
Sentiment Analysis and Opinion Mining

Lei Zhang¹ and Bing Liu²

¹LinkedIn, San Francisco, CA, USA

²University of Illinois at Chicago, Chicago, IL, USA

Abstract

With the rapid growth of social media, sentiment analysis, also called opinion mining, has become one of the most active research areas in natural language processing. Its application is also widespread, from business services to political campaigns. This article gives an introduction to this important area and presents some recent developments.

Definition

Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, appraisals, attitudes, and emotions toward entities and their aspects expressed in text.

Motivation and Background

Sentiment and opinion and their related concepts, such as evaluation, appraisal, attitude, affect, emotion, and mood, are about people's sub-

jective beliefs and feelings. They are key influencers of human behaviors. Whenever we need to make a decision, we often seek out others' opinions. This is true for both individuals and organizations.

The development of sentiment analysis coincides with the growth of *social media* (i.e., reviews, forum discussions, and blogs) on the Web. For the first time in human history, we now possess a huge volume of opinion data recorded in digital forms. These *user-generated contents* (UGC) are full of people's opinions. Mining useful knowledge from these corpora gives rise to the task of sentiment analysis. Since the early 2000s, it has been one of the most active research areas in natural language processing (NLP) (Pang and Lee 2008; Liu 2012). The research and applications have also spread from computer science to management science and social sciences because of its importance to business and society as a whole. Sentiment analysis techniques have been widely applied in practice, from business services to political campaigns.

Structure of the Task

In a nutshell, the task of sentiment analysis is to mine people's opinions and emotions from text. The term *opinion* is used as a concept represented with a quadruple (s, g, h, t) covering four components (Liu 2012): sentiment orientation s , sentiment target g , opinion holder h , and time t . *Sentiment* is the underlying feeling,

attitude, evaluation, or emotion associated with an opinion. Sentiment orientation can be *positive*, *negative*, or *neutral*. *Sentiment target*, also known as the *opinion target*, is an *entity* or an *aspect* of the *entity* that the sentiment has been expressed upon. *Opinion holder* is an individual or organization that holds an opinion. *Time* is when the opinion is expressed. We will discuss emotion specifically later.

We use the following camera review as an example (an ID number is associated with each sentence for easy reference):

Posted by John Smith
Date: September 10, 2011

(1) *I bought a Canon G12 camera six months ago.* (2) *I simply love it.* (3) *The picture quality is amazing.* (4) *The battery life is also long.* (5) *However, my wife thinks it is too heavy for her.*

Given the review, the task of sentiment analysis aims to extract the following opinion quadruples from sentences 2, 3, 4, and 5, respectively:

- (positive, Canon G12 camera, author, 2011/09/10)
- (positive, picture quality of Canon G12 camera, author, 2011/09/10)
- (positive, battery life of Canon G12 camera, author, 2011/09/10)
- (negative, weight of Canon G12 camera, author's wife, 2011/09/10)

The opinion target can be an entity (*Canon G12 camera*) or an aspect of the entity (*picture quality*, *battery life*, and *weight* of the Canon G12 camera). An aspect can be *explicit* (e.g., *battery life*) or *implicit* (e.g., *weight* is indicated by *heavy*) (Hu and Liu 2004).

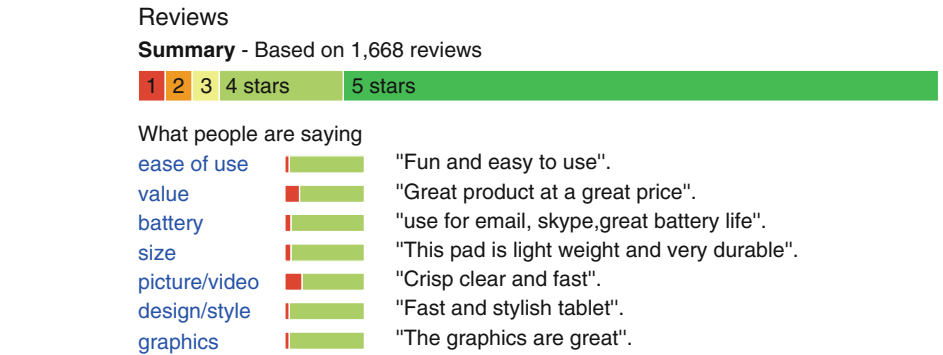
In many applications, it is useful to decompose opinion target to entity and aspect for more fine-grained analysis. Then, the above quadruples become the following quintuples, where GENERAL represents the entity itself (Liu 2012):

