OPINION MINING

Bing Liu
Department of Computer Science
University of Illinois at Chicago
851 S. Morgan Street
Chicago, IL 60607-0753
liub@cs.uic.edu

SYNONYMS

Sentiment analysis

DEFINITION

Given a set of evaluative text documents D that contain opinions (or sentiments) about an object, opinion mining aims to extract attributes and components of the object that have been commented on in each document $d \in D$ and to determine whether the comments are positive, negative or neutral.

HISTORICAL BACKGROUND

Textual information in the world can be broadly classified into two main categories, *facts* and *opinions*. Facts are objective statements about entities and events in the world. Opinions are subjective statements that reflect people's sentiments or perceptions about the entities and events. Much of the existing research on text information processing has been (almost exclusively) focused on mining and retrieval of factual information, e.g., information retrieval, Web search, and many other text mining and natural language processing tasks. Little work has been done on the processing of opinions until only recently. Yet, opinions are so important that whenever one needs to make a decision one wants to hear others' opinions. This is not only true for individuals but also true for organizations.

One of the main reasons for the lack of study on opinions is that there was little opinionated text before the World Wide Web. Before the Web, when an individual needs to make a decision, he/she typically asks for opinions from friends and families. When an organization needs to find opinions of the general public about its products and services, it conducts surveys and focused groups. With the Web, especially with the explosive growth of the user generated content on the Web, the world has changed. One can post reviews of products at merchant sites and express views on almost anything in Internet forums, discussion groups, and blogs, which are collectively called the *user generated content*. Now if one wants to buy a product, it is no longer necessary to ask one's friends and families because there are plentiful of product reviews on the Web which give the opinions of the existing users of the product. For a company, it may no longer need to conduct surveys, to organize focused groups or to employ external consultants in order to find consumer opinions or sentiments about its products and those of its competitors.

Finding opinion sources and monitoring them on the Web, however, can still be a formidable task because a large number of diverse sources exist on the Web and each source also contains a huge volume of information. In many cases, opinions are hidden in long forum posts and blogs. It is very difficult for a human reader to find relevant sources, extract pertinent sentences, read them, summarize them and organize them into usable forms. An automated opinion mining and summarization system is thus needed. *Opinion mining*, also known as *sentiment analysis*, grows out of this need. This article introduces this research area. In particular, it discusses the following topics: (1) the abstract model of opinion mining, (2) sentiment classification, (3) feature-based opinion mining and summarization, and (4) opinion mining from comparative sentences.

Research on opinion mining started with identifying *opinion* (or *sentiment*) *bearing words*, e.g., great, amazing, wonderful, bad, and poor. Many researchers have worked on mining such words and identifying their *semantic orientations* (i.e., positive or negative). In [5], the authors identified several linguistic rules that can be exploited to identify opinion words and their orientations from a large corpus. This method has been applied, extended and improved in [3, 8, 12]. In [6, 9], a bootstrapping approach is proposed, which uses a small set of given seed opinion words to find their synonyms and antonyms in WordNet (http://wordnet.princeton.edu/). The next major development is sentiment classification of product reviews at the document level [2, 11, 13]. The objective of this task is to classify each review document as expressing a positive or a negative sentiment about an object (e.g., a movie, a camera, or a car). Several researchers also studied sentence-level sentiment classification [9, 14, 15], i.e., classifying each sentence as expressing a positive or a negative opinion. The model of feature-based opinion mining and summarization is proposed in [6, 10]. This model gives a more complete formulation of the opinion mining problem. It identifies the key pieces of information that should be mined and describes how a structured opinion summary can be produced from unstructured texts. The problem of mining opinions from comparative sentences is introduced in [4, 7].

SCIENTIFIC FUNDAMENTALS

Model of Opinion Mining

In general, opinions can be expressed on anything, e.g., a product, a service, a topic, an individual, an organization, or an event. The general term *object* is used to denote the entity that has been commented on. An object has a set of *components* (or *parts*) and a set of *attributes*. Each component may also have its sub-components and its set of attributes, and so on. Thus, the object can be hierarchically decomposed based on the *part-of* relationship.

Definition (**object**)**:** An *object* O is an entity which can be a product, topic, person, event, or organization. It is associated with a pair, O: (T, A), where T is a hierarchy or taxonomy of *components* (or *parts*) and *sub-components* of O, and A is a set of *attributes* of O. Each component has its own set of sub-components and attributes.

In this hierarchy or tree, the root is the object itself. Each non-root node is a component or sub-component of the object. Each link is a part-of relationship. Each node is associated with a set of attributes. An opinion can be expressed on any node and any attribute of the node.

However, for an ordinary user, it is probably too complex to use a hierarchical representation. To simplify it, the tree is flattened. The word "*features*" is used to represent both components and attributes. Using features for objects (especially products) is quite common in practice. Note that in this definition the object itself is also a feature, which is the root of the tree.

