Sentiment Polarity Analysis based multi-dictionary

Jian Jiao\textsuperscript{a}, Yanquan Zhou\textsuperscript{a}

\textsuperscript{a}School of Computer Science and Technology Beijing University of Posts and Telecommunications, Beijing, China

Abstract

This paper presents a novel algorithm for Chinese online reviews, which identifies sentiment polarity. To determine the sentence is negative or positive, we extracted opinion words and identified their opinion targets by CRFs and establish the absolute emotional dictionary (AbED), the relative emotional dictionary (ReED), the field of emotional dictionary (FiED) and the field of targets and opinion words dictionary (TfED). With those emotional dictionary, negative dictionary and modified dictionary, we achieved an effective algorithm to discriminate sentiment polarity by multi-string pattern matching algorithm. For evaluation, we used car online reviews, hotel online reviews and computer online reviews which annotated positive or negative. Experimental results show that our proposed method has made a higher precision and recall rate.

Keywords: sentiment polarity; absolute emotional dictionary; relative emotional dictionary; field of emotional dictionary; field of targets and opinion words dictionary

1. Introduction

With the rapid development of Web technology, people are receiving more and more information from the Web. How to provide people concerned information from the vast amount of Web data, Sentiment Polarity Analysis is becoming more and more important.

Turney\textsuperscript{y} [1] applied point of view of the adjectives and adverbs sentence contains the phrase posed by the tendency to evaluate the article polarity. Kim and Hovy [2] automatic evaluation of the identification holder of this topic, and emotional expression by the given topic and calculated the emotional tendency of words to determine the sentence set the emotional tendency. Hiroshi [3] applied some techniques of deep language analysis for machine translation, proposing a new paradigm for...
sentiment analysis: translation from text documents to a set of sentiment units. Kennedy [4], in the field of movie reviews, two methods was applied to identify the polarity of the comment text: extending the term-counting method with contextual valence shifters improves the accuracy of the classification. The second method uses Support Vector Machines. Devitt [5] explored a computable metric of positive or negative polarity in financial news text. McDonald [6] investigated a structured model for jointly classifying the sentiment of text at varying levels of granularity. Bliter [7] investigated domain adaptation for sentiment classifiers, focusing on online reviews for different types of products.

The rest of the paper is organized as follows. Section 2 introduces the system framework. In section 3, we introduce how to build dictionaries. Section 4 will introduce the classification algorithm that applied multi-pattern matching [8] algorithm and integrate multiple dictionaries. Evaluation and experiment of algorithms in the section 5.

2. Approach

Different areas or different contexts, the word emotion is dynamic tendency in Chinese. Therefore, we marked emotion dictionary that has been divided manually into the absolute emotional dictionary (AbED) and relative emotional dictionary (ReED). There are many special emotional words. So we have established the field of emotional dictionary (FiED) and field of targets and opinion words dictionary (TfED). We applied the proposed classification algorithm to obtain good results in online product reviews as shown in Fig.1.

Fig.1. System Frame Diagram

We defined the field of the targets and opinion words dictionary TfED = {<op, target, value>|op ∈ FiED, target ∈ TaRD}, and the field of the targets dictionary TaRD = {t1, t2, t3, t4… …}; these also have the following properties:

Property.1: AbED ⊖ ReED= ∅, AbED and ReED is mutually exclusive.
Property.2: TfED ⊂ FiED × TaRD, TfED is a subset of the Cartesian product of FiED and TaRD.

Some training and testing data sets are used in the system. We define OTData that is a training corpus that words and emotions have been tagging. SPData is a sentence corpus set that have been marked sentiment polarity. BLData is a larger areas data.
3. The establishment of a dictionary

3.1. AbED and ReED

Hownet’s dictionary and manually marked emotion word dictionary was manually divided into AbED and ReED, the words of AbED were give emotional polarity, positive tendency to +1, negative tendency to -1. Due to relative emotional words with the different areas, different context, different expressions, the emotional tendency may be different. So we introduce a number of areas related data have been labeled and apply the Eq. (1) assess polarity of emotional words. Some don’t appear on the corpus of emotional words were initial reservations.

\[ P(w|d_+) = p(w|d_+)p(w|d_+) + p(w|d_-)p(w|d_-) \]

\[ P(w|d_-) \] is the probability of word \( w \) in the positive text, \( p(w|d_-) \) is the probability of word \( w \) in the negative text, \( p(w|d_-) \) is the probability of word \( w \) as positive word, \( p(w|d_-) \) is the probability of word \( w \) as negative word; \( p(w|d_-) \) is the value of word \( w \) as positive word; \( p(w|d_-) \) is the value of word \( w \) as negative word.

