Explanatory opinions: to whom or what is all the fuzz about?

Filipa Peleja
Yahoo Labs
Barcelona
Spain
peleja@yahoo-inc.com

Ioannis Arapakis
Yahoo Labs
Barcelona
Spain
arapakis@yahoo-inc.com

João Magalhães
NovaLincs, Dep. de Informática
Universidade Nova de Lisboa
Caparica, Portugal
jm.magalhaes@fct.unl.pt

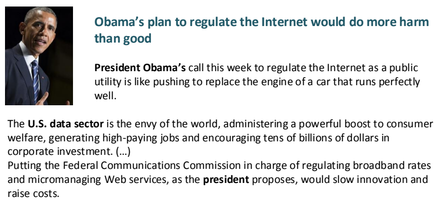
Exploiting sentiment relations to improve the accuracy of sentiment analysis has caught the interest of recent research. When expressing their opinions, users apply different sentence syntactic constructions styles. This analysis leverages on a sentiment lexicon that includes general sentiment words that characterize the overall sentiment towards the targeted named-entity. However, in most cases, target entities are themselves part of the sentiment lexicon, creating a loop from which it is difficult to infer the overall sentiment to the target entities. We propose the application of conditional random fields (CRF) to predict opinion target labels. More specifically, we exploit a set of opinion patterns to extend an opinion word lexicon and then propose to apply a CRF algorithm to detect the interactions between opinion expressions and opinion targets.

sentiment analysis, opinion target

1. INTRODUCTION

Social media has extended people's online interactions beyond simply sharing and commenting on what is happening around them, to exchanging advice and opinions with other members of the same sociosphere. This phenomenon has sparked a relationship between people's opinions and their opinion target. The information targeting the opinion targets is generally controlled by users and consumers (Jansen et al. 2009). Unlike user generated text, where the user (opinion holder) expresses freely her opinion, news articles contain a more structured text with one or more opinion holders targeting several opinion targets. This paper addresses the problem of classifying accurately the sentiment in news articles, as well as the respective sentiment target. The detection of opinion holders and targets in news articles will allow to have a better understanding of the relations between people, organizations and/or countries (Kim and Hovy (2006)). Figure 1 illustrates the opinions expressed in a news article about Internet regulations. In this example, we observe that the opinions expressed in the news article target multiple opinion targets, e.g., President Obama and U.S. data sector.

The analysis of opinionated text, also known as subjective text, involves the detection of words, phrases or sentences that express a sentiment. Although this area has been researched in academia, the problem is still far from being solved Liu (2012). One of the main challenges is that opinionated language varies over a broad range of discourse, and a system with a fixed vocabulary will not be enough to represent users' opinion. Another challenge is to identify relevant mentions to opinion targets which are accompanied by related sentiment words. From an algorithmic perspective, the challenge is to analyse how these sentiment words affect the public image of the opinion targets. Previous work (Hu and Liu (2004); Liu (2012)) has introduced significant advances in detecting product aspects or features, and it is reasonable to apply such methods by analysing how sentiment words affect named entities' reputation. However, unlike products, opinions about named entities are not structured around a fixed set of aspects or features, which implies a more challenging task (Albornoz et al. (2012)).


Obama's plan to regulate the Internet would do more harm than good

President Obama's call this week to regulate the Internet as a public utility is like pushing to replace the engine of a car that runs perfectly well.

The **U.S. data sector** is the envy of the world, administering a powerful boost to consumer welfare, generating high-paying jobs and encouraging tens of billions of dollars in corporate investment. (...)
Putting the Federal Communications Commission in charge of regulating broadband rates and micromanaging Web services, as the **president** proposes, would slow innovation and raise costs.

