Fundamentals of Sentiment Analysis and Its Applications

Chapter · March 2016

DOI: 10.1007/978-3-319-30319-2_1

2 authors:

Mohsen Farhadloo
University of California, Riverside

Erik Rolland
California State Polytechnic University, Pomona

Fundamentals of Sentiment Analysis and Its Applications

Mohsen Farhadloo and Erik Rolland

Abstract The problem of identifying people's opinions expressed in written language is a relatively new and very active field of research. Having access to huge amount of data due to the ubiquity of Internet, has enabled researchers in different fields—such as natural language processing, machine learning and data mining, text mining, management and marketing and even psychology—to conduct research in order to discover people's opinions and sentiments from the publicly available data sources. Sentiment analysis and opinion mining are typically done at various level of abstraction: document, sentence and aspect. Recently researchers are also investigating concept-level sentiment analysis, which is a form of aspect-level sentiment analysis in which aspects can be multi terms. Also recently research has started addressing sentiment analysis and opinion mining by using, modifying and extending topic modeling techniques. Topic models are probabilistic techniques for discovering the main themes existing in a collection of unstructured documents. In this book chapter we aim at addressing recent approaches to sentiment analysis, and explain this in the context of wider use. We start the chapter with a brief contextual introduction to the problem of sentiment analysis and opinion mining and extend our introduction with some of its applications in different domains. The main challenges in sentiment analysis and opinion mining are discussed, and different existing approaches to address these challenges are explained. Recent directions with respect to applying sentiment analysis and opinion mining are discussed. We will review these studies towards the end of this chapter, and conclude the chapter with new opportunities for research.

Keywords Sentiment Analysis · Opinion mining · Customer satisfaction · Probabilistic approaches

---

M. Farhadloo (✉) · E. Rolland
School of Engineering University of California, Merced, USA
e-mail: mfarhadloo@ucmerced.edu

E. Rolland
e-mail: erolland@ucmerced.edu

W. Pedrycz and S.-M. Chen (eds.), *Sentiment Analysis and Ontology Engineering*,
Studies in Computational Intelligence 639, DOI 10.1007/978-3-319-30319-2_1

1 Introduction

With the increasing popularity of analytics and data science, computational intelligence methods are proving to be important competitive tools in many industries. For example, in business analytics data is mined for patterns that would help better understand customers, and improve sales and marketing. Computational intelligence methods allow for probabilistic methods to be used in finding patterns in data. These methods typically work on low-level data, and are not guided by absolute knowledge as is the case with general AI methods [3]. In addition, a huge amount of the data that warrants analysis is now generated in written form. For example, users could leave written comments about a product or service in online sites like Yelp and TripAdviser. The written text is subject to interpretation, and representing the data in an absolute syntax (such as a binary system) is difficult. However, computational intelligence methods allows for such fuzziness, and may be the most appropriate methods for finding patterns in such data.

Sentiment analysis brings together various research areas such as natural language processing, data mining and text mining, and is fast becoming of major importance to organizations as they strive to integrate computational intelligence methods into their operations, and attempts to shed more light on, and improve, their products and services [32]. In sentiment analysis, or opinion mining, (SAOM), the goal is to discover people's opinions expressed in written language (text). Sentiment in term means "what one feels about something", "personal experience, one's own feeling", "an attitude toward something" or "an opinion".

Opinions are central to almost all human activities and are key influencers of our behaviors. Our beliefs and perceptions of reality, and the choices we make, are, to a considerable degree, conditioned upon how others see and evaluate the world. For this reason, when we need to make a decision we often seek out the opinions of others. This is not only true for individuals but also true for organizations [38]. Traditionally, closed-form customer satisfaction questionnaires would be used to determine the significant components, or aspects, of overall customer satisfaction [57]. However, the development and execution of questionnaires are expensive or may not be available. In some cases, public agencies are even prohibited by law from collecting satisfaction questionnaires from customers. In cases such as this, the only alternative may be to analyze publicly available free-form text comments, for example from sources such as Trip Advisor.

Organizations are increasingly focused on understanding how their value-creating activities are perceived by their customers. Customers' opinions drive the organization's image and the demand for their products or services. For non-profits and government organizations, better serving the customers' needs help generate political and taxpayer support, while in for-profit organizations improved customer understanding drives the organizational ability to generate revenues and compete in the marketplace. Thus, understanding customers' perception is a key to organizational success.

Customers' perceptions of products or services are often established through surveys, focus groups, observation, and other fairly labor intensive and expensive methods. With the advent of the Internet, survey tools have become more readily available (and cheaper to use), but obtaining accurate and relevant data from customers surveys is still a challenge. At the same time, the Internet has bred a whole industry of product/service review services, such as Yelp, that provides newfound capabilities and an ocean of information regarding customer's opinions. Examples include sites such as Trip Advisor for travel, Yelp and Urban Spoon for restaurants, Patagonia for outdoor clothing and gear, Lands' End for clothing and Epinions for product reviews have become commonplace. These sites allow customers to both read and provide reviews of the product or service. The customer reviews, which are often openly available on the web, contain a wealth of information usable by management, competitors, investors, and other stakeholders to discern customer concerns driving overall customer satisfaction for a particular service or product (for simplicity, we use the terms overall satisfaction and overall rating interchangeably). Usually the number of online reviews about an object is of a very large scale, such as hundreds of thousands, and the number is consistently growing as more and more people keep contributing online. Thus, we are now faced with a relatively new challenge of extracting customer opinions and sentiments from (often) unstructured data in the form of text comments.

An ideal method is one which would enable us to automatically identify the dimensions that drive customer satisfaction without *a priori* knowledge about those dimensions and their impact on overall customer satisfaction. The benefits of such a system would be that it would be able to discover the components and their impact on overall satisfaction without human intervention. Such a system enables analyzing large amounts of data which otherwise could not be analyzed by humans.

Both individuals and organizations can take advantage of sentiment analysis and opinion mining. When an individual wants to buy a product or decides whether or not use a service, he/she has access to a large number of user reviews, but reading and analyzing all of them could be a lengthy and perhaps frustrating process. Also when an organization or a manager seeks to elicit the public opinion about their products, market theirs products, identify new opportunities, predict sales trends, or manage its reputation, it needs to deal with an overwhelming number of available customers comments. With sentiment analysis techniques, we can automatically analyze a large amount of available data, and extract opinions that may help both customers and organization to achieve their goals. This is one of the reasons why sentiment analysis has been spread in popularity from computer science to management and social sciences. Sentiment analysis has also applications in review-oriented search engines, review summarization, and for fixing the errors in users ratings (such as for cases where users have clearly accidentally selected an incorrect rating when their review otherwise indicates a different evaluation) [43]. In public policy there are typically huge amounts of data (newspapers or other) documents containing opinions regarding an issue. For instance, during the recent financial crisis is Greece, there were vast amounts of documented opinions for or against Greece taking a new Eurozone loan.

One might establish a pretty good predictor for the Greek referendum by analyzing Greek media over the months leading up to the crisis. If sentiment analysis can be done, search engine could search for opinions and find the answers of these type of questions, as typically attempted with social media analytics.

Sentiment analysis can be used as a complement to other systems such as recommendation systems, information extraction and question answering systems [43]. The performance of recommendation systems may improve by not recommending items that receive a lot of negative feedback [51]. Information extraction systems may benefit from opinion mining systems by discarding certain types of information found in subjective sentences. In question answering systems different types of questions (opinion-oriented and definitional questions) may acquire different types of treatments. This is another field that can leverage from opinion mining.