3.2. Field of emotional dictionary (FiED)

There are still specific areas of emotional expression. In order to accurately explore this kind of emotional words, choose the following features: \( F_N = \{f_{1n}, f_{2n}, f_{3n}, f_{4n}\} \), \( f_{1n} = \) current word, \( f_{2n} = \) current pos, \( f_{3n} = \) distance from the adverb, \( f_{4n} = \) distance from the noun, OTData as the training corpus, using the CRFs effectively extracted emotional words. Application of Eq.(1) trained emotional polarity in the SPData and constitute the field of emotional dictionary (FiED).

3.3. Field of the targets dictionary (TaRD)

With the emotional words describing the different objects or targets, the polarity of emotional words may have with the changes. Some of the words themselves do not have the emotional, but it and the target together have some emotions tendency. Therefore, we selected the following features: \( T_N = \{t_{1n}, t_{2n}, t_{3n}, t_{4n}\} \), \( t_{1n} = \) current word, \( t_{2n} = \) current pos, \( t_{3n} = \) Named entity, \( t_{4n} = \) distance from opinion word. We applied CRF++ by training data OTData and testing on SPData. Extracting the targets or objects in product field, and we have established the field of the targets dictionary (TaRD).

3.4. TfED

Although we have taken FiED and TaRD, but we did not determine opinion word of sentiment polarity under the conditions in a given target. We select the center of emotional words to \([-n, +n]\) for the window, taking all the “< targets, opinion words >” in BLData, and using Eq.(2) to assess <opinion words, targets, polarity> to generate the field of the targets and opinion words dictionary in SPData.

\[ P(\geq) = p(\geq) p(\geq|d_-) + p(\geq) p(\geq|d_-) \]

\[ P(\leq) = p(\leq) p(\leq|d_-) + p(\leq) p(\leq|d_-) \]
\[ p(<>|d_+) \text{ is the probability of } <\text{targets, opinion words}> \text{ in the positive text } d_+, \quad p(<>|d_-) \text{ is the probability of } <\text{targets, opinion words}> \text{ in the negative text } d_- , \quad p(+|<>) \text{ is the positive probability of } <\text{targets, opinion words}>, \quad p(-|<>) \text{ is a negative probability of } <\text{targets, opinion words}>, \quad P(+|<>^+) \text{ is the positive value of } <\text{targets, opinion words}>, \quad P(-|<>^-) \text{ is the negative value of } <\text{targets, opinion words}>. \]

4. Sentiment Classification Algorithm

For product reviews, in the expression of the people, the study found that people carrying the major components of emotion is emotion words, the degree adverb modifying an adverb or negative, so we proposed a novel algorithm.

First of all, each a statement for the separation, the multi-pattern string matching algorithms [8] taken all the emotional words, adverbs modifying adverbs or negation in the sentence. The word appears in a sentence in order. For each word, if it is emotional, then in the window [-n, + n] to find whether it contains within the object, If it is true, the value of emotional words under the conditions in the given target and emotional words from the most recent modifier and negative adverbs are multiplied. If the window does not contain the object, emotional value of words multiplied. Computing the scores of sentence, we determine the emotional tendency by a given threshold.

Classification algorithm is as follows:

Input: test sentence S, set the threshold \( \theta \).
Output: sentence sentiment polarity, positive or negative.

Step1: the sentence S, according to the clause segmentation punctuation, composition clause set \( S_{set} \);
Step2: for each …
Step3: Find a sentence \( S_i \) in the presence of AbED, ReED, FiED of the word, constitute the word set \( W_{set} \)
Step4: words order in the sentence appeared to sort to form a new word set \( W'_{set} \) for each \( w_i \in W'_{set} \) do
    if \( w_i \) is negative adverbs or modifier \( \text{ then } S_{adv} = S_{w_i} ; \)
    if \( w_i \) is opinion word \( \text{ then } \) if (target is found in window n && targets is found in TfRD) then \( S_{w_i} = Starop ; \)
    Else \( S_{w_i} = S_{op} ; \)
endif;
S_value += S_{wi}*S_{adv};
S_adv = 1;
Done
Step5: Computing Score = S_value / | Wset |;
Step6: if (Score >= \theta ) then
S’s polarity is positive;
Else
S’s polarity is negative;

5. Experimental

5.1. Data

We grab 12,985 sentences in http://auto.sina.com.cn/ that include the number of positive is 6459 and negative is 6526. We randomly selected 3000 sentences that were manually marked targets or objects and opinion words to constitute OTData. While we randomly selected 1000 positive sentences and 1000 negative sentences from the corpus except OTData to build TestData. The rest is SPData.