Figure 1: *Multiple opinion targets in a news article about Internet regulation.*

2. RELATED WORK

Sentiment analysis employs various techniques for detecting words that communicate a positive or negative emotion. These words are commonly known as sentiment words or opinion words. Beyond words, n-grams (contiguous sequence of n words) and idiomatic expressions are commonly used as sentiment words, such as for example “terrible”, “quite wonderful”, and “break a leg”. At document- or sentence- level, sentiment words can be used to predict sentiment classes for users opinions (Liu (2012)). Unlike sentiment analysis at document- or sentence- level, entity- or aspect-level allows for a fine-grained analysis. Entity- or aspect- level sentiment analysis captures specific product features that users dislike or like (Hu and Liu (2004)). For example, Turney (2002) proposed a document level approach to evaluate reviews polarity in which an unsupervised learning algorithm was used to evaluate review’s polarity. For each review, the authors compute the average polarity of its constituent words or phrases. Other works (Pang et al. (2002); Heerschop et al. (2011)) have addressed the sentiment analysis task by using a document-level approach. A common use of sentence-level sentiment analysis is to capture subjective sentences (Wiebe et al. (1999)). When classifying subjectivity, the goal is to distinguish between sentences that express factual information (objective) and sentences that express an opinion (subjective) (Hatzivassiloglou and Wiebe (2000)).

The task of detecting overall sentiment, opinion holders and targets implies several steps (Liu (2012)). In a sentence-level sentiment analysis approach, Meena and Prabhakar (2007) showed that rules based on atomic sentiments of individual phrases can be helpful to decided the overall sentiment of a sentence. However, in Meena at al. work, only adjectives and verbs were considered as features, which implies that only those can be related to the opinion target. Furthermore, as Wilson et al. (2009) showed, other word families (e.g., nouns) may share dependency relations with opinion targets (also referred as aspects), which might be indicative of the sentiment expressed towards those terms. In another work by Gildea and Jurafsky (2002), the authors introduced a system based on statistical classifiers to identify semantic relationships. Their system analyses the prior probabilities of various combinations of semantic roles (predicate verb, noun, or adjective) to automatically label domain-specific semantic roles such as Agent, Patient, Speaker or Topic. Similarly to the semantic roles’ detection introduced by Gildea et al., we propose to analyze sentences lexical and syntactic relations to automatically label opinion targets.

3. OPINION WORDS AND OPINION-PHRASES

We employ Moghaddam and Ester (2012) semantic relationships between words to extract opinion-phrases. These have proven to be quite successful in asserting semantic relations between opinion phrases. Table 1 shows the applied rules. For example, rules number 1 and 5 are able to extract the opinion-phrases *(works, amazing)* and *(small, blurry)* from sentences ”The automode works amazing.” and ”The LCD is small and blurry.” respectively.

Table 1: *Patterns to capture opinion-phrases (N is a noun, A is an adjective, V is a verb, h is a head term, m is a modifier, and* $<$*h, m*$>$ *is an opinion phrase)*

1. amod(N, A) $\rightarrow <$ N, A $>$
2. acomp(V, A) + nsubj(V, N) $\rightarrow <$ N, A $>$
3. cop(A, V) + nsubj(A, N) $\rightarrow <$ N, V $>$
4. dobj(V, N) + nsubj(V, N0) $\rightarrow <$ N, V $>$
5. $<$ h1, m $>$ + conj and (h1,h2) $\rightarrow <$ h2, m $>$
6. $<$ h, m1 $>$ + conj and(h1, h2) $\rightarrow <$ h, m2 $>$
7. $<$ h, m $>$ + neg(m, not) $\rightarrow <$ h, not + m$>$
8. $<$ h, m $>$ + nn(h, N) $\rightarrow <$ N + h, m $>$
9. $<$ h, m $>$ + nn(N, h) $\rightarrow <$ n + N, m $>$

The proximity between an opinion target and a single opinion word is key to building the opinion target semantic roles. For this reason, we have used SentiWordNet, which is a popular sentiment dictionary introduced by Esuli and Sebastiani (2006). SentiWordNet is a lexicon created semi-automatically by means of linguistic classifiers and human annotation. In SentiWordNet, each synset is annotated with its degree of positivity, negativity and neutrality.

4. THE PROPOSED MODEL

An important first step to extracting opinion targets in news articles, is understanding how an opinion word is semantically related to an opinion target. To this end, we propose a sentence-level approach, where our method will identify the opinion words and opinion phrases (Section 3). Figure 2 provides an example on how we aim to decompose each sentence.