Let an evaluative document be d, which can be a product review, a forum post or a blog that evaluates a particular object O. In the most general case, d consists of a sequence of sentences $d = \langle s_1, s_2, \ldots, s_m \rangle$.

Definition (**opinion passage on a feature**)**:** The *opinion passage* on a feature f of the object O evaluated in d is a group of consecutive sentences in d that expresses a positive or negative opinion on f.

This means that it is possible that a sequence of sentences (at least one) together expresses an opinion on an object or a feature of the object. It is also possible that a single sentence expresses opinions on more than one feature, e.g., "The picture quality of this camera is good, but the battery life is short".

Definition (**opinion holder**)**:** The *holder* of a particular opinion is a person or an organization that holds the opinion.

In the case of product reviews, forum postings and blogs, opinion holders are usually the authors of the posts. Opinion holders are important in news articles because they often explicitly state the person or organization that holds a particular opinion [9]. For example, the opinion holder in the sentence "John expressed his disagreement on the treaty" is "John".

Digital_camera_1:

CAMERA:
    Positive: 125   <individual review sentences>
    Negative: 7   <individual review sentences>
Feature: **picture quality**
    Positive: 123   <individual review sentences>
    Negative: 6   <individual review sentences>
Feature: **size**
    Positive: 82   <individual review sentences>
    Negative: 10   <individual review sentences>
…

Fig. 1. An example of a feature-based summary of opinions

Definition (semantic orientation of an opinion): The *semantic orientation* of an opinion on a feature f states whether the opinion is positive, negative or neutral.

Putting things together, a *model* for an object and a set of opinions on the features of the object can be defined, which is called the *feature-based opinion mining model*.

Model of Feature-Based Opinion Mining: An object O is represented with a finite set of features, $F = \{f_1, f_2, \ldots, f_n\}$, which includes the object itself. Each feature $f_i \in F$ can be expressed with a finite set of words or phrases W_i, which are *synonyms*. That is, there is a set of corresponding synonym sets $W = \{W_1, W_2, \ldots, W_n\}$ for the n features. In an evaluative document d which evaluates object O, an opinion holder j comments on a subset of the features $S_j \subseteq F$. For each feature $f_k \in S_j$ that opinion holder j comments on, he/she chooses a word or phrase from W_k to describe the feature, and then expresses a positive, negative or neutral opinion on f_k. The opinion mining task is to discover all these hidden pieces of information from a given evaluative document d.

Mining output: Given an evaluative document d, the mining result is a set of quadruples. Each quadruple is denoted by (H, O, f, SO), where H is the opinion holder, O is the object, f is a feature of the object and SO is the semantic orientation of the opinion expressed on feature f in a sentence of d. Neutral opinions are ignored in the output as they are not usually useful.

Given a collection of evaluative documents D containing opinions on an object, three main technical problems can be identified (clearly there are more):

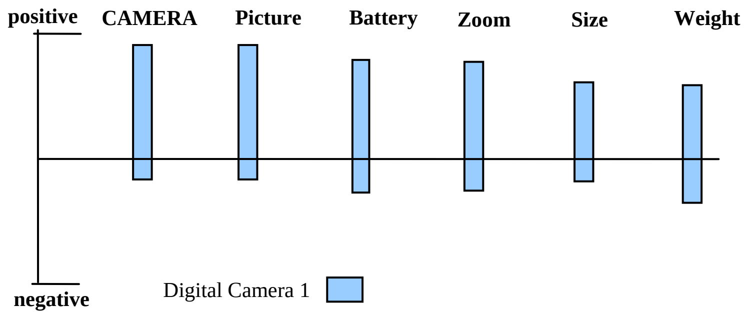
Problem 1: Extracting object features that have been commented on in each document $d \in D$.

Problem 2: Determining whether the opinions on the features are positive, negative or neutral.

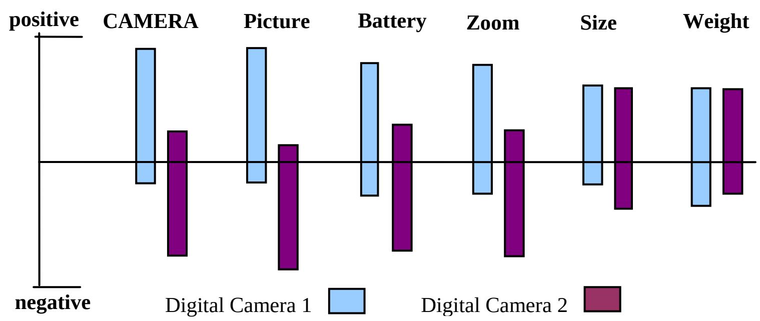
Problem 3: Grouping synonyms of features (as different opinion holders may use different words or phrase to express the same feature).

