjective from non-subjective adjectives. In the future, we will classify adjectives as positive or negative markers. This distinction would allow a better understanding of consumer reviews.

Better Adjective Identification In order to improve the accuracy of our scoring function we could incorporate a syntaxical method to better identify the related terms of subjective adjectives.

As user concern towards product features become a key issue for a large number of brands, we hope that this work become the starting point of more NLP research to address this problem.

References

1. Avancini, H., Lavelli, A., Zanoli, R.: Automatic expansion of domain specific lexicons by term categorization. ACM Transactions on Speech and Language Processing 3, 2006 (2004)
2. Murray, G., Carenini, G.: Subjectivity detection in spoken and written conversations. Natural Language Engineering 17, 397–418 (7 2011)
3. Nouvel, D., Antoine, J.Y., Friburger, N., Soulet, A.: Fouille de règles d'annotation partielles pour la reconnaissance d'entités nommées. In: Actes de la 20e conférence TALN (TALN'13) (2013)
4. Vincze, N., Bestgen, Y.: Identification de mots germes pour la construction d'un lexique de valence au moyen d'une procédure supervisée. In: Actes de la 18e conférence TALN (TALN'11) (2011)