We suggest to deal with the task of identifying opinion targets as a sequence labelling problem. The problem of opinion target extraction as a sequence labelling task using CRFs, is defined as follows. Given a sequence of tokens, $x = x_1x_2...x_n$ we need to generate a sequence of labels $y = y_1y_2...y_n$. To train the model, a set of labels are defined as ’OW’ and ’OT’, where ’OW’ corresponds to an opinion word or phrase, and ’OT’ to an opinion target. Similarly to Choi et al. (2005), opinion holders detection model, we create a linear-chain CRF

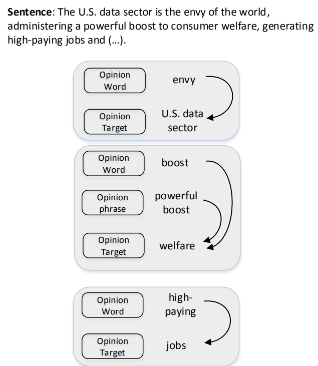


Figure 2: *An overview of our opinions words extraction.*

based on an undirected graph $G = (V, E)$, where for each n tokens of a sentence V is the set of random variables $Y = \{Y_i | 1 < i \leq n\}$. $E = \{(Y_{i-1}, Y_i) | 1 < i \leq n\}$ is the set of $n-1$ edges forming a linear chain. According to Lafferty et al. (2001) the conditional probability of a sequence of labels y given a sequence of tokens x is given by:

$$P(y|x) = \frac{1}{Z_x} exp\Big(\sum_{i,k} \lambda_k f_k(y_{i-1}, y_i, x) + \sum_{i,k} \lambda_k^{'} f_k^{'}(y_i, x)\Big) \tag{1}$$

$$Z_x = \sum_{y} exp\Big(\sum_{i,k} \lambda_k f_k(y_{i-1}, y_i, x) + \sum_{i,k} \lambda_k^{'} f_k^{'}(y_i, x)\Big) \tag{2}$$

where Z_x is a normalization constant for each sentence x, $f_k(\ldots)$ is a binary feature indicator function, λ_k is the weight of each feature function, and k are the number of features defined for edges and $k^{'}$ for nodes.

5. EXPERIMENTS

5.1. Dataset

The goals are experiments first, to see how accurately we can perform a binary sentiment classification, and second, to examine the correlation between opinion phrases and opinion targets. For this analysis, one challenge to overcome is the lack of labeled data. To this end, we have selected a labeled dataset from SemEval-2014 challenge. This dataset contains opinionated sentences from the restaurants domain, and it is part of the Task 4: *Aspect Based Sentiment Analysis* of the abovementioned challenge [1]. In addition the dataset has a total of 1601 annotated sentences in which 1198 and 403 are positive and negative respectively. In addition, the dataset presents a mean of 66 characters and 12 words per sentence.

[1] http://alt.qcri.org/semeval2014/task4/

5.2. Sentiment classification

For our sentiment classification task, the sentences are classified according to a deterministic binary classification in which sentences are classified as either positive or negative. TO classify the sentences we applied a 10-fold cross validation using the Weka [2] implementation of SVM (Support Vector Machines). Table 2 shows the initial sentiment classification results.

Table 2: *Sentiment classification of comments from restaurant reviews.*

Polarity	Precision	Recall	F1
positive	0.792	0.915	0.849
negative	0.583	0.331	0.422

We note, that the classifier performs better on positive sentences. One reason for this could the imbalanced nature of the dataset. Also, as it has been demonstrated by previous work (Liu (2012)), users tend to frequently apply the same opinion word both in positive and negative contexts. This sentiment classification experiment aims to validate the quality of the selected opinion words and opinion phrases in a sentiment classification task.

5.3. Opinion phrases and opinion targets

In the present work we argue that there are many semantic relations between opinion words and opinion phrases that semantic relations analysis is not able to capture, i.e. subject and object relations. Also, as expected, we notice an intersection between opinion phrases and opinion targets. For example, in the sentence ”The service was excellent and the food was delicious.” the labeled opinion targets are ”food” and ”service” and the extracted opinion phrases are ”food delicious”, ”service excellent” and ”service delicious”. In this context, an opinion phrase is defined as a pair (aspect, opinion), therefore aspect has a high probability to be an opinion target. For the extracted opinion phrases and labeled opinion targets we observe a Jaccart similarity of 0.28. Here, Jaccart similarity refers to the quotient between the intersection of opinion- phrases and targets. Although we observe intersection between these objects there are many opinion targets that are not within the obtained opinion phrases.

