Artif Intell Rev
https://doi.org/10.1007/s10462-017-9599-6

A survey on classification techniques for opinion mining and sentiment analysis

Fatemeh Hemmatian[1] · **Mohammad Karim Sohrabi**[1]

Abstract Opinion mining is considered as a subfield of natural language processing, information retrieval and text mining. Opinion mining is the process of extracting human thoughts and perceptions from unstructured texts, which with regard to the emergence of online social media and mass volume of users' comments, has become to a useful, attractive and also challenging issue. There are varieties of researches with different trends and approaches in this area, but the lack of a comprehensive study to investigate them from all aspects is tangible. In this paper we represent a complete, multilateral and systematic review of opinion mining and sentiment analysis to classify available methods and compare their advantages and drawbacks, in order to have better understanding of available challenges and solutions to clarify the future direction. For this purpose, we present a proper framework of opinion mining accompanying with its steps and levels and then we completely monitor, classify, summarize and compare proposed techniques for aspect extraction, opinion classification, summary production and evaluation, based on the major validated scientific works. In order to have a better comparison, we also propose some factors in each category, which help to have a better understanding of advantages and disadvantages of different methods.

Keywords Opinion mining · Sentiment analysis · Machine learning · Classification · Lexicon

1 Introduction

Due to the increasing development of web technology, different evaluation areas are growing in this field. The original web had the static pages and the users didn't allow manipulating

✉ Mohammad Karim Sohrabi
Amir_sohraby@aut.ac.ir

Fatemeh Hemmatian
Fatemehammatian@gmail.com

1 Department of Computer Engineering, Semnan Branch, Islamic Azad University, Semnan, Iran

Published online: 18 December 2017

its contents. Nevertheless, with the advent of new programming technologies, the possibility of interactions and getting feedback on the web pages grew increasingly. The major part of these interactions includes the users' comments, which lead to feedback for the owners of the web pages to benefit from the users' ideas to improve the future performances and causes the products and services adapt with their target group in an appropriate manner. However, manual analysis of such opinions, especially in the social networks with a lot of audience through the world, is very difficult, time consuming and in some cases impossible.

To overcome these limitations, the opinion mining has been introduced as an effective way to discover the knowledge through the expressed comments, especially in the context of the web. Opinion mining or sentiment analysis extracts the users' opinions, sentiments and demands from the subjective texts in a specific domain and distinguishes their polarity. The exponential and progressive increase of internet usage and the exchange of the public thoughts are the main motivations of researches in opinion mining and sentiment analysis. Since several data processing approaches (Sohrabi and Azgomi 2017a, b; Sohrabi and Ghods 2015), supervised and unsupervised machine learning techniques (Sohrabi and Akbari 2016), data mining and knowledge discovery methods, including association rule mining (Sohrabi and Marzooni 2016), frequent itemset mining (Sohrabi and Barforoush 2012, 2013; Sohrabi and Ghods 2014; Sohrabi 2018), and sequential pattern mining (Sohrabi and Ghods 2016; Sohrabi and Roshani 2017), with various applications (Arab and Sohrabi 2017; Sohrabi and Tajik 2017; Sohrabi and Karimi 2018), and web mining approaches (Zhang et al. 2004; Sisodia and Verma 2012), including web structure mining (WSM) (Velásquez 2013), web usage mining (WUM) (Yin and Guo 2013), and web content mining (WCN) (Mele 2013), have been represented in the literature, there are different choices to select techniques and provide methods for opinion mining and sentiment analysis.