We conducted a statistical about the size of n on OTData. Window distance of 1 accounted for 40.4788%. Window distance of 2 accounted for 59.4107%. Distance of 3 accounted for 79.7054%. Window distance of 4 accounted for 88.3242%. So we selected window distance of 4 in the following experiment.

There are many the proposed features can affect the effectiveness and performance of the system, but the main features are AbED, ReED, FiED and TfED, as shown in Table 1.

Table 1. Feature description

| Features | Description                          |
|----------|--------------------------------------|
| F1       | AbED: Absolute emotional dictionary  |
| F2       | ReED: Relative emotional dictionary  |
| F3       | FiED: The field of emotional dictionary |
| F4       | TfED: targets field of emotional dictionary |

In order to better demonstrate the effect of the proposed algorithm, we selected the baseline to be studied. We only used F1, F2, F3 and F4 in Table 2.

It is easy to see that the order of feature’s importance: F3 > F2 > F4 > F1. F3 will be applied only when the positive F was 0.84, negative F to 0.83. So we selected F3 classification as the experiment baseline. In considering the feature combination, they can consider adding F2, further F4, the last to join F1. The algorithm classification results under a given the baseline, as shown in Fig.2(a).

Table 2. Positive and Negative classification results
|       | Positive |       | Negative |
|-------|----------|-------|----------|
|       | P        | R     | F        | P        | R     | F        |
| F1    | 0.29     | 0.38  | 0.33     | 0.12     | 0.08  | 0.11     |
| F2    | 0.67     | 0.76  | 0.71     | 0.72     | 0.62  | 0.67     |
| F3    | 0.81     | 0.88  | 0.84     | 0.87     | 0.79  | 0.83     |
| F4    | 0.11     | 0.06  | 0.08     | 0.32     | 0.45  | 0.37     |

Fig. 2. (a) Positive and negative classification results under a given the baseline; (b) Threshold $\theta$ impacting

Fig. 3. (a) Positive classification results on car, hotel, computer; (b) Negative classification results on car, hotel, computer
With the increasing features, the algorithm's classification performance is constantly improving, and ultimately reached a positive $F = 0.88$ and negative $F = 0.87$.

In order to better determine the classification threshold, this paper at -0.5--0.5 range, the threshold value $F$ of positive and negative impact the value of the experiment. Experimental results are shown in Fig.2 (b). When the threshold is 0.02, positive and negative $F$ is up to 0.88 and 0.88 of the maximum.

To validate the robustness of the proposed algorithm and the domain adaptation, we experiment on the Hotel and Computer reviews corpus. The Hotel and Computer corpus are provided by Chen, the size of the hotel reviews corpus size is 1000 lines, we divided into the training data of 500 lines and test data of 500 lines; we divided the computer corpus of 2000 lines into the training data of 1000 lines and test data of 1000 lines.

In Fig.3 (a) and (b), the proposed algorithm achieved positive classification result $F = 0.92$, negative classification result $F = 0.91$ on hotel data and positive classification result $F = 0.85$, negative classification result $F = 0.87$ on computer data.

6. Conclusion

In this paper, we proposed a novel sentiment classification algorithm that integrates multiple dictionaries, including multiple emotion dictionaries and negative adverb modifiers dictionary. We apply multi-pattern string matching algorithm to classify the polarity of online reviews. We replaced the current common emotional dictionary that because of the different emotional words, the field, targets or object will be misclassified to different sentiment polarity in the context. The experimental results show that this method has achieved good results in different review areas.

References

[1] Turney, P. “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” In Proceedings of ACL, 2002.

[2] S.-M. Kim and E. Hovy.” Determining the sentiment of opinions, “ In Proceedings of COLING, 2004, pp.1367

[3] Hiroshi, N. Tetsuya and W. Hideo. “Deeper sentiment analysis using machine translation technology,” In Proceedings of COLING, 2004, pp.494.

[4] Kennedy, A. and Inkpen, D. “Sentiment classification of movie reviews using contextual valence shifters”. Computational Intelligence, 2006, pp. 110-125.

[5] Devitt and K. Ahmad. “Sentiment polarity identification in financial news: a cohesion-based approach,” In Proceedings of ACL2007, 2007, pp.984.

[6] McDonald, K. Hannan, T. Neylon, M. Wells and J. Reynar. “Structured models for fine-to-coarse sentiment analysis,” In Proceedings of ACL2007, 2007, pp.432.

[7] Blitzer, M. Dredze and F. Pereira. “Biographies, bollywood, boom-boxes and blenders: domain adaptation for sentiment classification,” In Proceedings of ACL2007, 2007, pp.440.

[8] Dharmapurikar, S. and Lockwood, J. “Fast and Scalable Pattern Matching for Content Filtering.” Proceedings of the 2005 ACM symposium on Architecture for networking and communications systems, 2005, pp.183-192