This chapter provides an introduction to the problem of sentiment analysis. The focus of this chapter is on highlighting the various challenges researchers encounter in aspect-level sentiment analysis and some of the recent methods to address them. Since recently some researchers have started addressing sentiment analysis and opinion mining by using, modifying and extending topic modeling techniques, in this chapter a comprehensive survey about these methods will be given. Topic models are probabilistic techniques for discovering the main themes existing in a collection of unstructured documents. There are previous books and surveys on sentiment analysis [16, 37]. In this book chapter we aim at addressing recent approaches to sentiment analysis, and explain these in the context of wider use.

The remainder of this chapter is structured as follows: in Sect. 2 we first define the problem of sentiment analysis and then in Sect. 3 we will describe various challenges and approaches to them explained in the literature.

2 Problem Definition

In this section we will define the problem of SAOM and in particular the problem of aspect-level sentiment analysis. It is useful to first define some of the terms and concepts that are related to the problem. Here we have borrowed some of the definitions from the [37].

Contributor: A Contributor is the person or organization who is expressing his/her/its opinions in written language or text. Contributor may also called opinion holder or opinion source.

Object: An object is an entity which can be a product, service, person, event, organization, or topic [37]. It may be associated with a set of components and attributes that in sentiment analysis literature these components and attributes are called aspects.

Review: A Review is a contributor-generated text that contains the opinions of the contributor about some aspects of the object. A review may also be called opinionated document.

Overall rating: This is a user reported overall satisfaction with the object for example on a Likert scale from 1 to 5.

Opinion: An opinion on an aspect is a positive, neutral or negative view, attitude, emotion or appraisal on that aspect from a contributor.

Aspect: An aspect is an important attribute of the object with respect to overall costumer satisfaction that the contributor has commented on in their review.

An opinionated review may be generated by a number of contributors or contains opinions from a number of sources. Also it should be noticed that the review may be a direct review about a single object or a comparative review in that compares 2 or more objects to each other. In general a review d_i is a collection of sentences $d_i = \{s_{i1}, s_{i2}, \ldots, s_{im}\}$ that hold opinions from different contributers about different aspects of different objects. Liu [37] has defined a direct opinion as a quintuple of $(o_j, f_{jk}, oo_{ijkl}, h_i, t_l)$ where oo_{ijkl} is the opinion orientation of contributer h_i about the k^{th} aspect f_{jk} of object o_j. Although this model is not a comprehensive model that contains all information about all cases, it is a sufficient model for practical applications. It is worth mentioning that most of the literature have considered the orientation of the opinion as a binary variable: either positive or negative. There are some works that include the neutral class as well. In general the opinion orientation could be considered in different scales rather than just binary or ternary.

The objective of SAOM can be written as:

> Given a collection of reviews $D = \{d_1, d_2, \ldots, d_D\}$ all about an object, discover all **aspects** and corresponding **sentiments** expressed in that collection.

This objective can be reached at different levels. If the focus is on each document and the sentiment orientation of the whole document is found, this is called **document-level** sentiment analysis. Although a document may convey an overall positive or negative sentiment, it is quite likely that not all sentences in a document are all positive or negative. For example if the reviews of interest are movie reviews, a particular contributor may give an overall positive rating to the movie, however they may not like all aspects of that movie and in their review they mention what aspects they like and what they do not like about that movie. When the focus of SAOM is on sentences and a sentiment orientation for sentences of the review is found, it is called **sentence-level** sentiment analysis. In a finer analysis, one may be interested in knowing what particular aspects of the object the contributer is commenting on and also whether they like them or not. This level os SAOM is called **aspect-level** SAOM.

The visualization and reporting of the SAOM results are essential issues that need particular attention. There are different ways to visualize the results of the analysis. One way could be making a list of discovered aspects and linking to each of them the corresponding positive and negative reviews. This could be augmented with a rating for each aspect. This way people can easily find out their peers' opinions about different aspects and specifically pick what particular reviews they want to read. It is likely that some contributors give a positive/negative score to an object just because they are (not) happy with a particular aspect of the object. Other users may/not care about those particular aspects as much as others. Having an aspect-centric review report makes it possible to be more focused on the specific aspects that you care and seek others' opinion regarding them. If one wants to get an idea about popular or unpopular aspects of an object, creating a frequency list of different discovered aspects could be beneficial. One of the benefits of SAOM is that it empowers individuals and organizations to track the people's opinions over time. How things change over time has critical values for individuals and organizations. For example, Blackberry was a market leader in the mobile phone market for years, but lost out to the market pressure due to lacking features in their mobile phones. Once the importance of the missing aspects of a Blackberry phone, such as the lack of ability to text, outweighed the importance of secure email, their market-share diminished severely.

3 Challenges

In order to illustrate the need for fuzziness in the computational intelligence approaches to extracting patterns from user comments, we summarize below some major challenges associated with SAOM.

3.1 Synonymy and Polysemy

Users in different contexts or with different needs, knowledge or linguistic habits will describe the same information using different terms (synonymy) [11, 13]. Furnas et al. [20] showed that people generate the same keyword to describe well-known objects only 20 percent of the time. Because searchers and authors often use different words, relevant materials are missed. People also use the same word to refer to different things (polysemy) [13]. Words like saturn, jaguar, or chip have several different meanings. In different contexts or when used by different people the same term takes on varying referential significance [11]. In general synonymy and polysemy are caused due to the possible variability in word usage. Synonymy and polysemy issues pose difficult challenges on sentiment analysis and opinion mining problem.

Several approaches have been proposed in order for addressing the synonymy and polysemy challenges. *Stemming* which can be seen as normalizing some kind of surface-level variability is a popular technique. The purpose of stemming is to bring variant forms of a word together (to their morphological roots). A stemming algorithm for English reduces the words “stemmer”, “stemming”, “stemming” and “stems” to the root word “stem”. Stemming algorithms have been studied in computer science since 1960s [39, 46, 59]. Stemming may sometimes help information retrieval but, it does not address cases where related words are not morphologically related (e.g., physician and doctor) [13].

Controlled vocabulary is another approach that has been shown effective in dealing with the issues caused by variability in word usage [2, 35, 50]. However since in controlled vocabulary approach it is a requirement that terms be restricted to a pre-determined list of words, it is not applicable in SAOM (it is not feasible to restrict contributors to limit their vocabulary to a pre-determined one). In review websites the contributors can usually express their opinions in any way that they prefer and it is not feasible to enforce a particular set of words.

Latent Semantic Analysis (LSA) [11] is an other approach to work around synonymy challenge. With advent of large scale collections of text data, statistical techniques are being used more and more to discover the relationship among terms and documents. LSA simultaneously models the relationship among documents based on their constituent words and relationship among words based on their occurrence in the documents. LSA can be seen as a linear dimensionality reduction technique based on singular value decomposition of *term-document* matrix. By reducing the dimensions and using fewer dimensions than the number of unique words, LSA induces similarities among words. LSA constructs a *semantic space* wherein terms and documents that are closely associated are place near one another [11]. It is worth noting that while LSA method deals nicely with the synonymy problem, it offers only a partial solution to polysemy problem. Since the meaning of a word can be conditioned by other words of the in the document, LSA provides some help regarding polysemy problem. However the failure arises from every single word having only a single representative point in the semantic space of LSA not more than one representatives.