- (positive, Canon G12 camera, GENERAL, author, 2011/09/10)
- (positive, Canon G12 camera picture quality, author, 2011/09/10)
- (positive, Canon G12 camera, battery life, author, 2011/09/10)
- (negative, Canon G12 camera, weight, author's wife, 2011/09/10)

An opinion from a single opinion holder is usually not actionable in an opinion mining application. The user often needs opinions from a large number of opinion holders, which leads to opinion summary. A summary of opinions is normally constructed based on positive and negative sentiments about opinion targets, which is called aspect-based opinion summary (or feature-based opinion summary) (Hu and Liu 2004). Figure 1 shows an opinion summary generated from product reviews of Apple iPad by Google products. Generally, opinion summary needs to be quantitative, which is reflected by the proportions or the numbers of positive and negative opinions for each sentiment target or aspect.

Sentiment Analysis Methods

Researchers have studied sentiment analysis at three main levels of granularities: document, sentence, and aspect levels.



Sentiment Analysis and Opinion Mining, Fig. 1 Opinion summary for iPad from Google products

Document Sentiment Classification

Document sentiment classification classifies an opinion document (e.g., a product review) as expressing a positive or negative sentiment. It does not study or extract any information within the document. The task is also known as the *document-level sentiment classification*.

Document sentiment classification is commonly formulated as a supervised learning problem with two classes (*positive* and *negative*) or rating scores (e.g., 1–5 stars). Standard supervised learning methods such naïve Bayesian classification and support vector machines (SVM) can be applied for classification directly. Pang et al. (2002) first adopted those classification methods to classify movie reviews into two classes. Since this work, numerous other works have been reported. Like most supervised learning approaches, the main task of these works is to engineer a set of effective features. See Liu (2012) for an overview of this line of research.

There are also unsupervised approaches to document sentiment classification, which are mainly based on sentiment words and language patterns. It is quite clear that *sentiment words* (also called *opinion words*) that indicate positive or negative sentiments (e.g., *good* and *nice* are positive sentiment words, and *horrible* and *bad* are negative sentiment words) play an important role in sentiment classification (Turney 2002; Hu and Liu 2004; Kim and Hovy 2004). Turney (2002) proposed an unsupervised approach based on syntactic opinion patterns and Web search. Taboada et al. (2011) adapted a *lexicon-based approach* for document sentiment classification. It basically uses a set of sentiment words and phrases with appropriate scores and an aggregation scheme to aggregate the scores of the sentiment words appeared in a document to perform the classification. The lexicon-based approach was originally proposed for aspect-level and sentence-level sentiment classification (Hu and Liu 2004; Kim and Hovy 2004).

Researchers found that supervised sentiment classification is domain-sensitive, that is, a classifier trained using opinion documents from one domain performs poorly when it is applied

or tested on opinion documents from another domain. The reason is that words used in different domains for expressing opinions can be different. Furthermore, the same word may mean positive in one domain but negative in another domain. Domain adaptation or transfer learning techniques have been employed to address the problem (Blitzer et al. 2007; Pan et al. 2010; Liu 2012).

Another interesting topic is cross-language sentiment classification, which focuses on using the extensive resources and tools available in English and automated translation to help build sentiment classifiers in other languages with few resources or tools (Wan 2009; Mihalcea et al. 2007). Existing research proposed three main strategies: (1) translate test documents in the target language into the source language and classify them using a source language classifier, (2) translate a source language training corpus into the target language and build a classifier in the target language, and (3) translate a sentiment lexicon in the source language to the target language and build a lexicon-based classifier in the target language.

Sentence Sentiment Classification

Sentence sentiment classification is similar to document sentiment classification as sentences can be regarded as short documents. However, sentence classification is usually harder because the information contained in a typical sentence is much less than that contained in a typical document. Furthermore, sentence sentiment classification needs to consider the neutral class (or no opinion) because there are many factual sentences that express no positive or negative opinion in an opinion document. Document classification normally does not consider the neutral class.

Document sentiment classification techniques can be naturally applied for sentence sentiment classification. Some sentence-specific approaches have also been proposed, e.g., hierarchical sequence learning model (McDonald et al. 2007) and deep learning methods (Socher et al. 2013). In addition, researchers found that

different types of sentences may need different kinds of classification methods, e.g., *conditional sentences* and *integrative sentences* (Liu 2012).