Opinion Summary: There are many ways to use the mining results. One simple way is to produce a *feature-based summary* of opinions on the object [6]. An example is used to illustrate what that means.

Fig. 1 summarizes the opinions in a set of reviews of a particular digital camera, *digital_camera_1*. The opinion holders are omitted. In the figure, “CAMERA” represents the camera itself (the root node of the object hierarchy). 125 reviews expressed positive opinions on the camera and 7 reviews expressed negative opinions on the camera. “picture quality” and “size” are two product features. 123 reviews expressed positive opinions on the picture quality, and only 6 reviews expressed negative opinions. The <individual review sentences> points to the specific sentences and/or the whole reviews that give the positive or negative comments about the feature. With such a summary, the user can easily see how existing customers feel about the digital camera. If he/she is very interested in a particular feature, he/she can drill down by following the <individual review sentences> link to see why existing customers like it and/or dislike it.



(A) Feature-based summary of opinions on a digital camera



(B) Opinion comparison of two digital cameras

Fig. 2. Visualization of feature-based opinion summary and comparison

The summary in Fig. 1 can be easily visualized using a bar chart [10]. Fig. 2(A) shows such a chart. In the figure, each bar above the *X*-axis gives the number of positive opinions on a feature (listed at the top), and the bar below the *X*-axis gives the number of negative opinions on the same feature. Obviously, other visualizations are also possible. For example, one may only show the percentage of positive (or negative) opinions on each feature. Comparing opinion summaries of a few competing objects is even more interesting [10]. Fig. 2(B) shows a visual comparison of consumer opinions on two competing digital cameras. One can clearly see how consumers view different features of each camera.

Sentiment Classification

Sentiment classification has been widely studied in the natural language processing (NLP) community [e.g., 2, 11, 13]. It is defined as follows: Given a set of evaluative documents D, it determines whether each document $d \in D$ expresses a positive or negative opinion (or sentiment) on an object. For example, given a set of movie reviews, the system classifies them into positive reviews and negative reviews.

This is clearly a classification learning problem. It is similar but also different from the classic topic-based text classification, which classifies documents into predefined topic classes, e.g., politics, sciences, and sports. In topic-based classification, topic related words are important. However, in sentiment classification, topic-related words are unimportant. Instead, opinion words that indicate positive or negative opinions are important, e.g., great, excellent, amazing, horrible, bad, worst, etc. There are many existing techniques. Most of them apply some forms of machine learning techniques for classification [e.g., 11]. Custom-designed algorithms specifically for sentiment classification also exist, which exploit opinion words and phrases together with some scoring functions [2, 13].

This classification is said to be at the document level as it treats each document as the basic information unit. Sentiment classification thus makes the following assumption: Each evaluative document (e.g., a review) focuses on a single object O and contains opinions of a single opinion holder. Since in the above opinion mining model an object *O* itself is also a feature (the root node of the object hierarchy), sentiment classification basically determines the semantic orientation of the opinion expressed on *O* in each evaluative document that satisfies the above assumption.

Apart from the document-level sentiment classification, researchers have also studied classification at the *sentence-level*, i.e., classifying each sentence as a subjective or objective sentence and/or as expressing a positive or negative opinion [9, 14, 15]. Like the document-level classification, the sentence-level sentiment classification does not consider object features that have been commented on in a sentence. Compound sentences are also an issue. Such a sentence often express more than one opinion, e.g., "The picture quality of this camera is amazing and so is the battery life, but the viewfinder is too small".

Feature-Based Opinion Mining

Classifying evaluative texts at the document level or the sentence level does not tell what the opinion holder likes and dislikes. A positive document on an object does not mean that the opinion holder has positive opinions on all aspects or features of the object. Likewise, a negative document does not mean that the opinion holder dislikes everything about the object. In an evaluative document (e.g., a product review), the opinion holder typically writes both positive and negative aspects of the object, although the general sentiment on the object may be positive or negative. To obtain such detailed aspects, going to the feature level is needed. Based on the model presented earlier, three key mining tasks are:

1. Identifying *object features*: For instance, in the sentence "The picture quality of this camera is amazing," the object feature is "picture quality". In [10], a supervised pattern mining method is proposed. In [6, 12], an unsupervised method is used. The technique basically finds frequent nouns and noun phrases as features, which are usually genuine features. Clearly, many information extraction techniques are also applicable, e.g., conditional random fields (CRF), hidden Markov models (HMM), and many others.
2. Determining opinion orientations*:* This task determines whether the opinions on the features are positive, negative or neutral. In the above sentence, the opinion on the feature "picture quality" is positive. Again, many approaches are possible. A lexicon-based approach has been shown to perform quite well in [3, 6]. The lexicon-based approach basically uses opinion words and phrases in a sentence to determine the orientation of an opinion on a feature. A relaxation labeling based approach is given in [12]. Clearly, various types of supervised learning are possible approaches as well.
3. Grouping synonyms: As the same object features can be expressed with different words or phrases, this task groups those synonyms together. Not much research has been done on this topic. See [1] for an attempt on this problem.

