

# Mining the Sentiment Expectation of Nouns Using Bootstrapping Method

**Miaomiao Wen**

Language Technologies Institute  
Carnegie Mellon University

mwen@andrew.cmu.edu

**Yunfang Wu**

Key Laboratory of Computational  
Linguistics (Peking University)  
Ministry of Education, China

wuyf@pku.edu.cn

## Abstract

We propose an unsupervised bootstrapping method to generate a new type of affect knowledge base: the sentiment expectation of nouns (e.g., “high salary” is desirable while “high price” is usually undesirable, because people have opposite sentiment expectation towards “salary” and “price”). A bootstrapping framework is designed to retrieve patterns that might be used to express complaints from the Web. The sentiment expectation of a noun could be automatically predicted with the output patterns. We evaluate the retrieved patterns and show that our method yields good results. Also, they are applied to improve both sentence and document level sentiment analysis results.

## 1 Introduction

In recent years, sentiment analysis has attracted considerable attention in the NLP community due to its wide applications. The task is mining positive and negative opinions from text and an in-depth review of its literature can be found in Pang and Lee (2008). Previous work on this problem falls into three groups: opinion mining of documents, sentiment classification of sentences and polarity prediction of words. Sentiment analysis both at document and sentence level rely heavily on word level.

The most frequently explored task at the word level is to determine the sentiment orientation (SO) of words or word senses in the lexicon. While adjectives and verbs are often considered, the sentiment classification of nouns still poses a challenge. This

paper aims at identifying people’s sentiment expectation towards a noun, even though the noun itself does not carry polarity. We propose three categories of sentiment expectation (SE) of nouns: positive expectation nouns ( $Pn$ ), negative expectation nouns ( $Nn$ ) and neutral. For example, “工资|salary” is a  $Pn$ , as “high salaries” is desirable for most people. Also, the noun “价格|price” describes an object that is generally neutral, but it is a  $Nn$ , as most people in most cases expect that the product prices become lower.

There are several significances lying in this study. First, the SE of noun reflects world knowledge about an object, which is not readily available in existing semantic resources. This knowledge is useful in determining the context dependent SO of adjectives or verbs. For example, “high salary” is desirable while “high price” is undesirable. Also, “receive money” will probably impart positive state onto its patient while “receive hepatitis” will impart negative state onto its patient. Second, our method requires very little human supervision.

We introduce an unsupervised bootstrapping approach. Our system is initialized with a very small seed set of nouns, and then iterates between (a) retrieving a set of *complaint patterns* (CPs) - lexico-syntactic patterns such as “<n> is a little a”<sup>1</sup> that tend to occur only in people’s complaints - from search engine snippets and (b) using the acquired patterns to determine the SE of new nouns.

The rest of this paper is organized as follows: The related work is discussed in Section 2. In Sections 3,

<sup>1</sup>In this paper, <n> represents a noun and *a* represents an adj.

we present our bootstrapping method. In Section 4, we conduct evaluation experiments at both sentence and document level sentiment analysis. The paper is concluded in Section 5.

## 2 Related Work

Bootstrapping and pattern-based methods have been shown to be very effective in previous information extraction research (Riloff, 1996; Riloff and Jones, 1999; Ravichandran and Hovy, 2002; Thelen and Riloff, 2002; Riloff et al., 2003; Mooney and Bunescu, 2005; Wiebe and Mihalcea, 2006; Kozareva et al., 2008). These previous works derive patterns that reveal direct relationship between two words or the property of the word. Though similar in methodology, we focused on patterns that express an implicit relationship between the target noun and the opinion-bearing adjective.

There has been a large body of work on automatic SO prediction of words (Hatzivassiloglou and McKeown, 1997; Turney and Littman, 2003; Kim and Hovy, 2004; Esuli and Sebastiani, 2006), but unfortunately they did not consider the SE of nouns in their research and regarded most of the nouns as “neutral”. Recently, some studies try to disambiguate the context dependent SO of adjectives (e.g. distinguish between “the battery life is very *long*” and “it takes a *long* time to focus”) (Ding et al., 2008). They infer the context dependent SO by inferring with intra-sentence conjunction rule. Our task is more challenging as we have no global or domain information. Thus our method could be applied to isolated phrase or sentence and is domain independent. Wu and Wen (2010) present the first algorithm for retrieving SE of nouns automatically but the results critically depends on access to a high quality, carefully chosen collection of CPs.