3.2 Sarcasm

Identifying sarcasm is a very hard task for humans and it is even harder for machines. The ability to reliably identify sarcasm in text can improve the performance of many natural language processing tasks particularly SAOM. Sarcasm is a form of expression where the literal meaning is opposite to the intended. “The restaurant was great in that it will make all future meals seem more delicious,” is an example of a sarcastic sentence, in which, although there is technically no negative term in the language, it is intended to convey a negative sentiment. This example clearly shows some difficulties in dealing with sarcastic phrases. Dealing with a sarcastic situation requires a good understanding of the context, the culture of the situation, the topic, the people and also the language involved in the sarcastic statement. Having access to all of these pieces of information is a difficult task in itself, but trying to make use of them is especially challenging for a machine. Although the phenomenon of sarcasm has extensively been studied in fields such as psychology, cognitive science and linguistics [22, 23, 53], very few attempts have been done on analyzing it computationally. Lack of a dataset with reliably labeled instances of sarcasm and not-sarcasm is one of the reason that computational analysis of sarcasm is very young. Recently [18] has generated a corpus consisting of Amazon product reviews that can be used for understanding sarcasm on two levels: document and text utterance level. The corpus contains a pair of sarcastic and non-sarcastic reviews written about the same object. Filatova [18] has used quality control algorithms to quantify the quality of their collection. There are some works [18, 24, 41] that have tried to automatically detect sarcasm in small text phrases such as Twitter data. However the drawback of working on short phrases in order to detect sarcasm is the ignorance of the fact that context plays an important role in sarcastic situations. Indeed [24] has shown that lexical features are not sufficient for sarcasm detection and pragmatic and contextual information are necessary in order to enhance the performance. Analyzing sarcasm is an area that is in need of more thorough research in SAOM research community.

3.3 Compound Sentences

A compound sentence has two independent clauses or sentences. Two independent clauses can be joined by a coordinating conjunction (such as "and", "or", "but" and "for") or a semicolon. Dealing with compound sentences makes the problem of SAOM difficult. For instance a sentences like "The kids enjoyed the beach but we did not," or "Despite a pleasant experience I cant support the many reviews that it was a great restaurant," are challenging for sentiment analysis. Dealing with compound sentences is still largely an open area of research in SAOM.

3.4 Unstructured Data

The reviews that the contributors have written for each object are the input which are in the plain text format. A challenge in the problem is the transformation of the unstructured input data which is available in the form of written reviews into a semi-structured data. Semi-structured data is data that is neither raw data, nor conformal with the formal structure of data models associated with relational databases or other forms of data tables, but nonetheless contains tags or other markers to separate semantic elements.

3.5 Aspect Identification

A body of works in the literature have addressed the problem of aspect-level SAOM in two stages: first, aspect identification and second, sentiment identification. The goal of the aspect identification is to discover the particular aspects of the object that contributors are expressing their opinions about. There are 2 categories of works in the literature addressing aspect identification:

1. **Automatic extraction**: There is no prior knowledge about the aspects an the aspects are automatically extracted from the given reviews [4, 15, 21, 29, 40]. This category can be divided into *supervised* and *unsupervised* subcategories.
2. **(Semi) Manual extraction**: Either a subset of aspects or the whole set of desired aspects are known a priori. In cases where a subset of aspects are known, the subset is being used as a seed set and expanded in some ways [55].

Of the early works in the aspect level sentiment analysis, is the research by Hu and Liu [29]. Hu and Liu propose to first find the frequent aspects using association mining, and then using them, they extract the infrequent aspects. Any sentence that contains one of the frequent aspects is being analyzed to find out the infrequent aspects.

In [21] authors aim at presenting an unsupervised aspect identification algorithm that employs clustering over sentences with each cluster representing an aspect. They finally proposed applying an empirical weighting scheme to the list of terms, which are sorted according to their frequency of occurrence.

Clustering over sentences has been used in [15] in order to find similar sentences which are most likely about similar aspects. In [15] instead of representing the sentences using the commonly Bag-Of-Word (BOW) method, they propose to use Bag-Of-Nouns that makes clustering more effective.

Blair-Goldensohn et al. [4] propose a sentiment summarization system for local services. In their system the aspects are divided into two types: dynamic aspects (string-based frequent nouns/noun phrases) which are extracted similar to the method of [29] and static aspects (generic and coarse-grained ones), which are extracted by designing classifier for each one using hand-labeled sentences.

PLSA [28] is the probabilistic version of LSA that has evolved from it. PLSA is a model for a collection of documents in which each document is modeled as a mixture of topics. Lu et al. [40] have taken advantage of PLSA and the structure of the sentences for aspect identification. Each document is represented as a bag of phrases and each phrase is a pair of head and modifier terms $<h_i, m_i>$. Each aspect is also modeled as a distribution over *head terms*. In the unstructured version a document is considered as a collection of head terms (the modifiers are ignored). In this model the log likelihood of all documents can be written as:

$$logp(D|\Lambda) = \sum_{d=1}^{D} \sum_{w_h \in V_h} \{c(w_h, d) log \sum_{k=1}^{K} [\pi_{d,k} p(w_h|\Theta_k)]\} \tag{1}$$

where V_h is the set of all head terms in the vocabulary, $\pi_{d,k}$ is the proportion of topic k in document d, $c(w_h, d)$ is the number of times head term w_h has been occurred in document d, Θ_k is the k^{th} aspect-topic and Λ is the set of all model parameters. In the structured version (*structured* PLSA) since a single modifier term can modify different head terms, each modifier term is being modeled as a mixture of K topics. Each modifier term w_m is represented as a set of head terms that it modifies and again each aspect is a distribution over head terms (Θ_k). The modifier can be regarded as a sample of the following mixture model:

$$p_{d(w_m)}(w_h) = \sum_{k=1}^{K}[\pi_{d(w_m),k}\, p(w_h)|\Theta_k)] \tag{2}$$

and the log likelihood of the collection of modifiers V_m is:

$$logp(V_m|\Lambda) = \sum_{w_m \in V_m} \sum_{w_h \in V_h} \{c(w_h, d(w_m))log \sum_{k=1}^{K}[\pi_{d(w_m),k}\, p(w_h|\Theta_k)]\} \tag{3}$$

In both equations of (1) and (3) Λ is the model parameters set (including Θ_k) and is estimated using Expectation-Maximization (EM) algorithm. The details of the EM algorithms can be found in [40]. It is worth noting that the structured PLSA estimated the parameters based on the co-occurrence of head terms at the level of modifiers, not at the level of documents. Since the reviews are usually short reviews and have few phrases, the structured PLSA is typically more informative.

3.6 Sentiment Identification

The sentiment identification is the process of discovering the opinion orientation of the text fragments of interest. The sentiment orientation can be expressed in different scales. Rating scales are used widely online in an attempt to provide indications of consumer opinions of products. Many sites like *TripAdvisor.com, Amazon.com, Epinions.com* use the 5-star overall rating scale. Users can cast their vote about movies in a $1-10$ rating scale on *IMDb* and an on a 5-star scale on *rottentomatos.com*.

In SAOM dealing with negative sentiments is particularly difficult. As Leo Tolstoy said in his book *Anna Karenina* “All happy families resemble one another, each unhappy family is unhappy in its own way”, in SAOM all happy sentiments are pretty much alike, each unhappy sentiment is uniquely expressed. In some cases, people choose to cover their criticisms with qualifiers to be polite (even on the Internet), and more broadly, people just tend to be more creative in how they choose to describe the things that they do not like. All of these make the automatic solution of the problem challenging. Automatic techniques make use of machine learning algorithms to address the problem and these algorithms aim at finding patterns from data and try to learn from training examples. When there are not enough training examples or there are training examples with very limited variety, one can imagine how hard the pattern extraction would be.