For example, a conditional sentence describes implications or hypothetical situations and their consequences. Such a sentence typically contains two clauses that are dependent on each other: the condition clause and the consequent clause. Their relationship has significant impact on whether the sentence expresses a positive or negative sentiment (Narayanan et al. 2009). For example, the sentence *If someone makes a reliable car, I will buy it* expresses no sentiment toward any particular car, although it contains the positive sentiment word *reliable*. In Narayanan et al. (2009), supervised learning was used to deal with the problem using a set of linguistic features, e.g., sentiment words or phrases and their locations, part-of-speech tags of sentiment words, tense patterns, and conditional connectives.

Another type of difficult sentences is the *sarcasm sentences*. Sarcasm is a sophisticated form of speech act in which the speakers or the writers say or write the opposite of what they mean. In the context of sentiment analysis, it means that when one says something positive, one actually means negative, and vice versa. Sarcastic sentences are very difficult to deal with because commonsense knowledge and discourse analysis are often required to recognize them (Tsur et al. 2010; Riloff et al. 2013).

At the sentence level, another popular research problem is to identify *subjectivity* and *objective sentences*. Subjective expressions express opinions, appraisals, evaluations, allegations, desires, beliefs, suspicions, speculations, or stances (Wiebe et al. 2004). Some of these concepts indicate positive or negative sentiments. Some of them do not, e.g., *I want to buy a camera that can take good photos* which is a subjective sentence but does not express a positive or negative sentiment about anything. Objective sentences state facts. However, we should note that objective sentences can imply positive or negative sentiments of their authors because there are desirable facts and undesirable facts (Zhang and Liu 2011). For example, the sentence *I bought the mattress a week ago and a*

valley has formed in the middle states a fact, but the fact is undesirable. It thus implies a negative opinion about the quality of the mattress.

Aspect Sentiment Classification

Aspect-level classification classifies or determines sentiment on individual targets, which both the document-level and the sentence-level classification do not do because no sentiment target is involved at these two coarse levels of analysis. However, in applications, one often needs to know opinion targets. Without targets, any positive or negative sentiment is of limited use. For example, the sentence *trying out Chrome because Firefox works poorly* expresses a negative sentiment. But if we do not know that the negative sentiment is toward Firefox, not Chrome, the sentiment is of little use and can even be misleading. Many sentences also have mixed sentiments, e.g., *The performance of the car is great but the price is too high*. Aspect sentiment classification should find the opinion on the *performance* aspect of the car to be positive and the opinion on the *price* aspect of the car to be negative. In short, aspect sentiment classification determines sentiments expressed on entities and aspects of entities, which gives more useful information than document or sentence classification.

Although supervised learning can be applied, the kinds of features used for document and sentence sentiment classification are no longer sufficient or appropriate. The key reason is that those features do not consider (or are independent of) opinion targets and are thus unable to determine to which target an opinion refers. To remedy this problem, opinion target needs to be considered in learning. Two kinds of approaches have been proposed. The first one is to generate a set of features that are dependent on each opinion target in the sentence, e.g., weighing features based on their distances to a target. The second approach is to check the application scope of each sentiment expression to determine whether it covers the target in the sentence. For example, in the sentence *Apple is doing very well in this bad economy*, the sentiment word *bad's* application

scope covers only *economy*, not *Apple*. Current supervised learning methods mainly use the first approach but also have a flavor of the second approach (Jiang et al. 2011).

The lexicon-based approach can be employed as well. It computes the sentiment orientation on a target in a sentence by using a *sentiment aggregation function* that takes into account the distances of the sentiment expressions (sentiment words or phrases) and the opinion target in the sentence and/or by exploiting syntactic relationships of sentiment expressions and opinion targets to find the *application scope* of each sentiment expression. At the high level, the lexicon-based approach works as follows: it uses (1) a lexicon of sentiment expressions including sentiment words, phrases, idioms, and composition rules, (2) a set of rules for handling different language constructs (e.g., sentiment shifters and but-clauses) and different types of sentences, and (3) a sentiment aggregation function or a set of sentiment and target relationships derived from the parse tree to determine the sentiment orientation on each target (Ding et al. 2008; Liu 2012).