## 3 Our Approach

Our main insight is to make use of CPs<sup>2</sup>, which co-occur frequently with the target noun and the adjective that is opposite to the noun’s SE. E.g. people

<sup>2</sup>Wish patterns (WP), which are usually used in people’s praises and wishes, might also serve our purpose (Goldberg et al., 2009). But using the Web as a corpus, we found that the CPs are much easier to be retrieved than WPs. Thus we will just focus on CPs in this paper and leave WPs for future work.

---

**Input:** Noun\_Pool =  $\{Pn, Nn\}$ , A =  $\{Pa, Na\}$ ;

**Output:** CP\_Pool =  $\{N\text{ CPs}\}$ ;  $i = 0$ ;

### Bootstrapping

1. Candi\_CP\_Pool  $\leftarrow$  patterns extracted from snippets. For each pattern  $p$ ,

$score(p) = TScore(p.CPFreq, p.WPFreq)$

2. CP\_Pool = the  $N$  patterns with the highest score in Candi\_CP\_Pool. If CP\_Pool remains unchanged for two succeeding loops or if  $i > K$ , stop bootstrapping.

3. Candi\_Noun\_Pool = Extract new nouns from snippets.

4. Train SVM with CP\_Pool. Training set  $\leftarrow$  Noun\_Pool, testing set  $\leftarrow$  Candi\_Noun\_Pool

For each noun  $n$ , the feature vector  $F_i = TScore(\sum_{a \in Na} Hits(p_i(n, a)), \sum_{a \in Pa} Hits(p_i(n, a)))$ ,  $i = 1, \dots, N$ .

5. Add the nouns with the largest posterior probability and the smallest posterior probability to  $Pn$  and  $Nn$  respectively.

6.  $i = i + 1$ , Go to Step 1.

---

Table 1: The Whole Bootstrapping Process.

might say “工资有点低| salary is a little low” but seldom say “工资有点高| salary is a little high”. Utilizing this property, we try to (a) extract patterns like “ $\langle n \rangle$  is a little  $a$ ” from snippets returned queries like “工资低|salary low”;<sup>3</sup> (b) Use SVM to determine the SE of new nouns with page counts based features. E.g. “价格有点低| price is a little low” obtains 1080 hits while “价格有点高| price is a little high” obtains 19400 hits.

We use a bootstrapping method to automatically discover CPs and predict sentiment expectation of nouns (Table 1). In iteration phase 1, with a few seed nouns, the candidate CPs are retrieved from search engine snippets. We rank these CPs according to their ability to express SE. In iteration phase 2, we infer the sentiment expectation of a noun by mining the Web with CPs. The SVM is trained to classify positive expectation nouns and negative expectation nouns.

**Initiation:** The bootstrapping begins with a seed noun set Noun\_Pool =  $\{Pn, Nn\} = \{\{\text{“工资|salary”}\}, \{\text{“价格|price”}\}\}$ , and an adjective set A. A is grouped into two sets: positive-like adjectives ( $Pa$ ) and negative-like adjectives ( $Na$ ):

<sup>3</sup>We use Baidu as our search engine in this paper, <http://baidu.com.cn>

(1)  $Pa = \{\text{大|large, 多|many, 高|high, 厚|thick, 深|deep, 重|heavy}\}$

(2)  $Na = \{\text{小|small, 少|few, 低|low, 薄|thin, 浅|shallow, 轻|light}\}$

**Phase 1. Extract CPs from Snippets:** For each snippet returned by the query “ $n$   $a$ ”  $\in$   $Noun\_Pool : A$ , extracts word  $n$ -grams after word segmentation. We select  $n$ -grams which contain exactly one  $\langle n \rangle$  and one  $a$ . Counts the frequency of a pattern  $p$  appears as a CP (where  $(\langle n \rangle, a) \in (Pn, Na) \text{ or } (Nn, Pa)$ ) and the frequency of  $p$  appears as a WP (where  $(\langle n \rangle, a) \in (Nn, Na) \text{ or } (Pn, Pa)$ ). Finally, we adopt T-Score to measure the confidence with which we can assert whether this pattern is a CP (Step 1 in Table 1). The top ranking  $N$  patterns are selected as CPs experimentally.