Most of the work in the literature have treated sentiment in a binary scale of negative and positive [29, 44], some have added a third class of sentiment (the neutral sentiment) into their consideration [15]. As in some actual cases that each review comes with an overall rating in a 5-star scale, the number of sentiment orientations can vary. [33] have considered 5-level sentiment categories in their experiments.

There are 2 types of fragments in a textual document: *objective fragments* are those text fragments that express factual data and *subjective fragments* that express personal feelings. Although both subjective and objective fragments may carry an opinion about an object, it is more likely for a subjective sentence to be opinionated. The process of sentiment identification can be written as the following steps:

1. Extract all the opinionated fragments.
2. Identify the sentiment of each opinionated fragments.
3. Infer the overall sentiment from the opinion of individual fragments.

3.6.1 Extract Opinionated Fragments

Subjectivity classification is the task of determining whether a fragment of text is subjective or objective. If a text fragment is being categorized as subjective then it is very likely that it is an opinionated piece of text and then its opinion orientation needs to be identified. There are works on subjectivity determination [19, 27], that can be used in this regard.

It is sensible to conjecture that opinionated words and phrases are dominant sentiment indicators. Since adjective, adverbs and also some nouns and verbs are strong sentiment indicators, Part-Of-Speech (POS) tags and syntactic structure of the sentences are also beneficial in extracting opinionated fragments. POS tagging is the process of marking up a word in a text (corpus) as corresponding to a particular grammatical category such as nouns, adjective, adverb, verb or etc. categories. The POS tagging algorithms are based on terms definitions as well as the context i.e. the term's relationship with adjacent and related words and phrases.

In aspect-level SAOM the sentiment orientation of all text fragments that contains one of the extracted aspects has to be identified. So if in the first stage the aspects have been identified they can be used in order for extracting opinionated fragments.

3.6.2 Identification of Sentiment Using Unsupervised Techniques

Once the opinionated fragment has been extracted in the next step the sentiment orientation of it should be identified. Turney [52] has determined the sentiment polarity of a word by computing its similarity to 2 seed terms: "Excellent" for positive and "Poor" for negative polarity. The similarity between an unknown word and each of the seed terms is computed by counting the number of occurrences and co-occurrences of them using the results of a web search engine. Depending upon which one of the seed terms the unknown term is more similar to, the sentiment orientation is determined.

3.6.3 Identification of Sentiment Using Supervised Techniques

Sentiment identification can easily been formulated as a classification problem. In the literature of SAOM various classification techniques have been employed from machine learning to identify sentiments. Pang et al. [44] have experimented a variety of features with both naive Bayes classifiers and also Support Vector Machines (SVM) to classify the whole input document (movie reviews) as either positive or negative. The best results came from the unigrams in a presence-base frequency model run through SVMs. For sentiment classification of sentences [21] implement a naive Bayes classifier using expectation maximization and bootstrapping from a small set of labeled data to a large set of unlabeled data. Feature engineering, finding the most suitable set of features for sentiment identification, is a critical issue in designing classifiers. Different features have been used in SAOM such as *terms and their frequencies, POS tags, opinion words and phrases, syntactic structures and negations*. In [15] a new feature set called *score representation* was introduced for classification of sentiments. Score representation is a low dimensional representation and is computed using the scores of the terms in the vocabulary that are learned from the data. Score representation has been used for designing 3-class SVM classifiers and outperforms the classifiers that are based on BOW representation.

Another feature engineering approach based on a hierarchical deep learning framework that simultaneously discovers aspects and their corresponding sentiments has been proposed in [34]. In their framework classifiers based on deep learning (recursive neural networks) have been designed that jointly predict the aspect and sentiment labels of the input phrase. The main idea is to learn a vector (or a matrix or a tensor) representation for words using hierarchical deep learning that can explain the aspect-sentiment labels at the phrase level. In this framework each node in the parse tree of the given phrase is represented by a vector (or a vector and a matrix) and the model is defined by leveraging the syntactic structure of the phrase encoded in the parse tree. Figure 1 shows an example parse tree in which each node is associated by a d-dimensional vector.

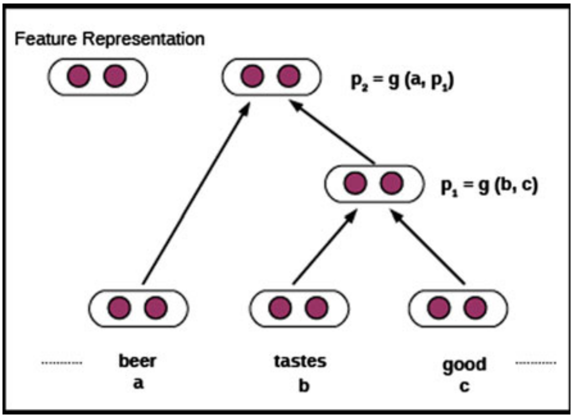


Fig. 1 An example parse tree. Each node is associated with e vector and the vector representation of the nodes are computed *bottom-up*

The vector representation of the nodes are calculated bottom-up as:

$$p_1 = f\left(W\begin{bmatrix} b \\ c \end{bmatrix}\right), \qquad p_2 = f\left(W\begin{bmatrix} a \\ p_1 \end{bmatrix}\right) \tag{4}$$

where p_1 and p_2 are the parent vectors and b and c are the leaf nodes and $f = tanh$ is a standard element-wise non-linear function. The goal is to learn the combination matrix $W \in R^{d\times 2d}$ and also the compositional feature representation of the words (the associated vectors and matrices). In this model the class label of each given phrase is predicted using the vector representation of its root node of the parse tree as:

$$y_i = softmax(W_s p_i^{\text{root}})$$

where $W_s \in R^{C\times d}$ is the classification matrix and needs to be estimated. The true label of the given phrase in the training set is $t_i \in R^{C\times 1}$ that has an entry 1 at the correct label and a 0 at other indices. Since the aspects and sentiments meant to be captured jointly the labels are supposed to be provided as aspect-sentiment pairs. For instance (Taste, Positive) correspond to one class label. If the whole set of model's parameters is denoted by Θ then Θ is estimated in a way that the softmax function of the vector representation at the root level of the parse tree $y_i \in R^{C\times 1}$ matches the class label of the i^{th} text snippet as closely as possible. This can be achieved by minimizing the cross entropy error between y_i and t_i. The error function to be minimized is:

$$E(\Theta) = \sum_i \sum_j t_{i,j} log y_{i,j} + \lambda \|\Theta\|^2 \tag{5}$$

The non-convex objective function in Eq. (5) is minimized using Adagrad optimization [12] procedures. The estimation involves computation of the sub-gradients of the $E(\Theta)$ and the forward calculation of the vectors and matrices and back-propagation of the softmax error at the root node.

3.6.4 Identification of Sentiment Using Lexicons

Many work in the literature have identified the sentiment orientation of the fragment of interest using some opinion lexicons [1, 14, 26, 29, 49, 58]. Lexicon based approaches discover the sentiment orientation of an opinionated phrase of interest by looking-up it from their existing lexicons. The lexicon generation process usually start with a seed list of opinion words that their opinion orientations are known a priori and then using different approaches the seed list is expanded and more opinion words are added to it. There exist 2 main strategies to expand the initial seed list of opinion words: *dictionary-based* and *corpus-based* strategies. Dictionary-based methods make use of online dictionaries like *WordNet* [8, 14, 17] and some additional information in them (like their glosses) in order to expand the initial seed set. These approaches take advantage of synonym and antonym relationships among words and also some machine learning techniques to generate better list. The generated lexicon based on online dictionaries is a context independent opinion lexicon. In order to generate domain specific lexicon that can capture the opinion words in a particular domain, corpus-based strategies have been proposed. In this strategy the initial seed list of opinion words is expanded using some syntactic rules and co-occurrence patterns. In this approach rules are designed for connective terms such as *And, Or, But, Neither-or, Neither-nor* and leveraging them the seed set is expanded. Using the particular corpus of interest domain specific lexicons can be generated.

