# Sentiment analysis survey and final report

Nasrin Baratali

December 2014

## 1 Introduction

Sentiment Analysis is defined as the automatic detection and classification of opinions expressed in text written in natural language (Wiegand et al., 2010). In semantic analysis task, typically the bag-of-words model is used to represent a text. However, the bag-of-words model cannot represent many aspect of semantic represent by texts. Recent studies have shown that there are several phenomena that invert, intensify or neutralize the polarity of an opinion expressed by words. These phenomena are categorized in four groups (Wiegand et al., 2010):

1. Negations: A polarity of a word is inverted by a negator. For example, in "I do *not* like this new Nokia model", the word *not* invert the polarity of the word *like*.

2. Diminisher/intesifier valence shifters: A word diminishes the polarity of another word. For example, in "I find the functionality of the new phone *less* practical", the word *less* diminishes the polarity of the word *practical*.

3. Discourse connectives: the more salient textual unit that affects the polarity of sentences can be identified with discourse connectives. For example, in "perhaps it is a great phone, *but* I fail to see why.", the phrase that the word *but* associated with dominates the polarity of the sentence.

4. Modals: In theory, the phone *should* have worked even under water.

Negation is one of the most important linguistic construction that can affect the sentiment of texts and typically is considered in sentiment analysis approaches in practice. For example, Jia et al. (2009) proposed an approach based on parse tree and their corresponding dependencies to determine the scope of negation. They used ML approach to show that modeling negation scope can improve sentiment classification. As another example, Kennedy and Inkpen (2006) evaluated a negation model which is almost identical to the one proposed by Polanyi and Zaenen (2006) in document-level polarity classification. In this works, they used some simple negation words and reversed the polarity of the word followed the negation words in the text.

Intensifiers such as the word *rather* in the expression 'rather efficient' act to weaken or strengthen the base valence of the term. Polanyi and Zaenen (2006) presented the model of contextual valence shifter. In this work, linguistic analysis were done in order to show how the base attitudinal valence of a lexical item can be modified by context. They concluded that by extending the term-counting method with contextual valence shifters, sentiment classification accuracy can be improved.

Polanyi and Zaenen (2006) combine the idea of diminishers/intensifiers and negations and proposed an approach to calculate the local interactions between those words in order to achieve the polarity of texts. In this work, a parser was used to determine which negations/intensifiers/diminishers apply to terms with a polarity, if any. Then, with respect to intensifiers/diminishers assigned to the terms, the polarities of terms are increased/decreased. In this model if a polar expression is negated, its polarity score is simply inverted. Final results showed that modeling contextual valence shifters, especially negations, can be very beneficial in the identification polarity of text.

Discourse connectives can guide a sentiment analysis system to identify more important textual unit of sentences. Wu and Qiu (2012) conducted several methods to incorporate discourse structure knowledge to the task of sentiment analysis. They automatically detect the most influential discourse relations and connectives and validate the effectiveness of using discourse relations in Chinese sentiment analysis.

Modal operators can affect sentiment of a text by setting up a context of possibility or necessity in texts and removing entities that do not necessarily reflect the author's attitude in an actual situation under discussion. Therefore, in computing an evaluation of the author's attitude, terms in a modal context should not be treated precisely as terms in a real context.

Generally, sentiment classification approaches are categorized into two groups:

1. Rule-based methods: A simple rule-based polarity classifiers or lexicon based approaches count the number of positive and negative polar expressions in a text and assigns the polarity type with the majority of polar expressions to texts. Note that these methods have the advantage that they do not require training data; thus it can be applied to any dataset where training data are not available.

2. Machine Learning (ML) methods: In this method, several features, such as the counts of polar expressions are defined. Then these features are supplied to a classifier and let the classifier to decide about the polarity of each text. One of the most common features in the ML methods are the words appear in a text. These models can be expanded to n-grams of texts (e.g. bigrams) in order to model syntactic structure.

Modeling the four above phenomena in ML approach needs to take certain points in consideration. Firstly, it is vital to model the negation scope with negation triggers. Socher et al. (2013) showed that solely modeling negation as a binary feature in sentences cannot improve results of sentiment analysis systems and negation triggers should be combined with negation scopes in order to enhance results.