**Phase 2. Determine the SE of new nouns:** We create a feature vector  $F$  using the harvested CPs for each noun. The two-class support vector machine (SVM) is trained to find the optimal combination of the page counts-based features (Step 4 in Table 1). We define the SE of a noun as the posterior probability (converted from SVM output with a sigmoid function (J. Platt., 2000)) that they belong to the positive expectation noun class. Then the nouns with the largest and smallest probability are added to the  $Noun\_Pool$ .

## 4 Evaluation

### 4.1 Direct Evaluation

We examine the harvested CPs directly to determine whether they are actually CPs or not. Our evaluation metric is the precision of the output CPs at each bootstrapping iteration. We recruited two Chinese speakers to label them as “being CP”, “not CP” or “hard to decide”.

Figure 1 plots the number of correctly retrieved CPs in the output at each iteration. Clearly, we see general substantial improvement along with the bootstrapping process, although the increases level off in later iterations. As the number of the output CPs increases, there are more patterns that are labeled as “hard to decide”, but some of these patterns could still serve our purpose. E.g., “ $\text{因} \langle n \rangle \langle a \rangle | \text{because} \langle n \rangle \text{is} \langle a \rangle$ ” and “ $\text{我} \langle n \rangle \langle a \rangle | \text{my} \langle n \rangle \text{is} \langle a \rangle$ ”, as people often

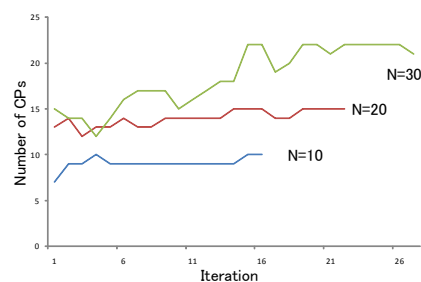


Figure 1: Number of CPs in results returned by the bootstrapping method at each iteration.  $N$  is the number of output patterns. Items labeled as “hard to decide” are not included.

post their problems and complaints on Web. These patterns may serve as possible indicators of the presence of a noun and the reversed expectation adjective, regardless of whether they are actually restricted in negative contexts.

The top ranked 10 CPs are of high quality. The manually chosen patterns in Wu and Wen (2010) have been successfully retrieved. This shows that our method yields good results. The 10 CPs and the nouns added in the bootstrapping process are listed below:

**CPs:**  $\langle n \rangle$ 有点  $a$  |  $\langle n \rangle$  is a little  $a$ ,  $\langle n \rangle$ 太  $a$  |  $\langle n \rangle$  is too  $a$ ,  $\langle n \rangle$ 是不是太  $a$  | isn’ t  $\langle n \rangle$  too  $a$ ,  $\langle n \rangle$ 实在太  $a$  |  $\langle n \rangle$  is really too  $a$ , 解决  $\langle n \rangle$   $a$  | solve a  $\langle n \rangle$ ,  $\langle n \rangle$   $a$ 怎么办 |  $\langle n \rangle$  is  $a$ , what should I do, 嫌  $\langle n \rangle$   $a$  | hate a  $\langle n \rangle$ , 因  $\langle n \rangle$   $a$  | because  $\langle n \rangle$  is  $a$ ,  $\langle n \rangle$ 过  $a$  |  $\langle n \rangle$  is too  $a$

**Pn:** 空间|room, 素质|quality, 水平|level, 效率|efficiency, 内存|PC memory, 收入|income, 孩子|children, 时间|time  
**Nn:** 要求|requirement, 压力|pressure, 成本|cost, 风险|risk, 问题|problem, 花费|expense, 代价|cost, 房价|house price, 损失|lost

### 4.2 Sentiment Analysis at Sentence Level

We apply the SE of nouns to predict the SO of sentiment ambiguous adjectives, which is SemEval-2010 Task 18 (Wu and Jin, 2010).