For sentiment identification, [29] follow a lexicon based approach, which involves identifying the closest adjectives to the nouns in the under process sentence and look-up those adjectives or their synonyms in their lists of positive and negative adjectives. Blair-Goldensohn et al. [4] first compute a single sentiment score for each term in their lexicon. These scores are computed starting with an initial seed with arbitrary scores, and then propagation of these scores using a modified version of standard label propagation algorithms over a graph [62]. A shortcoming of this approach is that it is an iterative algorithm and it is not clear when to stop propagating the scores. Using the computed scores and considering the neighbors of each sentence, they design maximum entropy classifiers for positive and negative classes. Concept-based approaches use Web ontologies or semantic networks to accomplish semantic text analysis [9]. In this approach the document is represented with bag-of-concepts instead of bag-of-words. In concept-based approach for identifying the sentiment of each document a dictionary (SenticNet) which contains the affective labels of concepts is used [45]. However it is not clear how to identify the aspects using the bag-of-concepts representation.

3.7 Topic Model Based Approaches

Fundamental characteristics of human word usage is that people use a wide variety of words to describe the same object or concept [13]. As it was mentioned in Sect. 3.1 dealing with synonymy and polysemy in natural language is a challenging issue in many text mining tasks. This illustrates the need for the probabilistic features of computational intelligence to address SAOM. In topic modeling, each topic is defined as a probability distribution over terms (terms of considered vocabulary list). One of the outputs of topic modeling algorithms is a probability distribution over terms for each topic that determines which words have higher probability in the context of a particular topic. It is reasonable to expect that in each context there are certain terms that have higher probability of occurrence than other terms. Topic modeling techniques by definition are able to capture this characteristic of natural language.

Topic models are probabilistic techniques based on hierarchical Bayesian networks for discovering the main themes existing in a collection of unstructured documents. Most of the existing works in the literature solve the problem of SAOM in 2 stages: first aspect identification and second sentiment identification [4, 15, 29].

One of the advantages of methods that are based upon topic modeling is that they are able to find aspects and sentiments simultaneously. Also these algorithms do not require labeled training data and they find the topics from the analysis of the original texts.

Latent Dirichlet Allocation (LDA) was proposed in order to find short descriptions of the members of a collection. LDA enables efficient processing of large collections while preserving the essential statistical relationships that are useful for basic tasks such as classification, novelty detection, summarization and similarity and relevance judgments [6]. LDA assumes that each document has been generated from a mixture of topics. Therefore for each document to be generated, the proportion of each topic in that document should be known and also for each word of a document to which topic that word belongs, should be known. The generative process of LDA can be summarized as [5]:

(1) **For each document**:

   (a) Randomly choose a distribution over topics.

(2) **For each word in the document**:

   (a) Randomly choose a topic from the distribution over topics in step 1.
   (b) Randomly choose a word from the corresponding distribution over the vocabulary (topic).

It is worth mentioning that LDA is based on the assumption that there exist a hidden structure that has generated the given collection of documents. Given the collection of documents (observation) the goal is to find the topics ($\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_K\}$), distribution over topics for each document ($\boldsymbol{\theta}_d$) and topic assignment ($z_{d,n}$) for each word $w_{d,n}$ in each document d. $\{\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{Z}\}$ construct the hidden structure behind the observed data and that is why the model is called *Latent* Dirichlet Allocation.

A neat way to describe the topic models is by using graphical models that provide a graphical language for describing families of probability distributions. The graphical model of LDA is depicted in Fig. 2.

In every graphical model a number of statistical dependencies are encoded that define that particular model. For instance in Fig. 2, the probability of n^{th} word in document d ($w_{d,n}$) depends on the topic assignment $z_{d,n}$ of that word and all topics $\{\boldsymbol{\beta}_{1:K}\}$. Topic assignment determines which topic to be used to get the probability of $w_{d,n}$. Also the topic assignments of the words in document d depend on distribution of topics $\boldsymbol{\theta}_d$ in that document. The joint distribution of the hidden and observed variable of the graphical model in Fig. 2 can be written as:

$$p(\boldsymbol{\beta}_{1:K}, \boldsymbol{\theta}_{1:D}, \mathbf{Z}_{1:D}, \mathbf{W}_{1:D}) = \prod_{k=1}^{K} p(\boldsymbol{\beta}_k) \prod_{d=1}^{D} p(\boldsymbol{\theta}_d) \prod_{n=1}^{N} p(z_{d,n}|\boldsymbol{\theta}_d) p(w_{d,n}|z_{d,n}, \boldsymbol{\beta}_{1:K})) \tag{6}$$

Given a collection of documents $D = \{d_1, d_2, \ldots, d_D\}$ the goal of training is to learn the hidden structure $\{\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{Z}\}$. In Bayesian framework, the particular computational problem in order to use the model is the *inference problem* in which computing

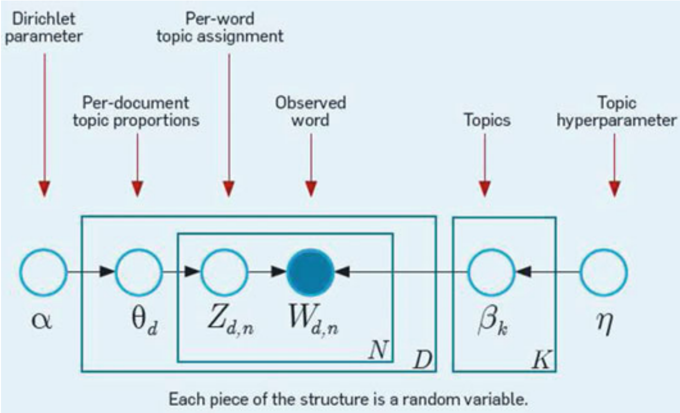


Fig. 2 Latent Dirichlet Allocation graphical model. The *shaded* node is an observed random variable and the *unshaded* ones are hidden random variables

the posterior distribution of the hidden variables given the observed variables is of interest. The exact computation of the posterior probability is intractable and approximation methods have been considered.

There are 2 general approximation methods in topic modeling algorithms:

(1) **Sampling-based algorithms**: These algorithms are Markov Chain Monte Carlo (MCMC) techniques which provide a principled way to approximate an integral (expected value). Monte Carlo techniques are algorithms that aim at obtaining a desired value by performing simulations involving probabilistic choices [47]. Sampling-based algorithms like *Gibbs sampling* try to approximate the posterior distribution by drawing samples from it without explicitly computing the posterior distribution. In Gibbs sampling the hidden structure defines a *vector* that represents the *state* of the system. The basic idea in Gibbs sampling is instead of sampling a multivariate random vector all at once, each dimension is being sampled using the samples of the other dimensions. For an introduction to Gibbs sampling you can refer to [47] and see [48] for a good description of Gibbs sampling for LDA.