The research about the opinion mining began from the early 2000, but the phrase "opinion mining" was firstly used in Dave et al. (2003) (Liu 2012). In the past 15 years, various researches have been conducted to examine and analyze the opinions within news, articles, and product and service reviews (Subrahmanian and Reforgiato 2008). Nowadays, most people benefit from the opinions of different people by a simple search on the Internet when buying a commodity or selecting a service. According to the study conducted in Li and Liu (2014), 81% of the Internet users have searched related comments before buying a commodity at least once. The search rates in related comments before using restaurants, hotels and a variety of other services have been reported from 73 to 87%. It should be noted that these online investigations had a significant impact on the customer's decisions. People's sentimental ideas and theories can be extracted from different web resources, such as blogs (Alfaro et al. 2016; Bilal et al. 2016), review sites (Chinsha and Joseph 2015; Molina-González et al. 2014; Jeyapriya and Selvi 2015), and recently micro-blogs (Balahur and Perea-Ortega 2015; Feng et al. 2015; Pandarachalil et al. 2015; Da Silva et al. 2016; Saif et al. 2016; Wu et al. 2016; Ma et al. 2017; Li et al. 2017; Keshavarz and Abadeh 2017; Huang et al. 2017). Micro-blogs, such as Twitter, have become very popular among users and provides the possibility of sending tweets up to a specified limited number of characters (Liu 2015).

Opinion mining, can take place in three levels of the document (Sharma et al. 2014; Moraes et al. 2013; Tang et al. 2015; Sun et al. 2015; Xia et al. 2016), sentence (Marcheggiani et al. 2014; Yang and Cardie 2014) and aspect (Chinsha and Joseph 2015; Marrese-Taylor et al. 2014; Wang et al. 2017b). Also all techniques which used to sentiment analysis can be categorized into three main classes as: machine learning techniques (Pang et al. 2002; Moraes et al. 2013; Saleh et al. 2011; Habernal et al. 2015; Riaz et al. 2017; Wang et al. 2017a), lexicon-based approaches (Kanayama and Nasukawa 2006; Dang et al. 2010; Pandarachalil

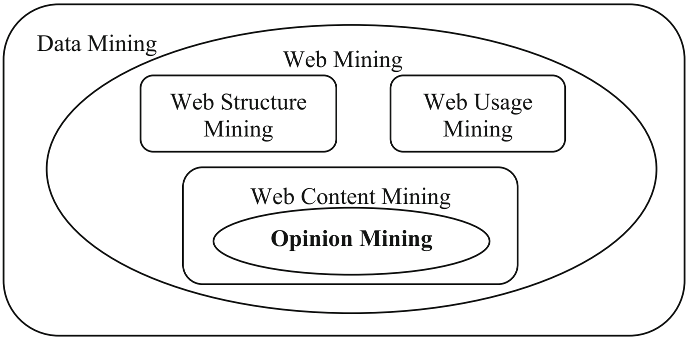

Fig. 1 The position of opinion mining

et al. 2015; Saif et al. 2016; Taboada et al. 2011; Turney 2002; Molina-González et al. 2015; Qiu et al. 2011; Liao et al. 2016; Bravo-Marquez et al. 2016; Muhammad et al. 2016; Khan et al. 2017) and hybrid methods (Balahur et al. 2012; Abdul-Mageed et al. 2014; Keshavarz and Abadeh 2017). The machine learning-based opinion mining techniques which have the benefit of using well-known machine learning algorithms, can be divided into three groups: supervised (Jeyapriya and Selvi 2015; Habernal et al. 2015; Severyn et al. 2016; Anjaria and Guddeti 2014), semi-supervised (Hajmohammadi et al. 2015; Hong et al. 2014; Gao et al. 2014; Carter and Inkpen 2015; Lu 2015) and unsupervised (Li and Liu 2014; Claypo and Jaiyen 2015; De and Kopparapu 2013) methods. Lexicon-based method relies on a dictionary of sentiments and has been highly regarded in the recent studies which can be divided into the dictionary-based method (Chinsha and Joseph 2015; Pandarachalil et al. 2015; Saif et al. 2016; Sharma et al. 2014) and corpus-based method (Turney 2002; Molina-González et al. 2015; Keshtkar and Inkpen 2013; Vulić et al. 2015). There are also very few works that are used both corpus-based and dictionary-based methods to improve the results (Taboada et al. 2011). Some literature reviews and books on opinion mining and sentiment analysis techniques and methods have also been represented before, which have investigated the problem from different points of views (Bouadjenek et al. 2016; Liu 2015).