Additionally, with ML algorithms it is more difficult to show the effect of valence shifters on the performance of a system, because they are already included, to some degree, in the basic classifier and can not leavee out from a feature-set. Even combinations of very simple feature such as unigrams can capture some aspects of the valence shifters.

Recently the sentiment studies have been expanded to exploit more syntactic structures of texts. For example, Shaikh et al. (2007) proposed a method to recognize sentiment at the sentence level by using dependencies. In this work, they performed dependency analysis on the words and output triplet(s) of subject, verb, and object according to each semantic verb frame of the input sentence. They also define some rules to assign contextual valences to each triple. Finally in the last step, they calculate the contextual valence of the whole sentence based on triples.

Joshi and Penstein-Rosé (2009) presented new features for ML approaches which is based on dependency. In this work, they try to generalize the dependency relations in order to avoid sparseness and help learning algorithm to learn a weight for a more general feature. For example, instead of considering 'amod-camera-great' as a feature they attempt to generalize it to 'amod-NN-great' feature.

Last but not least, subjectivity and objectivity are important in sentiment analysis. Subjectivity is defined as the linguistic expression of somebody's opinions, sentiments, emotions, evaluations, beliefs and speculations while objectivity is the expression of facts (Wiebe, 1994). Wilson et al. (2005) detected whether a segment of text, either a whole document or a sentence, is subjective or objective and remove all of objective text segments before start to classify the sentiment. This approach significantly improved the results of their system.

# 2 Common dataset in Sentiment analysis

Sentiment analysis has been expanded in different domains such as movie reviews, product reviews, and customer feedback reviews.

**Pang and Lee Moview reviews** is a data set of classified movie reviews prepared by Pang and Lee (2004). This data set contains 2,000 movie reviews: 1,000 positive and 1,000 negative. The reviews were originally collected from the Internet Movie Database (IMDb) archive rec.arts.movies.reviews. Their classification as positive or negative is automatically extracted from the ratings, as specified by the original reviewer.

Pang and Lee (2004) took movie reviews as data and showed that standard machine learning techniques definitively outperform human-produced baseline. In addition, they found that SVMs generally gave better results than other classifiers.

**Stanford Sentiment Treebank** is a corpus of fully labeled parse trees. This corpus consists of 11,855 single sentences extracted from movie reviews.It

was parsed with the Stanford parser and includes a total of 215,154 unique phrases from those parse trees, each annotated by 3 human judges.
Since *semantic word spaces* or *Semantic vector spaces for single words* cannot capture the meaning of longer phrases properly, *compositionality in semantic vector spaces* received a lot of attention. For this purpose it was needed a large and labeled compositionality resource, and a model to capture the underlying phenomena in such data. To address this need, Socher et al. (2013) introduce the **Stanford Sentiment Treebank** and a powerful Recursive Neural Tensor Network that can accurately predict the compositional semantic effects present in this new corpus.

Another data for opinion mining and sentiment analysis are microblogs that Pak and Paroubek (2010) introduce them as an attractive source of data. In their work a large amount of tweets were collected automatically and used this corpus as a training corpus. They analysed the effect of POS-tags in sentiment polarity and showed authors use syntactic structures to describe emotions or state facts. They used multinomial Naiive Bayes classifier to determine positive, negative and neutral sentiment. The classifier uses N-grams and POS-tags as features.

**AFINN wordlist** is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive). The words have been manually labeled by Finn Årup Nielsen in 2009-2011. This word list is incorporated of 2477 words and phrases. Nielsen (2011)

**MPQA wordlist** contains of subjectivity clues were collected from a number of sources. Some were culled from manually developed resources. Others were identified automatically using both annotated and unannotated data.
Each clue is specified by five features, type(strongsubj or weaksubj), length of the clue in words, token or stem of the clue, part of speech of the clue, stemmed (yes or no), priorpolarity (positive, negative, both, neutral).
In our running MPQA list, We removed the clues with the POS of anypos, means this clue with these features are applicable for any part of speech.

# 3 Sentiment Analysis Project Report

The goal of this project is to find sentiment of each text by using a sentiment dictionary. Term-counting is the basic approach in sentiment classification in which simply summing up each words priorpolarity occured in the text by looking up into a sentiment dictionary.