(2) **Variational algorithms**: These algorithms approximate the posterior probability by a parameterized family of distribution over hidden variables and using optimization techniques find the set of parameters that makes the approximated distribution closest to the exact posterior. We can put variational methods in other words: these algorithms are based on Jensen's inequality and try to obtain an adjustable lower bound on the log likelihood. Essentially one considers a family of lower bounds, indexed by a set of variational parameters. Th variational parameters are chose by an optimization procedure that attempts to find the tightest possible lower bound [6].

The LDA model is the simplest topic model that provides a powerful tool for discovering the hidden structure in a large collection of documents. Since its introduction LDA has been extended and modified in many ways. One of the areas that topic models can very well fit is the area of SAOM. In particular in aspect-level sentiment analysis, people are talking about different aspects of an object. Different contributors may use different terms in order to point to the same aspect of an object. For example in reviews about a camera, you may see people commenting about *image quality* and people commenting about *resolution* of the camera and most likely they are speaking about the same aspect. Therefore when people are talking about a particular aspect, a number of terms are most likely to be exploited. In topic modeling terminology, this means that each aspect has a distribution over terms or should be considered as a topic in our model. Also when contributors want to give a particular rating to an aspect in their review, a specific range of words have higher probability and for another level of rating a different set of words have higher probability mass than the rest of the vocabulary. This means that each rating (sentiment) like each aspect should be considered as a topic in order to extend topic modeling algorithms for SAOM.

Just simply increasing the number of topics (to count for both aspect-topics and rating-topics) in LDA is not sufficient to apply it to SAOM. Researchers have started modifying LDA in many ways in order to adapt it to address the problem of aspect-level sentiment analysis.

One of the assumptions of LDA is the Bag-of-Word (BOW) assumption. In LDA it is assumed that the order of the terms in the document can be neglected and the document can be represented as a collection of words. Since this assumption is not realistic, researchers have tried to address this shortcoming. To go beyond BOW [54] has suggested bigram topic model. In bigram-topic model [54] has integrated bigram-based and topic-based approaches to document modeling. In this model each topic instead of being a single distribution over words, has multiple distributions over words depending on the previous word. If the size of the vocabulary list is N then each topic is characterized by N distributions specific to that topic. It can easily be realized that the parameter space in this approach will expand rapidly. The LDA collocation (LDA-Col) model [25] is another attempt to go beyond BOW assumption of LDA. In LDA-Col a new hidden variable is introduced for each word in each document that determines whether or not the word should be drawn from a unigram model or should be drawn from a bigram model considering its previous word. Topical N-grams (TNG) model [56] is very similar to LDA-Col and the only difference is that in TNG it is possible to decide whether to form a bigram for the same two consecutive word tokens depending on their nearby context. The TNG model automatically determines to form an n-gram (and further assign a topic) or not, based on its surrounding context. Examples of topics found bt TNG are more interpretable than its LDA counterpart [56].

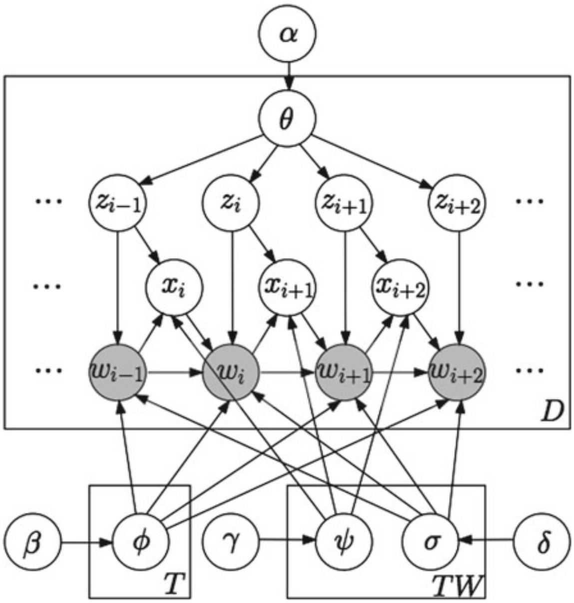


Fig. 3 Topical N-grams graphical model. The *shaded* node is an observed random variable and the *unshaded* ones are hidden random variables

In bigram topic model there is no extra hidden variable to decide whether or not to form a bigram or a unigram (all the terms are drawn from bigram models), in LDA-Col the binary random variable x_i for each word w_i determines unigram or bigram status however the value of x_i does not depends on the previous topic z_{i-1}. The graphical model of LDS-Col is exactly as Fig. 3 except that the links from z_i to x_{i+1} are not there.

In another attempt in the direction of relaxing BOW assumption of LDA in that it ignores the order of the words, [33] has introduced the notion of aspects and sentiments coherency and has also tried to leverage from the existing syntactic structure in natural sentences. In that model (CFACTS) they have considered each document as a collection of *windows* and all the words inside a window have the same aspect or sentiment topic (coherency). Therefore in spite of LDA that each word of a document has a topic assignment, each window has a topic or sentiment assignment variable and all the words in that window have the same assignment. In the CFACTS model each word could be either aspect word, sentiment word or background word, so they have considered a new hidden variable c_i for each word that determines to which category that word belongs. The syntactic dependencies among words have been captured through c_{i-1} and c_i in the model of [33].

Since the sentiments and aspects for adjacent windows may still be dependent for a specific review document in many ways, in the CFACTS model in Fig. 4, the multinomial variable $\psi_{d,x}$ for each window x of a review d has been introduced. $\psi_{d,x}$ can take on 3 different values that captures the following scenarios:

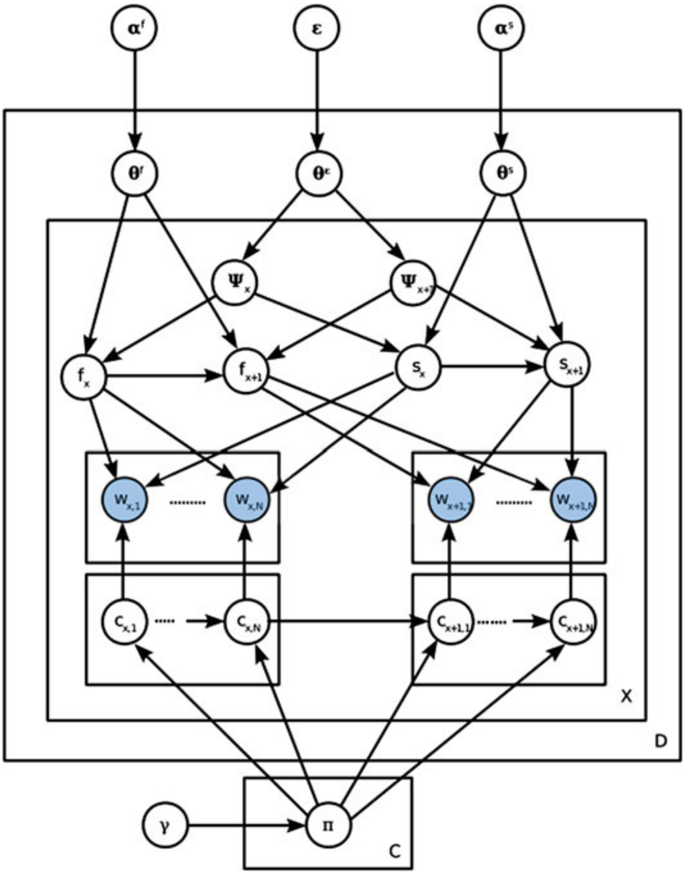


Fig. 4 Coherent Facet and Sentiments graphical model. The *shaded* node is an observed random variable and the *unshaded* ones are hidden random variables. This model introduces two new variables: $\psi_{d,x}$ which determines the dependency between windows and c_i that specifies to which category each word belongs: aspects, sentiments or background category

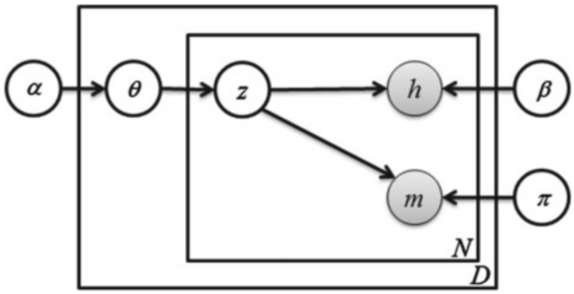
$\phi_{d,x} = 0$: indicates that both aspect and sentiment of current window is the same as those of the previous window.