Comparative Sentences

Unlike a regular opinion sentence, a comparative sentence expresses a relation based on similarities or differences of more than one entity. In English, comparisons are usually conveyed using the *comparative* or *superlative* forms of adjectives or adverbs, e.g., *The picture quality of Canon cameras is better than that of Sony cameras*. To mine comparative opinions, aspect sentiment analysis is necessary because it does not make much sense to classify a comparative sentence as expressing a positive, negative, or neutral sentiment. See Jindal and Liu (2006) and Liu (2015) for more details.

Supervised Learning vs. Lexicon-Based Approach

The key advantage of supervised learning for sentiment classification is that it can automatically

learn from all kinds of features for classification through optimization. Most of these features are difficult to use by a lexicon-based method. However, supervised learning depends on the training data, which needs to be manually labeled for each domain. A shortcoming of the approach is that a supervised classifier trained from the labeled data in one domain often does not work in another domain. Thus, for each domain, new training data needs to be labeled, which is time consuming. Another shortcoming is that it is hard to learn things that do not occur frequently. The lexicon-based approach is able to avoid these issues to some extent and has been shown to perform well in a large number of applications. Its main advantage is domain independence, that is, it can be applied to any domain without manual labeling of a large amount of training data as required in the supervised learning approach. The lexicon-based method is also flexible in the sense that the system can be easily extended and improved. If an error occurs, the user simply corrects some existing rules and/or adds new rules to the system's rule base. However, the lexicon-based approach also has its disadvantages. It needs heavy investments in time and effort to build the initial knowledge base of lexicon, patterns, and rules. Furthermore, although the lexicon-based approach is supposed to be domain independent, some additional work is still needed to take care of the idiosyncrasies of each domain. The main issue is that it is quite hard to deal with domain-dependent or context-dependent sentiment words and phrases (see below and Liu 2015).

Aspect and Entity Extraction

The task of aspect and entity extraction is to identify and extract opinion targets (aspect or entity) from opinion documents. Since aspect extraction and entity extraction are closely related tasks, ideas and methods proposed for aspect extractions can also be shared with the entity extraction task. Much of the existing research focused on aspect extraction. Current aspect extraction methods can be roughly grouped into four categories: mining frequent noun phrases, utilizing syntactic

relations of sentiment words and their targets, and applying supervised sequence learning models and topic modeling. All these approaches are used in practice.

Finding Frequent Noun Phrases

Since people often use the same words when they comment on the same product aspects, Hu and Liu (2004) makes use of this observation to mine aspects by simply finding frequent nouns and noun phrases in reviews using frequent item-set mining (Agrawal and Srikant 1994). Those more frequent noun phrases are also likely to be more important aspects because people usually comment on those more important aspects more frequently.

Exploiting Syntactic Relations of Sentiment and Target

It was observed in Hu and Liu (2004) that adjective sentiment words often modify (or describes) noun aspects (e.g., *great picture*). Hu and Liu (2004) used such relations to identify aspects that are hard to find by the frequency-based method above. Zhuang et al. (2006) formulated the idea based on the dependency grammar and extracted aspect and sentiment word pairs from movie reviews using a set of dependency relations. Qiu et al. (2011) developed the idea further and proposed an algorithm called *double propagation* (DP). DP uses a set of manually compiled dependency rules derived from some dependency relations to identify both aspects and sentiment words simultaneously through a bootstrapping process. These methods are all based on the idea that opinion always has target, and there are often syntactic relations that connect sentiment words and targets in a sentence. Thus, sentiment words can be recognized by identified aspects, and aspects can be identified by known sentiment words. The extracted sentiment words and aspects are utilized to identify new sentiment words and new aspects, which are used again to extract more sentiment words and aspects. This

is the approach used in the DP method. Recently, this method was improved with automated rule selection (Liu et al. 2015) and word alignment from the machine translation research (Liu et al. 2013).

Applying Supervised Sequence Learning Models

Sequence learning models such as Hidden Markov models (HMM) and conditional random fields (CRF) are widely used in information extraction. They are thus also used for aspect extraction. Aspect extraction can be regarded as a sequence labeling task since entity, aspect, and opinion expressions are often interdependent and occur in a sequence in a sentence. Jin and Ho (2009) utilized lexicalized HMM to extract product aspects and opinion expressions from reviews. Different from traditional HMM, they integrated linguistic features such as part of speech and lexical patterns into HMM. Jakob and Gurevych (2010) utilized CRF to extract opinion aspects from opinion sentences.