The rest of paper is organized as follows: The clear explanation of the problem, process, tasks and applications of opinion mining has been represented in Sect. 2. Section 3 defines the levels of opinion mining. Section 4 focuses on extraction of aspects. The classification and comparison of sentiments analysis techniques are presented in Sect. 5. The evaluation criteria in the opinion mining are discussed in Sect. 6, the future direction of opinion mining are represented in Sect. 7, and finally Sect. 8 concludes the review.

2 Opinion mining: process, tasks, and applications

Opinion mining can be considered as a new subfield of natural language processing (Daud et al. 2017), information retrieval (Scholer et al. 2016), and text mining (Singh and Gupta 2017). Figure 1 represents the position of opinion mining. Opinion mining is actually considered as a subset of the web content mining process in the web mining research area. Since the web content mining focuses on the contents of the web and texts have formed large volume of web content, text mining techniques are widely used in this area. The most important challenge of using text mining in web content is their unstructured or semi-structured nature that requires the natural language processing techniques to deal with. Web mining itself is also considered a subset of the data mining research area. Here, the use of data mining is to discover the knowledge from massive data sources of the web.

2.1 Opinion mining definitions

The main goal of opinion mining is to automate extraction of sentiments expressed by users from unstructured texts. Two major definitions of opinion mining can be seen in the literature. The first definition is proposed in Saleh et al. (2011), as "The automatic processing of documents to detect opinion expressed therein, as a unitary body of research". The second major definition says: "Opinion mining is extracting people's opinion from the web. It analyzes people's opinions, appraisals, attitudes, and emotions toward organizations, entities, person, issues, actions, topic and their attribute" (Jeyapriya and Selvi 2015; Liu 2012; Liu and Zhang 2012).

Opinion mining contains several tasks with different names which all of them are covered by it (Liu 2012):

- **Sentiment Analysis** The purpose of sentiment analysis is the sentiment recognition and public opinion examination that is considered as a research area in the field of text mining.
- **Opinion extraction** The process of extraction of users' opinions from the web documents is called opinion extraction. The main purpose of opinion extraction is to find out the users' ways of thinking.
- **Sentiment mining** Sentiment mining has two main goals: first, it determines whether the given text contains objective or subjective sentences. A sentence is called objective (or factual), when it contains the factual information about the product. The subjective sentences represent the individual emotions about the desired product. In the opinion mining we consider the subjective sentences. Second, it extracts opinions and classifies them into three categories of positive, negative and neutral (Farra et al. 2010).
- **Subjection analysis** Subjection analysis provides the possibility to identify, classify, and collect subjective sentences.
- **Affect or emotion analysis** Many of the words at the text are emotionally positive or negative. Affect analysis specifies the aspects that are expressing emotions in the text using the natural language processing techniques (Grefenstette et al. 2004).
- **Review mining** Review mining is a sub-topic of text sentiment analysis and its main purpose is to extract aspects from the authors' sentiments and is to produce a summary of the sentiments. More researches in the review mining have been focused on the product reviews (Zhuang et al. 2006).

2.2 Opinion mining procedure

The main objective of the opinion mining is to discover all sentiments exist in the documents (Saleh et al. 2011); in fact, it determines the speaker's or writer's attitude about the different aspects of a problem. We have modeled the opinion mining process in Fig. 2, in which, each part has some obligations which are as follows:

1. **Data collection** Having a comprehensive and reliable dataset is the first step to perform opinion mining process. The necessary information could be collected from various web resources, such as weblogs, micro blogs (such as Twitter[1]), social networks (such as

[1] https://www.twitter.com/.