## 3.1 Resources

In this project we used afinn as a sentiment dictionary. AFINN wordlist is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive). The words have been manually labeled by Finn Årup Nielsen in 2009-2011. This word list is incorporated of 2477 words and phrasesNielsen (2011).
Our used test benchmark is rt-polarity, which contains 5331 positive snippets and 5331 negative snippets and it is extracted From Rotten Tomatoes webpages.

## 3.2 Motivation

In traditional lexicon-based methods, all words and sentences are treated equally, ignoring the structural aspects of a text. As we explained before simple term-counting approach is a task of Bag of word model, which discards the order of the words and the syntactic relations between them. However, semantic is dependent to the syntactic structure of the text.

## 3.3 Natural Language Analsis Project Report

In our bottom-up viewpoint we first write a grammar for parsing the input sentences, then, by uploading EarleyChartParser attempt to parse the dataset sentences with the written grammar.
For creating the grammar, we started from general English grammar in James Allen book and improved it with some features for checking the agreement and verb forms. Meanwhile we added extra features for determining sentiment of each sentence.

For sentiment analysing, we added SENT feature to our grammar rules in which every constituent can take 3 value of Positive, Negative or Neutral. For this purpose, we first initialize SENT feature value of each lexical rules with its priorpolarity in Afinn sentiment dictionary. After initialization, we define more rules for propagating up the sentiment of each constituent and determine the sentiment of each sentence.
In the term sentiment initialization phase, for each term we look up its priorpolarity from AFFIN sentiment dictionary, if this value was greater than zero we assigned Positive to the SENT feature, if it was less than zero, the polarity value initialized to Negative, in the case it couldn't find the term in dictionary, its sentiment value got Neutral.

In language, some words such as 'like' can be used both as a verb or a preposition, however, words in different position have different strength in sentiment classification. So it is not reasonable to count such words as a positive term when it is using as a preposition. For this issue, we initialize all of the prepositions with Neutral sentiment value.
Proper nouns are one of the units which in our designed feature grammar carried neutral polarity. However, we know there is some proper nouns which carry positive or negative polarity. Like the name of a popular actor could hold posi-

tive priorpolarity, however, the name of country that a war happening on it is potential to hold negative priorpolarity. It is not far from access, if we construct our sentiment dictionary over a wide domain annotated dataset.

In this sentiment classification project, our major focus is on contrast connectives especially the term of 'but', and its influence on whole the sentence sentiment. Meanwhile, we consider other terms like negations and modals which affect the sentiment of higher level constituent in a sentence. We will explain more about the mentioned ideas as follow.

- Discourse Connectives
  (Wu and Qiu, 2012) stated that the overall sentiment of a text is critically affected by its discourse structure. They make use of discourse structure knowledge to improve sentiment classification and improve the baseline. According to this study discourse relations split texts into spans with different semantic relations. It showed that 'but', which denotes a contrast relation, is one of the most influential connectives. it also proved that the span introduced by connective 'but' has higher degree of importance.

  In our sentiment analysis project we are inspired by this idea and added a rule to our designed feature grammar.

$$S[SENT = s2] \rightarrow S[SENT = s1]CCS[SENT = s2]$$
$$CC \rightarrow \text{“}but\text{”}$$

  According to these rules, we take sentiment of the second part of the text followed the contrast connective, for a sentence parsed with this rule.

- Modals Modal is one of the mechanism to convey irrealis in English and it has the potential to influence the contextual polarity. In this work our preferences for handling modality is to neutralize the sentiment value of the phrase in the scope of the modals. In this purpose, we define a feature named TRIGGER which can take the value of Modal. So, we initialize our modal lexicon with TRIGGER feature valued Modal, and every time this modal terms occur in the text we do neutralize the sentiment of its following phrase. This is the rules we added for handling modality issue:

$$\begin{aligned} &VP[SENT = Neutral, AGR = ?a] \rightarrow \\ &\qquad MD[TRIGGER = Modal, SENT = Neutral, AGR = ?a] \\ &\qquad VP[VFORM = base] \end{aligned}$$