Topic Modeling

Topic models such as PLSA (probabilistic latent semantic analysis) and LDA (latent Dirichlet allocation) have been popularly used to mine hidden topics from the document corpora. In the context of aspect extraction, aspects are basically topics in topic modeling. Mei et al. (2007) proposed a model for extracting both aspects and sentiment words. Titov and McDonald (2008) pointed out that global topic models such as PLSA and LDA might not be suitable for detecting aspects from reviews. To tackle this problem, they proposed some multigrain topic models to discover aspects, which models two distinct types of topics: global topics and local topics. Lin and He (2009) proposed a joint topic-sentiment model, which extended LDA by adding a sentiment layer. It detects sentiment and aspect simultaneously from the corpus. Further works along a similar line have been done in Brody and

Elhadad (2010), Wang et al. (2010), Zhao et al. (2010), and Jo and Oh (2011). Recently, two new types of models were proposed: *knowledge-based models* (Mukherjee and Liu 2012) which can exploit prior domain knowledge to produce better results and *lifelong topic models* (Chen and Liu 2014) which exploit the big data to automatically mine prior knowledge to be used in the modeling process.

Sentiment Lexicon

It is quite clear that sentiment words are instrumental for sentiment analysis. Positive sentiment words are used to express some desired states, while negative ones are used to express some undesired states. Examples of positive sentiment words are *beautiful*, *wonderful*, and *good*. Examples of negative sentiment words are *bad*, *poor*, and *terrible*. Apart from individual words, there are also sentiment phrases and idioms. To compile a sentiment word list or lexicon, two approaches have been studied: dictionary-based approach and corpus-based approach.

Dictionary-Based Approach

This approach is based on bootstrapping using a small set of seed sentiment words and an online dictionary, e.g., WordNet or thesaurus. The strategy is to first collect a small set of sentiment words manually with known orientations and then to grow this set by searching in the WordNet or an online thesaurus to find their synonyms and antonyms. The newly found words are added to the seed list. The next iteration starts. The iterative process stops when no more new words are found (Hu and Liu 2004; Kim and Hovy 2004; Kamps et al. 2004).

Corpus-Based Approach

The corpus-based approach relies on syntactic patterns and also a seed list of sentiment words to find other sentiment words in a large corpus.

One of the key ideas was proposed in Hatzivassiloglou and McKeown (1997). The technique uses the set of seed sentiment words and a set of linguistic constraints or conventions on connectives to identify additional sentiment words and their orientations. One of the constraints is about the conjunction AND, which says that conjoined adjectives usually have the same sentiment orientation. For example, if *beautiful* is known to be positive, we can infer that *spacious* is also positive from the sentence *This car is beautiful and spacious*. Rules or constraints are also designed for other connectives, OR, BUT, EITHER-OR, and NEITHER-NOR. This constraint is called *sentiment consistency*. Kanayama and Nasukawa (2006) and Ding et al. (2008) expanded this approach to intra-sentential and inter-sentential sentiment consistency. Ding et al. (2008) further showed that the same word may indicate positive sentiment in one sentence context but negative sentiment in another sentence context. For example, in the domain of car reviews, the word “quiet” expresses opposite sentiments or opinions in the following two sentences: *This car is very quiet* (positive) and *The audio system in the car is very quiet* (negative). The authors proposed to consider both the sentiment word and the aspect together in determining the sentiment orientation of the sentiment word. To determine the sentiment orientation of the pair, the above sentiment consistency idea is still used. In a similar vein, Choi and Cardie (2009) studied the problem of adapting a general-purpose sentiment lexicon to a specific domain of application.

Sentiment Analysis of Emotions

Emotions are human feelings. They are similar and also different from opinions. An opinion expresses an evaluation or appraisal about some objects, whereas an emotion expresses a human inner feeling. Human beings have many different types of emotions. However, there is still no agreement among researchers on how many kinds of emotions there are and what they are. According to Parrott (2001), humans have six basic emotions: joy, love, anger, fear,

sadness, and surprise. Existing sentiment analysis of emotions is focused on classification of emotion types expressed in sentences. Both supervised learning and lexicon-based approaches have been attempted by researchers.