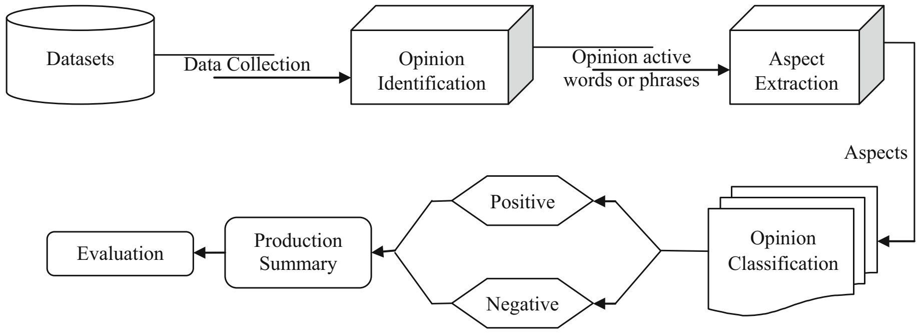

Fig. 2 Opinion mining process

Facebook[2]) and review websites (such as Amazon,[3] Yelp,[4] and Tripadvisor[5]). Using tools that are developed for extracting data through web, and using various techniques such as web scraping (Pandarachalil et al. 2015), can be useful to collect appropriate data. Some datasets are provided in English which can be used as references (Pang et al. 2002; Pang and Lee 2004; Blitzer et al. 2007). Researchers can apply their methods on these datasets for their simplicity. The first dataset[6] prepared by Pang et al. (2002) includes 1000 positive movie reviews and 1000 negative movie reviews. This dataset is the most important and the oldest dataset in this area. The second dataset[7] prepared by Pang and Lee (2004), which includes 1250 positive reviews, 1250 negative reviews, and 1250 neutral reviews. The Third one is Blitzer (Blitzer et al. 2007),[8] which includes 1000 positive movie review and 1000 negative movie reviews. Table 1 shows the obtained accuracies of different researches on the benchmark datasets.

2. **Opinion identification** All the comments should be separated and identified from the presented texts in this phase. Then the extracted comments should be processed to separate the inappropriate and fake ones. What we mean by opinions is all the phrases representing the individual emotions about the products, services or any other desired category.
3. **Aspect extraction** In this phase, all the existing aspects are identified and extracted according to the procedures. Selecting the potential aspects could be very effective in improving the classification.
4. **Opinion classification** After opinion identification and aspect extraction which can be considered as the preprocessing phase, in this step the opinions are classified using different techniques which this paper summarizes, classifies and compares them.
5. **Production summary** Based on the results of the previous steps, in the production summary level, a summary of the opinion results is produced which can be in different forms such as text, charts etc.
6. **Evaluation** the performance of opinion classification can be evaluated using four evaluation parameters, namely accuracy, precision, recall and f-score.

2 https://www.facebook.com/.

3 http://www.amazon.com/.

4 http://www.yelp.com/.

5 http://www.tripadvisor.com/.

6 http://www.cs.cornell.edu/people/pabo/movie-review-data/.

7 http://www.cs.cornell.edu/people/pabo/movie-review-data/ (review corpus version 2.0).

8 http://www.cs.jhu.edu/~mdredze/datasets/sentiment/.

Table 1 Obtained accuracies on the benchmark datasets

Datasets	Papers	Accuracy (%)
Pang et al. (2002)	Chen et al. (2011)	64
	Li and Liu (2012)	77
Pang and Lee (2004)	Penalver-Martinez et al. (2014)	89.6
	Fernández-Gavilanes et al. (2016)	69.95
	Fersini et al. (2016)	81.7
	Saleh et al. (2011)	85.35
	Boiy and Moens (2009)	87.40
Blitzer et al. (2007)	Xia et al. (2016)	80
	Xia et al. (2011)	85.58
	Poria et al. (2014)	87

2.3 Opinion mining applications

Sentiment analysis tries to describe and assess the expressed sentiments about the issues of interest to web users which have been mentioned in textual messages. These issues can include a range of brands or goods up to the broader favorite topics such as social, political, economic and cultural affairs. We note to the several major applications of the opinion mining in this section.

2.3.1 Opinion mining in the commercial product areas

The usage of opinion mining in the area of commercial products (Chen et al. 2014; Marrese-Taylor et al. 2014; Jeyapriya and Selvi 2015; Li et al. 2012; Luo et al. 2015) is important from three viewpoints:

1. The individual customers' point of view: when someone wants to buy a product, having a summary of the others' opinions can be more useful than studying the massive amounts of others' comments about this product. Moreover, the customer will be able to compare the products easily by having a summary of the opinions.
2. The business organizations and producers' point of view: this issue is important for the organizations to improve their products. This information is used not only for the product marketing and evaluation but also for product design and development. The manufacturing companies can even increase, decrease or change the products based on customer's opinions.
3. The advertising companies' point of view: the opinions are important for advertising companies because they can obtain ideas of the market demand. The public perspective of the people and type of products that they are interested in can be found among the items that extracted by opinion mining.