**Data:** We use the benchmark dataset of SemEval-2010 Task 18. The task consists of 14 sentiment ambiguous adjectives (SAA) (divided to  $Pa$  and  $Na$  sets same as in Section 3), which are all high-frequency words in Mandarin Chinese. Each of the 2917 sentences in the dataset contains a target noun and a SAA.

**Methods:** The SO of SAA can be determined by the target noun in noun-adjective phrases. If the SAA has the same polarity as the SE of noun, then the SAA has positive sentiment; if the SAA has the op-

posite polarity to the SE of noun, the SAA has negative sentiment.

**Results:** Compared with the other 16 systems that participated in Task 18, our system ranks fifth and is substantially better than baseline (Table 2). Note that all the top ranked 11 systems are supervised or incorporate manually built library. Our system also outperforms Wu and Wen (2010), indicating our bootstrapping method works better than the manually selected patterns.

	Micro Acc.	Macro Acc.
Our Method	77.63	79.52
Wu & Wen	75.83	71.67
Baseline	61.20	62.37

Table 2: The scores on SemEval-2010 Task 18.

### 4.3 Sentiment Analysis at Document Level

We also investigated the impact of recognizing SE of nouns and CPs on the sentiment classification of product reviews. SAAs are frequently used in product reviews and could be sentiment disambiguated by the SE of nouns. Also, CPs usually indicate the speaker is complaining and unsatisfied with the product (i.e. negative reviews). For example, “按键设计太紧密| keyboards are designed *too* close” and “价格偏贵| It is a *little* expensive”.

**Data:** Following the work of Wan (2008) and Wan (2009), we selected the same dataset<sup>4</sup>. The dataset of Wan (2008) contains 886 Chinese product reviews. The dataset of Wan (2009) contains another 1000 unlabeled Chinese product reviews. We manually annotated these product reviews with positive or negative polarity labels. We use both these two datasets as our test set, which includes 1886 reviews. In order to examine the impact of recognizing SE of nouns, we extracted the files that contain the following strings, where the nouns are modified by SAAs in most cases:

- (3) noun+adjective (adjective∈SAA)
- noun+adverb+adjective
- noun+adverb+adverb+adjective.

We obtained 449 files (SAA-set for short), up to 24% of the overall data.

**Methods:** The baseline method is the same algorithm with Wan (2008). The semantic orientation

<sup>4</sup>Available here: <http://sites.google.com/site/wanxiaojun1979/publicationlist-1>

value for a review is computed by summing the polarity values of all words in the review, making use of both the word polarity defined in the positive and negative lexicons and the contextual valence shifters defined in the negation and intensifier lexicons. We also use the same parameter setting and the same sentiment lexicon<sup>5</sup>.

**Our method:** (a) Add the disambiguation of SO of SAAs to the algorithm. When a word  $\in$  SAA, compute its SO with our method in Section 4.2, rather than using its prior polarity specified in the sentiment lexicon. (b) Use the CPs as indicators of negative comments. If any CP appears in a review, then it is judged as negative SO.

**Results:** Our method obviously outperforms the baseline by 12.16% in f-score and 17.02% in accuracy (on SAA-set, see Table 3). The improvement in recall is especially obvious. The results also indicate using more CPs could bring further improvement.

		Base line	SE N=10	SE N=20	SE N=30
Pos.	Pre.	65.69	81.93	83.44	85.28
	Rec.	76.40	74.16	76.40	75.96
	F	70.16	79.07	79.77	80.35
Neg.	Pre.	87.43	84.35	84.53	83.77
	Rec.	60.96	88.05	89.24	90.24
	F	71.83	86.16	86.82	86.89
Total	MacroF	70.98	82.46	83.14	83.49
	Acc.	66.90	83.46	83.92	84.15

Table 3: The sentiment classification results at document level. SE denotes our method. N is the number of CPs.