In supervised learning, Alm et al. (2005) classified the emotional affinity of sentences in the narrative domain of children's fairy tales. The features are not the traditional word n-grams but fourteen groups of Boolean features about each sentence and its context in the document. The classes are only two: neutral and emotional. In Mohammad (2012), a Twitter data set was annotated with emotion types based on emotion words or hashtags in Twitter posts. The author then performed classification of emotions using SVM with binary features that capture the presence or absence of unigrams and bigrams. Additional references can be found in Liu (2015).

In the lexicon-based approach, Yang et al. (2007) first constructed an emotion lexicon and then performed emotion classification at the sentence level using the lexicon. To construct the emotion lexicon, the proposed algorithm uses only sentences with a single user-provided emoticon. For each word, it computes a collocation (or association) strength of the word with each emoticon using a measure similar to pointwise mutual information (PMI). Those top-scoring words are very likely to indicate different types of emotions. For emotion classification of sentences, it experimented with two approaches: the lexicon-based approach and the supervised learning approach. For supervised learning, only the top k emotion words were used as features.

Summary

This article gave a brief introduction to sentiment analysis. Interested readers can refer to Liu (2015) for an in-depth and comprehensive coverage of the topic. Sentiment analysis is a highly challenging research problem with almost unlimited applications. It has been one of the most active research areas in natural language processing for many years. Although significant progresses have been made and numerous in-

dustrial systems have been built, the problem remains to be very difficult. The accuracy results in many cases are still unsatisfactory. However, the practical application needs and technical challenges will keep the field vibrant and lively for years to come.

Recommended Reading

- Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Proceedings of the international conference on very large databases (VLDB-1994), Santiago de Chile
- Alm CO, Roth D, Sproat R (2005) Emotions from text: machine learning for text-based emotion prediction. In: Proceedings of conference on human language technology and empirical methods in natural language processing (HLT/EMP/NLP-2005), Vancouver
- Blitzer J, Dredze M, Pereira F (2007) Biographies, Bollywood, boom-boxes and blenders: domain adaptation for sentiment classification. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2007), Prague
- Brody S, Elhadad S (2010) An unsupervised aspect-sentiment model for online reviews. In: Proceedings of the annual conference of the North American chapter of the ACL (NAACL-2010), Los Angeles
- Chen Z, Liu B (2014) Topic modeling using topics from many domains, lifelong learning and big data. In: Proceedings of the international conference on machine learning (ICML-2014), Beijing
- Choi Y, Cardie C (2009) Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2009), Singapore
- Ding X, Liu B, Yu PS (2008) A holistic lexicon-based approach to opinion mining. In: Proceedings of the conference on web search and web data mining (WSDM-2008), Palo Alto
- Hatzivassiloglou V, McKeown K (1997) Predicting the semantic orientation of adjectives. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-1997), Madrid
- Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD-2004), Seattle
- Jakob N, Gurevych I (2010) Extracting opinion targets in a single- and cross-domain setting with conditional random fields. In: Proceedings of the conference on empirical methods in natural language processing (EMNLP-2010), MIT, Massachusetts
- Jiang L, Yu M, Zhou M, Liu X, Zhao T (2011) Target-dependent twitter sentiment classification. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2011), Portland