The important achievements of opinion mining in the commercial products are as follows (Tang et al. 2009):

- **Products comparison** Online sellers want their customers to comment about the purchased products. Due to the increasing use of the online marketing and such web services, these sentiments are growing. These sentiments are useful both for product manufacturers and consumers because they can have a better decision making by comparing the sentiments and ideas of others on this product. More researches have been carried out in this area, which have focused on the issue of automatic classification of the products in two categories of recommended and non-recommended.

- **Sentiments summarization** When the number of sentiments increase, its recognition is difficult either for producers or consumers. With the sentiments summarization, customers find out easier the sentiments of other customers about the product and also manufacturers realize easier to the customers' sentiments about the products as well.
- **Exploring the reason of opinion** The reason of the user to give an opinion can also be extracted in the opinion mining process. It is extremely important to determine the reason why consumers like or dislike the product.

2.3.2 Opinion mining in the politics area

Along with the comments on the sale and purchase of goods, with the widespread and comprehensive use of the Internet services by people, users can also comment on various political, social, religious, and cultural issues. Collecting and analyzing these comments helps greatly to politicians, managers of social issues or religious and cultural activists to take appropriate decisions for improving the social life of the community. One of the significant applications among these areas is in the political elections that individuals can benefit from the sentiments of others to make decision in their voting. Analyzing opinions existed in social networks related to election is addressed in (Tsakalidis et al. 2015; Unankard et al. 2014; Kagan et al. 2015; Mohammad et al. 2015; Archambault et al. 2013).

2.3.3 Opinion mining in the stock market and stock forecast

Achieving sustained and long-term economic growth requires optimal allocation of resources at the national economy level and this is not easily possible without the help of appropriate information and knowledge. Investing in supplied stocks in the stock exchange is one of the profitable options in the capital market which plays an important role in the individuals' better decision making and having its own particular audience which predicting the stock. Among the studies representing the application of the opinion mining in the stock market it can be pointed out to (Bollen et al. 2011; Nofer and Hinz 2015; Bing et al. 2014; Fortuny et al. 2014) that the opinions have been used to predict the stock market. For example, Daily comments of Twitter have been analyzed using OpinionFinder and GPOMS as two important moods tracking tools by Bollen et al. (2011) and showed the correlation to daily changes in Dow Jones Industrial Average closing values.

3 Levels of opinion mining

As shown in Fig. 3, opinion mining is possible on four different levels, namely document level, sentence level, aspect level, and concept level.

Document level (Moraes et al. 2013) of opinion mining is the most abstract level of sentiment analysis and so is not appropriate for precise evaluations. The result of this level of analysis is usually general information about the documents polarity which cannot be very accurate. Sentence level opinion mining (Marcheggiani et al. 2014) is a fine-grain analysis that could be more accurate. Since the polarity of the sentences of an opinion does not imply the same polarity for the whole of opinion necessarily, aspect level of opinion mining (Xia et al. 2015) have been considered by researchers as the third level of opinion mining and

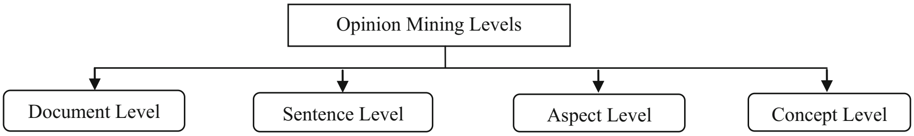

Fig. 3 Different levels of opinion mining

sentiment analysis. Concept level opinion mining is the forth level of sentiment analysis which focuses on the semantic analysis of the text and analyzes the concepts which do not explicitly express any emotion (Poria et al. 2014). Several recent surveys and reviews on sentiment analysis consider these levels of opinion mining from this point of view (Medhat et al. 2014; Ravi and Ravi 2015; Balazs and Velasquez 2016; Yan et al. 2017; Sun et al. 2017; Lo et al. 2017).