- Negation
  Negation is one the critical structure in sentiment classification. In this work, we determine the modals which are negative(e.g. wont, cannot) and reverse the sentiment value of the direct phrase followed them. For this

purpose we added following rules to our grammar:

$$
\begin{aligned}
&VP[SENT = Negative, AGR = ?a] \rightarrow \\
&\qquad MD[TRIGGER = Negation, SENT = Negative, AGR = ?a] \\
&\qquad VP[SENT = Positive, VFORM = base] \\
&VP[SENT = Positive, AGR = ?a] \rightarrow \\
&\qquad MD[TRIGGER = Negation, SENT = Negative, AGR = ?a] \\
&\qquad VP[SENT = Negative, VFORM = base] \\
&VP[SENT = Neutral, AGR = ?a] \rightarrow \\
&\qquad MD[TRIGGER = Negation, SENT = Negative, AGR = ?a] \\
&\qquad VP[SENT = Neutral, VFORM = base]
\end{aligned}
$$

After initializing lexical sentiment feature and making rules for handling contrast connectives, negation and modality, we define some default behaviour for classifying the sentiment of each sentence during propagating up the sentiment of each constituent in the sentence.

According to this fact that verbs are the head of a sentence, we give our highest priority to the main verb in a sentence. For example, if it carries a positive sentiment or negative sentiment it is reasonable to define whole sentence sentiment from it. So, for each sentence we first look at the verb phrase and then verb of this phrase. If the verb is neutral, we give the second priority to the adjectives. in the language people try to state their feeling in the adjective words. According to our knowledge adjective phrases are most potential unit after verb to defining a sentence sentiment. The next candidate units for determining sentiment of a sentence are nouns.

In summary we have following prefrences for assigning sentiment to each consituent.

$$VP >> ADJP >> NP >> PP >> ADVP$$

$$verb >> adjective >> noun >> adverb$$

You can find our designed grammar in 'newGrammar.fcfg' in the root directory of this project.

### 3.4 Evaluation

After running ASSP baseline on the rt-polarity dataset, we found 1968 of negative sentences recognized as positive and 813 positives sentences recognized as negative. We ran EarleyChartParser initialized with our designed feature grammar over the negative sentences and successfully parsed fifteen sentences contained our specific discourse connective, which is but.

You can find full parsed tree sentences in parseTreesAll folder or parseTrees folder in the root directory, in which it is properly shown in the parse tree that how each constituent takes their sentiment and how the sentiments value porpagating up in the parse tree.

According to our studies and observations in this project, we found discourse connectives, modals and negations are happening frequently in the texts, so taking them into consideration and presenting ideas for deal with them can improve the sentiment classification task significantly. Additionally, it is critical to consider syntactic structure for sentiment classification and presenting the ideas to determine the sentiment of a text by compositioning the polarity of simple constituents carrying polarity.

# References


Lifeng Jia, Clement Yu, and Weiyi Meng. The effect of negation on sentiment analysis and retrieval effectiveness. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 1827–1830. ACM, 2009.

Mahesh Joshi and Carolyn Penstein-Rosé. Generalizing dependency features for opinion mining. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, pages 313–316. Association for Computational Linguistics, 2009.

Alistair Kennedy and Diana Inkpen. Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2):110–125, 2006.

Finn Nielsen. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*, 2011.

Alexander Pak and Patrick Paroubek. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *LREC*, 2010.

Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, page 271. Association for Computational Linguistics, 2004.

Livia Polanyi and Annie Zaenen. Contextual valence shifters. In *Computing attitude and affect in text: Theory and applications*, page 1–10. Springer, 2006.

Mostafa Al Masum Shaikh, Helmut Prendinger, and Ishizuka Mitsuru. Assessing sentiment of text by semantic dependency and contextual valence analysis. In *Affective Computing and Intelligent Interaction*, pages 191–202. Springer, 2007.

Richard Socher, Alex Perelygin, Jean Y. Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, page 1631–1642. Citeseer, 2013.

Michael Wiegand, Alexandra Balahur, Benjamin Roth, Dietrich Klakow, and Andrés Montoyo. A survey on the role of negation in sentiment analysis. In *Proceedings of the workshop on negation and speculation in natural language processing*, pages 60–68. Association for Computational Linguistics, 2010.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics, 2005.

Fei Wang1 Yunfang Wu and Likun Qiu. Exploiting discourse relations for sentiment analysis. In *24th International Conference on Computational Linguistics*, page 1311. Citeseer, 2012.