Mining Comparative and Superlative Sentences

Directly expressing positive or negative opinions on an object or its features is only one form of evaluation. Comparing the object with some other similar objects is another. Comparisons are related to but are also different from direct opinions. For example, a typical opinion sentence is "The picture quality of camera x is great." A typical comparison sentence is "The picture quality of camera x is better than that of camera y." In general, a comparative sentence expresses a relation based on similarities or differences of more than one object. In English, comparisons are usually conveyed using the *comparative* or the *superlative* forms of adjectives or adverbs. The structure of a comparative normally consists of the stem of an adjective or adverb, plus the suffix *-er*, or the modifier "more" or "less" before the adjective or adverb. The structure of a superlative normally consists of the stem of an adjective or adverb, plus the suffix *-est*, or the modifier "most" or "least" before the adjective or adverb. Mining of comparative sentences basically consists of identifying what features and objects are compared and which objected are preferred by their authors (opinion holders). Details can be found in [4, 7].

KEY APPLICATIONS

Opinions are so important that whenever one needs to make a decision, one wants to hear others' opinions. This is true for both individuals and organizations. The technology of opinion mining thus has a tremendous scope for practical applications.

Individual consumers: If an individual wants to purchase a product, it is useful to see a summary of opinions of existing users so that he/she can make an informed decision. This is better than reading a large number of reviews to form a mental picture of the strengths and weaknesses of the product. He/she can also compare the summaries of opinions of competing products, which is even more useful.

Organizations and businesses: Opinion mining is equally, if not even more, important to businesses and organizations. For example, it is critical for a product manufacturer to know how consumers perceive its products and those of its competitors. This information is not only useful for marketing and product benchmarking but also useful for product design and product developments.

CROSS REFERENCES

Text mining

RECOMMENDED READING

1. Carenini, G., Ng, R. and Zwart, E. Extracting Knowledge from Evaluative Text. Proceedings of the Third International Conference on Knowledge Capture (K-CAP'05), 2005.
2. Dave, D., Lawrence, A., and Pennock, D. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. *Proceedings of International World Wide Web Conference* (*WWW'03*), 2003.
3. Ding, X., Liu, B. and Yu, P. A Holistic Lexicon-Based Approach to Opinion Mining. *Proceedings* of *the first ACM International Conference on Web search and Data Mining* (*WSDM'08*), 2008.
4. Ganapathibhotla, G. and Liu, B. Identifying Preferred Entities in Comparative Sentences. To appear in *Proceedings of the 22nd International Conference on Computational Linguistics (COLING'08)*, 2008.
5. Hatzivassiloglou, V. and McKeown, K. Predicting the Semantic Orientation of Adjectives. *ACL-EACL'97*, 1997.
6. Hu, M and Liu, B. Mining and Summarizing Customer Reviews. *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (*KDD'04*), 2004.
7. Jindal, N. and Liu, B. Mining Comparative Sentences and Relations. *Proceedings of National Conference on Artificial Intelligence* (*AAAI'06*), 2006.
8. Kanayama, H. and Nasukawa, T. Fully Automatic Lexicon Expansion for Domain-Oriented Sentiment Analysis. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing* (*EMNLP'06*), 2006.
9. Kim, S. and Hovy, E. Determining the Sentiment of Opinions. *Proceedings of the 20th International Conference on Computational Linguistics* (*COLING'04*), 2004.
10. Liu, B., Hu, M. and Cheng, J. Opinion Observer: Analyzing and Comparing Opinions on the Web. *Proceedings of International World Wide Web Conference* (*WWW'05*), 2005.
11. Pang, B., Lee, L. and Vaithyanathan, S. Thumbs up? Sentiment Classification Using Machine Learning Techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* (*EMNLP'02*), 2002.

12. Popescu, A.-M. and Etzioni, O. Extracting Product Features and Opinions from Reviews. *Proceedings of the 2005 Conference on Empirical Methods in Natural Language Processing* (*EMNLP'05*), 2005.
13. Turney, P. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *ACL'02*, 2002.
14. Wiebe, J. and Riloff, E. Creating Subjective and Objective Sentence Classifiers from Unannotated Texts. *Proceedings of International Conference on Intelligent Text Processing and Computational Linguistics* (*CICLing'05*), 2005.
15. Wilson, T., Wiebe, J. and Hwa, R. Just How Mad Are You? Finding Strong and Weak Opinion Clauses. *Proceedings of National Conference on Artificial Intelligence* (*AAAI'04*), 2004.