$\phi_{d,x} = 1$: indicates that the aspect topic of the current window is independent from the previous window but the sentiment topic of the current window is the same as previous one.

$\phi_{d,x} = 2$: indicates that both aspect and sentiment topics of the current window are independent form the previous window.

On the same line as coherency idea the model of [7] assumes all words in a single sentence are generated from a single topic and applies LDA on each sentence to extract topics. The authors of [31] further extend the idea to discover sentiments related to each aspect. In this model, each review has a distribution over sentiments

Fig. 5 Phrase Latent Dirichlet Allocation graphical model. The *shaded* node is an observed random variable and the *unshaded* ones are hidden random variables



and each sentiment has a distribution over aspects. To generate each sentence of a review, first a sentiment is drawn from the distribution over sentiments of that review, and then an aspect is drawn from the distribution of that sentiment over aspects. Each word of the sentence is then generated based on the selected aspect and sentiment.

Another work around the BOW assumption of LDA is Phrase-LDA (PLDA). PLDA assumes a bag-of-phrase model of reviews (Fig. 5). An opinion phrase is a pair of $<h_n, m_n>$ which leads to two observed variables head term h_n and modifier term m_n. In [61] PLDA is applied to extract topics from reviews with the further assumption that each sentence of the review is related to only one topic. Compared to PLDA, separate PLDA (SPLDA) model introduces a separate rating variable which is conditionally independent from the aspect. In this model a review is generated by first drawing a distribution over aspect/rating pairs θ, and then repeatedly sampling N aspects and ratings as well as opinion phrases $<h_n, m_n>$ conditioned on the chosen value of θ.

In [42] a set of guidelines for designing LDA-based models are presented by comparing a series of increasingly sophisticated probabilistic graphical models based on LDA. In order to get familiar with some aspects of LDA-based methods and compare the impact of various design decisions such as adding latent and observed variables and dependencies we refer the readers to [42].

Latent Dirichlet Allocation is a member of a larger family of methods: latent variables models. Although there are so many works on aspect-level sentiment analysis based on LDA, there are other latent variable model approaches based on conditional random fields [10, 36] and Hidden Markov Models [30, 60]. Interested readers should look into them.

4 Conclusions

Computational intelligence methods play a central role in sentiment analysis, and have proven to be powerful tools in helping to understand customers' perceptions related to products and services. While there have been many advances in the short history of this field, there is still much work to be done. Most of the work thus far has been focused on deciphering the semantics of written text, and this research is impacted by various linguistic challenges. Nevertheless, we see that past research has been able to propose methods that uncover sentiments, opinions, and aspects and that correlate very well with customer satisfaction scores. What remains less clear is the generalizability of the methods across contexts or domains in order to really test to what extent the probabilistic computational intelligence methods are generalizable. Thus, the opportunities for continued research is large, and this research area will in turn bring about a sea-change in how organizations understand their customers, and possibly in how customers understand and evaluate products and services.

5 Questions for Further Research

While research will continue to improve on methodologies for sentiment discovery and aspect identification, there is also a need for bringing in other types of context specific information into the analysis. For example, one could imagine including more information about the reviewer, as such information often is readily available (many reviewing systems contain potentially useful reviewer information). Also, one could also incorporate location specific information (geotagging) to improve on understanding sentiments and opinion, and possibly quantitative or qualitative information related to non-textual media (such as pictures or videos) for further enhance the analytical capabilities of sentiment analysis. Further, we believe that the dynamic aspect of sentiment analysis will need to be better addressed. That is, sentiments, opinions, and aspects are dynamic in nature, and change with changing competition, technology, use, and so on. We stress that the computation efforts associated with sentiments analysis will have to be efficient enough to handle large data in a dynamic fashion. Finally, the we believe that sentiment analysis as discussed herein will be strongly dependent on better understanding information quality. That is, garbage-in may result in garbage-out for sentiment analysis. As such, Information quality as a research field, also holds many opportunities for future research, and those may be strongly coupled to sentiment analysis.

References

1. Andreevskaia, A. Bergler, S.: Mining wordnet for a fuzzy sentiment: Senti-ment tag extraction from wordnet glosses. In: EACL, vol. 6, pp. 209–216, (2006)
2. Bates, M.J.: Subject access in online catalogs: a design model. J. Am. Soci. Inf. Sci. **37**(6), 357–376 (1986)
3. Bezdek, J.C. Sankar, K.: Fuzzy models for pattern recognition. Technical report, USDOE Pittsburgh Energy Technology Center, PA (United States); Oregon State University, Corvallis, OR (United States). Department of Computer Science; Naval Research Lab., Washington, DC (United States); Electric Power Research Institute, Palo Alto, CA (United States); Bureau of Mines, Washington, DC (United States) (1994)

4. Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G.A., Reynar, J.: Building a sentiment summarizer for local service reviews. In: WWW Workshop on NLP in the Information Explosion Era (2008)
5. Blei, D.M.: Probabilistic topic models. Commun. ACM **55**(4), 77–84 (2012)
6. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
7. Brody, S., Elhadad, N.: An unsupervised aspect-sentiment model for online reviews. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, pp. 804–812 (2010)
8. Cambria, E., Havasi, C., Hussain, A.: Senticnet 2: a semantic and affective resource for opinion mining and sentiment analysis. In: FLAIRS Conference, pp. 202–207 (2012)
9. Cambria, E., Schuller, B., Xia, Y., Havasi, C.: New avenues in opinion mining and sentiment analysis. Intell. Syst. IEEE **28**(2) (2013)
10. Choi, Y., Cardie, C.: Hierarchical sequential learning for extracting opinionsand their attributes. In: Proceedings of the ACL 2010 conference: short papers, pp. 269–274. Association for Computational Linguistics (2010)
11. Deerwester, S.C., Dumais, S.T., Landauer, T.K., Furnas, G.W., Harshman, R.A.: Indexing by latent semantic analysis. JASIS **41**(6), 391–407 (1990)
12. Duchi, J., Hazan, E., Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res. **12**, 2121–2159 (2011)
13. Dumais, S.T.: Latent semantic analysis. Annu. Rev. Inf. Sci. Technol. **38**(1), 188–230 (2004)
14. Esuli, A., Sebastiani, F.: Sentiwordnet: a publicly available lexical resource for opinion mining. In: Proceedings of LREC, vol. 6, pp. 417–422. Citeseer (2006)
15. Farhadloo, M., Rolland, E.: Multi-class sentiment analysis with clustering and score representation. In: 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW), pp. 904–912. IEEE (2013)
16. Feldman, R.: Techniques and applications for sentiment analysis. Commun. ACM **56**(4), 82–89 (2013)
17. Fellbaum, C.: WordNet. Wiley Online Library, (1998)
18. Filatova, E.: Irony and sarcasm: Corpus generation and analysis using crowdsourcing. In: LREC, pp. 392–398 (2012)
19. Finn, A., Kushmerick, N., Smyth, B.: Genre classification and domain transfer for information filtering. In: Advances in Information Retrieval, pp. 353–362. Springer (2002)
20. Furnas, G.W., Landauer, T.K., Gomez, L.M., Dumais, S.T.: The vocabulary problem in human-system communication. Commun. ACM **30**(11), 964–971 (1987)
21. Gamon, M., Aue, A., Corston-Oliver, S., Ringger, E.: Pulse: mining customer opinions from free text. Advances in Intelligent Data Analysis VI, pp. 121–132. Springer, Berlin (2005)
22. Gibbs, R.W.: On the psycholinguistics of sarcasm. J. Exp. Psychol. Gen. **115**(1), 3 (1986)
23. Gibbs, R.W., Colston, H.L.: Irony in language and thought: a cognitive science reader. Psychology Press (2007)
24. González-Ibánez, R., Muresan, S., Wacholder, N.: Identifying sarcasm in twitter: a closer look. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers, vol. 2, pp. 581–586. Association for Computational Linguistics (2011)
25. Griffiths, T.L., Steyvers, M., Tenenbaum, J.B.: Topics in semantic representation. Psychol. Rev. **114**(2), 211 (2007)
26. Hatzivassiloglou, V. McKeown, K. R.: Predicting the semantic orientation of adjectives. In: Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics, pp. 174–181. Association for Computational Linguistics (1997)
27. Hatzivassiloglou, V., Wiebe, J.M.: Effects of adjective orientation and gradability on sentence subjectivity. In: Proceedings of the 18th conference on Computational linguistics, vol. 1, pp. 299–305. Association for Computational Linguistics (2000)