- Jin W, Ho HH (2009) A novel lexicalized HMM-Based learning framework for web opinion mining. In: Proceedings of international conference on machine learning (ICML-2009), Montreal
- Jindal N, Liu B (2006) Mining comparative sentences and relations. In: Proceedings of national conference on artificial intelligence (AAAI-2006), Boston
- Jo Y, Oh A (2011) Aspect and sentiment unification model for online review analysis. In: Proceedings of the conference on web search and web data mining (WSDM-2011), Hong Kong
- Kamps J, Marx M, Mokken RJ, De Rijke M (2004) Using WordNet to measure semantic orientation of adjectives. In: Proceedings of international conference on language resources and evaluation (LREC-2004), Lisbon
- Kanayama H, Sukawa T (2006) Fully automatic lexicon expansion for domain-oriented sentiment analysis. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2006), Sydney
- Kim S-M, Hovy E (2004) Determining the sentiment of opinions. In: Proceedings of international conference on computational linguistics (COLING-2004), University of Geneva, Geneva
- Li S, Lin C, Song Y, Li Z (2010) Comparable entity mining from comparative questions. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2010), Uppsala University, Uppsala
- Lin C, He Y (2009) Joint sentiment/topic model for sentiment analysis. In: Proceedings of ACM international conference on information and knowledge management (CIKM-2009), Hong Kong
- Liu B (2012) Sentiment analysis and opinion mining. Morgan & Claypool, San Rafael
- Liu B (2015) Sentiment analysis: mining opinions, sentiments, and emotions. Cambridge University Press, Cambridge
- Liu K, Xu L, Zhao J (2013) Syntactic patterns versus word alignment: extracting opinion targets from online reviews. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2013), Sofia
- Liu Q, Gao Z, Liu B, Zhang Y (2015) Automated rule selection for aspect extraction in opinion mining. In: Proceedings of international joint conference on artificial intelligence (IJCAI-2015), Buenos Aires
- McDonald R, Hannan K, Neylon T, Wells M, Reynar J (2007) Structured models for fine-to-coarse sentiment analysis. In: Proceedings of the annual meeting of the association for computational linguistics (ACL2007), Prague
- Mei Q, Ling X, Wondra M, Su H, Zhai C (2007) Topic sentiment mixture: modeling facets and opinions in weblogs. In: Proceedings of international conference on world wide web (WWW-2007), Banff
- Mihalcea R, Banea C, Wiebe J (2007) Learning multilingual subjective language via cross-lingual projections. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2007), Prague
- Mohammad SM (2012) #Emotional tweets. In: Proceedings of the first joint conference on lexical and computational semantics, Montreal
- Mukherjee A, Liu B (2012) Aspect extraction through semi-supervised modeling. In: Proceedings of the annual meeting of association for computational linguistics (ACL-2012), Jeju Island
- Narayanan R, Liu B, Choudhary A (2009) Sentiment analysis of conditional sentences. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2009), Singapore
- Pan SJ, Ni X, Sun J, Yang Q, Chen Z (2010) Cross-domain sentiment classification via spectral feature alignment. In: Proceedings of international conference on world wide web (WWW-2010), Raleigh
- Pang B, Lee L (2008) Opinion mining and sentiment analysis. Foundations and trends in information retrieval. Now Publishers, Hanover, MA
- Pang B, Lee L, Vaithyanathan S (2002) Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2002), Philadelphia
- Parrott WG (2001) Emotions in social psychology: essential readings. Psychology Press, Hove
- Qiu G, Liu B, Bu J, Chen C (2011) Opinion word expansion and target extraction through double propagation. *Comput Linguist*, 37(1):9–27
- Riloff E, Qadir A, Surve P, De Silva L, Gilbert N, Huang R (2013) Sarcasm as contrast between a positive sentiment and negative situation. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2013), Seattle
- Socher R, Perelygin A, Wu J, Manning C, Ng A, Chuang J (2013) Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the conference on empirical methods on natural language processing (EMNLP-2013), Seattle
- Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. *Comput Linguist* 37(2):267–307
- Titov I, McDonald R (2008) Modeling online reviews with multi-grain topic models. In: Proceedings of international conference on world wide web (WWW-2008), Beijing
- Tsur O, Davidov D, Rappoport A (2010) A great catchy name: semi-supervised recognition of sarcastic sentences in online product reviews. In: Proceedings of the international AAAI conference on weblogs and social media (ICWSM-2010), Washington, DC
- Turney PD (2002) Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2002), Philadelphia
- Wan X (2009) Co-training for cross-lingual sentiment classification. In: Proceedings of the annual meeting of the ACL and the IJCNLP of the AFNLP (ACL-IJCNLP-2009), Singapore
- Wang H, Lu Y, Zhai C (2010) Latent aspect rating analysis on review text data: a rating regression approach. In:

- Proceedings of ACM SIGKDD international conference on knowledge discovery and data mining (KDD-2010), Washington, DC
- Wiebe J, Wilson T, Bruce R, Bell M, Martin M (2004) Learning subjective language. *Comput Linguist* 30(3):277–308
- Yang C, Lin KH-Y, Chen H-H (2007) Building emotion lexicon from weblog corpora. In: Proceedings of the annual meeting of the ACL on interactive poster and demonstration sessions, Prague
- Zhang L, Liu B (2011) Identifying noun product features that imply opinions. In: Proceedings of the annual meeting of the association for computational linguistics (ACL-2011), Portland
- Zhao W, Jiang J, Yan H, Li X (2010) Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In: Proceedings of conference on empirical methods in natural language processing (EMNLP-2010). MIT, Massachusetts
- Zhuang L, Jing F, Zhu X (2006) Movie review mining and summarization. In: Proceedings of ACM international conference on information and knowledge management (CIKM-2006), Arlington