5.4. Future work: Opinion targets in news articles

We observed that grammatical dependencies can be used to extract aspects and opinion phrases. However, it is noticeable that a more in-depth approach should be applied to improve the opinion

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

targets extraction. As future work, we aim at developing the proposed CRF model to obtain a higher coverage of the opinion targets. Finally, for the experiments shown in this section, we used a dataset from the restaurants' domain. In addition, the sentences were extracted from users' reviews, which have a structure that is considerable different to that observed in news articles. We also aim at obtaining an labeled news articles dataset to extend the opinion prediction model to this domain as well.

6. CONCLUSIONS

In this paper we discussed techniques to detect opinion targets. In opinionated sentences, an opinion target is the entity that is targeted by the sentiment expressed in the sentence. Our experimental results show that opinion phrases present an clear intersection with opinion targets. However, it is evident that there are many opinion targets that are not captured by this method. We believe that this is because fixed language pattern rules are not enough to cover the range of discourse used to express an opinion, as well as the respective target. In the future we plan to extend our work to a news articles dataset that is characterized by a different type of discourse, and apply a method based on CRF to detect opinion targets language patterns.

REFERENCES

Albornoz, J., I. Chugur, and E. Amigó (2012). Using an Emotion-based Model and Sentiment Analysis Techniques to Classify Polarity for Reputation. In *Conf. and Labs of the Evaluation Forum, Online Working Notes (CLEF)*.

Choi, Y., C. Cardie, E. Riloff, and S. Patwardhan (2005). Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. In *Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP)*.

Esuli, A. and F. Sebastiani (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. *Proc. of the 5th Conf. on Language Resources and Evaluation (LREC)*.

Gildea, D. and D. Jurafsky (2002). Automatic Labeling of Semantic Roles. *Computational Linguistics*.

Hatzivassiloglou, V. and J. Wiebe (2000). Effects of adjective orientation and gradability on sentence subjectivity. *Proc. of the 18th Conf. on Computational Linguistics (COLING)*.

Heerschop, B., F. Goossen, A. Hogenboom, F. Frasincar, U. Kaymak, and F. De Jong (2011). Polarity analysis of texts using discourse structure. *Proc. of the 20th Conf. on Information and Knowledge Management (CIKM)*.

Hu, M. and B. Liu (2004). Mining opinion features in customer reviews. *Proc. of the Association for the Advancement of Artificial Intelligence 19th Conf. on Artifical Intelligence*.

Jansen, B., M. Zhang, K. Sobel, and A. Chowdury (2009). Twitter Power: Tweets As Electronic Word of Mouth. *Journal of the American Society for Information Science and Technology*.

Kim, S. and E. Hovy (2006). Extracting opinions, opinion holders, and topics expressed in online news media text. *Proc. of the Workshop on Sentiment and Subjectivity in Text*.

Lafferty, J., A. McCallum, and F. Pereira (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proc. of the 18th Conf. on Machine Learning (ICML)*.

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*.

Meena, A. and T. Prabhakar (2007). Sentence Level Sentiment Analysis in the Presence of Conjuncts Using Linguistic Analysis. In *Proc. of the 29th European Conf. on Advances in Information Retrieval (ECIR)*.

Moghaddam, S. and M. Ester (2012). On the Design of LDA Models for Aspect-based Opinion Mining. *Proc. of the 21st Conf. on Information and Knowledge Management (CIKM)*.

Pang, B., L. Lee, and S. Vaithyanathan (2002). Thumbs up?: sentiment classification using machine learning techniques. *Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP)*.

Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proc. of the 40th Annual Meeting on Association for Computational Linguistics (ACL)*.

Wiebe, J., R. Bruce, and T. O'Hara (1999). Development and use of a gold-standard data set for subjectivity classifications. *Proc. of the 37th of the Association for Computational Linguistics on Computational Linguistics (ACL)*.

Wilson, T., J. Wiebe, and P. Hoffmann (2009). Recognizing Contextual Polarity: An Exploration of Features for Phrase-level Sentiment Analysis. *Journal Computational Linguistics*.