28. Hofmann, T.: Probabilistic latent semantic indexing. In: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp. 50–57. ACM (1999)
29. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD'04, pp. 168–177, ACM, New York, USA (2004). ISBN:1-58113-888-1
30. Jin, W., Ho, H.H., Srihari, R.K.: Opinionminer: a novel machine learning system for web opinion mining and extraction. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1195–1204. ACM (2009)
31. Jo, Y., Oh, A.H.: Aspect and sentiment unification model for online review analysis. In: Proceedings of the fourth ACM international conference on Web search and data mining, pp. 815–824. ACM (2011)
32. King, R.: Sentiment analysis gives companies insight into consumer opinion. Business Week: technology (2011)
33. Lakkaraju, H., Bhattacharyya, C., Bhattacharya, I., Merugu, S.: Exploiting coherence for the simultaneous discovery of latent facets and associated sentiments. In: SDM, pp. 498–509. SIAM (2011)
34. Lakkaraju, H., Socher, R., Manning, C.: Aspect specific sentiment analysis using hierarchical deep learning. In: Deep Learning and Representation Learning Workshop, NIPS (2014)
35. Lancaster, F.W.: Vocabulary control for information retrieval (1972)
36. Li, F., Han, C., Huang, M., Zhu, X., Xia, Y.-J., Zhang, S., Yu, H.: Structure-aware review mining and summarization. In: Proceedings of the 23rd international conference on computational linguistics, pp. 653–661. Association for Computational Linguistics (2010)
37. Liu, B.: Sentiment analysis and subjectivity. In: Handbook of natural language processing vol. 2, pp. 627–666 (2010)
38. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**(1), 1–167 (2012)
39. Lovins, J.B.: Development of a stemming algorithm. MIT Information Processing Group, Electronic Systems Laboratory (1968)
40. Lu, Y., Zhai, C., Sundaresan, N.: Rated aspect summarization of short comments. In: Proceedings of the 18th international conference on World wide web, pp. 131–140. ACM (2009)
41. Maynard, D., Greenwood, M.A.: Who cares about sarcastic tweets? investi-gating the impact of sarcasm on sentiment analysis. In: Proceedings of LREC (2014)
42. Moghaddam, S., Ester, M.: On the design of lda models for aspect-based opinion mining. In: CIKM'12: Proceedings of the 21st ACM international conference on Information and knowledge management, pp. 803–812, ACM, New York, USA (2012). ISBN:978-1-4503-1156-4, http://doi.acm.org/10.1145/2396761.2396863
43. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retr. **2**(1–2), 1–135 (2008)
44. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, vol. 10, pp. 79–86 (2002)
45. Poria, S., Gelbukh, A., Hussain, A., Das, D., Bandyopadhyay, S.: Enhanced senticnet with affective labels for concept-based opinion mining. Intell. Syst. IEEE, **28**(2) (2013)
46. Porter, M.F.: An algorithm for suffix stripping. Program **14**(3), 130–137 (1980)
47. Resnik, P., Hardisty, E.: Gibbs sampling for the uninitiated. Technical report, Maryland University College Park Institute For Advanced Computer Studies, DTIC Document (2010)
48. Steyvers, M., Griffiths, T.: Probabilistic topic models. Handb. Latent Sem. Anal. **427**(7), 424–440 (2007)
49. Subasic, P., Huettner, A.: Affect analysis of text using fuzzy semantic typing. IEEE Trans. Fuzzy Syst. **9**(4), 483–496 (2001)
50. Svenonius, E.: Unanswered questions in the design of controlled vocabularies. J. Am. Soc. Inf. Sci. **37**(5), 331–340 (1986)

51. Tatemura, J.: Virtual reviewers for collaborative exploration of movie reviews. In: Proceedings of the 5th international conference on Intelligent user interfaces, pp. 272–275. ACM (2000)
52. Turney, P.D.: Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th annual meeting on association for computational linguistics, pp. 417–424. Association for Computational Linguistics (2002)
53. Utsumi, A.: Verbal irony as implicit display of ironic environment: distinguishing ironic utterances from nonirony. J. Pragmat. **32**(12), 1777–1806 (2000)
54. Wallach, H.M.: Topic modeling: beyond bag-of-words. In: Proceedings of the 23rd international conference on Machine learning, pp. 977–984. ACM (2006)
55. Wang, H., Lu, Y., Zhai, C.: Latent aspect rating analysis on review text data: a rating regression approach. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 783–792. ACM (2010)
56. Wang, X., McCallum, A., Wei, X.: Topical n-grams: phrase and topic discovery, with an application to information retrieval. In: Seventh IEEE International Conference on Data Mining, 2007. ICDM 2007, pp. 697–702. IEEE (2007)
57. Ward, K.F., Rolland, E., Patterson, R.A.: Improving outpatient health care quality: understanding the quality dimensions. Heal. Care Manag. Rev. **30**(4), 361–371 (2005)
58. Wiebe, J.: Learning subjective adjectives from corpora. In: AAAI/IAAI, pp. 735–740 (2000)
59. Willett, P.: The porter stemming algorithm: then and now. Program **40**(3), 219–223 (2006)
60. Wong, T.-L., Bing, L., Lam, W.: Normalizing web product attributes and discovering domain ontology with minimal effort. In: Proceedings of the fourth ACM international conference on Web search and data mining, pp. 805–814. ACM (2011)
61. Zhan, T.-J., Li, C.-H.: Semantic dependent word pairs generative model forfine-grained product feature mining. In: Advances in Knowledge Discovery and Data Mining, pp. 460–475
62. Zhu, X., Ghahramani, Z.: Learning from labeled and unlabeled data with label propagation. Technical report, Technical Report CMU-CALD-02-107, Carnegie Mellon University (2002)