## 5 Conclusions

This paper presents an unsupervised bootstrapping method to retrieve the sentiment expectation of nouns from the Web. We utilize the predicted SE of nouns in determining the SO of sentiment ambiguous adjectives. For the sentiment analysis at sentence level, our method achieves promising result that is significantly better than baseline and comparable to the supervised methods. For the sentiment analysis at document level, our method also achieves obvious improvement in performance, which validates the effectiveness of our approach.

<sup>5</sup>Sentiment Hownet, a manually constructed Chinese opinion lexicon: [http://www.keenage.com/html/c\\_index.html](http://www.keenage.com/html/c_index.html)

## References

- Ding, X., Liu, B. and Yu, P. 2008. A holistic lexicon-based approach to opinion mining. *In Proceedings of WSDM'08*.
- Esuli, A. and Sebastiani, F. 2006. SentiWordNet: a publicly available lexical resource for opinion mining. *In Proceedings of LREC'06*.
- Goyal, A., Riloff, E., Daume III, H. 2010. Automatically Producing Plot Unit Representations for Narrative Text. *In Proceedings of EMNLP2010*.
- Goldberg, A., Fillmore, N., Andrzejewski, D., Xu, Z., Gibson, B., and Zhu, X.. 2009. May all your wishes come true: A study of wishes and how to recognize them. *In Proceedings of NAACL-HLT2009*.
- Hatzivassiloglou, V. and McKeown, K. 1997. Predicting the semantic orientation of adjectives. *In Proceedings of ACL1997*.
- J. Platt. 2000. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. *Advances in Large Margin Classifiers*.
- Kozareva, Z., Riloff, E. and Hovy, E. 2008. Semantic Class Learning from the Web with Hyponym Pattern Linkage Graphs. *In Proceedings of ACL-HLT2008*.
- Kim, S. and Hovy, E. 2004. Determining the sentiment of opinions. *In Proceedings of COLING2004*.
- Pang, B. and Lee, L. 2008. Opinion mining and sentiment analysis *Foundations and Trends in Information Retrieval*.
- Ravichandran, D and Hovy, E. 2002. Learning surface text patterns for a question answering system. *In Proceedings of ACL2002*.
- Riloff, E. 1996. Automatically generating extraction patterns from untagged text. *In Proceedings of the 13th National Conference on Artificial Intelligence*.
- Riloff, E., Wiebe, J., and Wilson, T.. 2003. Learning subjective nouns using extraction pattern bootstrapping. *In Proceedings of CoNLL2003*.
- Riloff, Ellen and Jones, Rosie. 1999. Learning dictionaries for information extraction by multi-level bootstrapping. *In Proceedings of AAAI1999*.
- Raymond J. Mooney and Razvan Bunescu. 2005. Mining knowledge from text using information extraction. *In ACM SIGKDD Exploration Newsletter, 7(1):3 - 10*.
- Thelen, M. and Riloff, E. 2002. A Bootstrapping Method for Learning Semantic Lexicons Using Extraction Pattern Contexts. *In Proceedings of EMNLP2002*.
- Turney, P. and Littman, M. 2003. Measuring praise and criticism: inference of semantic orientation from association. *In ACM transaction on information systems*.
- Turney, P. D., and Pantel, P. 2010. From frequency to meaning: Vector space models of semantics. *In Journal of Artificial Intelligence Research*.
- Wan, X. 2008. Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. *In Proceedings of EMNLP2008*.
- Wan, X. 2009. Co-Training for Cross-Lingual Sentiment Classification. *In Proceedings of ACL-IJCNLP2009*.
- Wiebe, J. and Mihalcea, R. 2006. Word Sense and Subjectivity. *In Proceedings of OLING-ACL2006*.
- Wu, Y. and Jin, P. 2010. SemEval-2010 task 18: disambiguating sentiment ambiguous adjectives. *In Proceedings of SemEval 2010*.
- Wu, Y and Wen, M. 2010. Disambiguating Dynamic Sentiment Ambiguous Adjectives. *In Proceedings of COLING2010*.