3.1 Document level

The sentiment analysis may be used in the document level. In this level of the opinion mining, sentiments are ultimately summarized on the whole of the document as positive or negative (Pang et al. 2002). The purpose of categorizing comments at the document level is the automatic classification of information based on a single topic, which is expressed as a positive or negative sentiment (Moraes et al. 2013). Since this level of opinion mining does not enter into details and the review process takes place in an abstract and general view, the mining process can be done much faster. In early works, most of the researches conducted at the document level and focused on datasets such as the news and the products review. By increase in the popularity of the social networks, different types of datasets were created which made increasing the studies of this level (Habernal et al. 2015; Gupta et al. 2015). Since the entire document is considered as a single entity in document level opinion mining, this level of opinion mining is not suitable for precise evaluation and comparison. Most of the techniques carrying out in the opinion classification at this level are based on supervised learning methods (Liu and Zhang 2012).

3.2 Sentence level

Since, document level sentiment analysis is too coarse, researchers investigated approaches to focus on the sentence (Wilson et al. 2005; Marcheggiani et al. 2014; Yang and Cardie 2014; Appel et al. 2016). The goal of this level of opinion mining is to classify opinions in each sentence. Sentiments analysis on the sentence level constitutes of two following steps (Liu and Zhang 2012):

- Firstly it is determined that the sentence is subjective or objective.
- Secondly the polarity (positive or negative) of sentence is determined.

In the classification of comments at the sentence level, since the documents are broken into several sentences, they provide more accurate information on the polarity of the views and naturally entail more challenges than the level of the document.

3.3 Aspect level

Although the classification of text sentiments on the document and sentence level is helpful in many cases but it does not provide all the necessary details. For example, being positive of the sentiments on a document in relation to a particular entity, does not imply that the author's opinion is positive about all the aspects of an entity. Similarly, negative sentiments do not represent the author negative opinion about all the aspects of an entity (Liu and Zhang 2012). The classification on the document level (Moraes et al. 2013) and sentence level (Marcheggiani et al. 2014) does not provide these kinds of information and we need to perform opinion mining in aspect level (Xia et al. 2015) to achieve these details. When the considered comment does not include a single entity or aspect, this level of opinion mining is the appropriate option, which is an important advantage of this level of classification and distinguishes it from the two previous levels. Aspect level opinion mining actually considers the given opinion itself instead of looking to the language structures (document, sentence or phrase) (Liu 2012). The objective of this level is to identify and extract the aspects from the sentiments text and then specify their polarity. This level of sentiments analysis can produce a summary of the sentiments about different aspects of the desired entity. It can be seen that this level of opinion mining provides a more accurate result (Chinsha and Joseph 2015).

3.4 Concept level

Cambria (2013) introduced the concept level opinion mining as a deep understanding of the natural language texts by the machine, in which, the opinion methods should go beyond the surface level analysis. Cambria et al. (2013) has also presented the concept level of opinion mining as a new avenue in the sentiment analysis. The analysis of emotions at the concept level is based on the inference of conceptual information about emotion and sentiment associated with natural language. Conceptual approaches focus on the semantic analysis of the text and analyze the concepts which do not explicitly express any emotion (Poria et al. 2014). An enhanced version of SenteicNet have been proposed in Poria et al. (2013), which assign emotion labels to carry out concept level opinion mining. Poria et al. (2014) have proposed a new approach to improve the accuracy of polarity detection. An analysis of comments at the conceptual level has been introduced that integrates linguistic, common-sense computing, and machine learning techniques. Their results indicate that the proposed method has a desirable accuracy and better than common statistical methods. A concept level sentiment dictionary has been built in Tsai et al. (2013) based on common-sense knowledge using a two phase method which integrates iterative regression and random walk with in-link normalization. A concept level sentiment analysis system has been presented in Mudinas et al. (2012), which combined lexicon-based and learning based approaches for concept mining from opinions. EventSensor system is represented in Shah et al. (2016) to extract concept tags from visual contents and textual meta data in concept-level sentiment analysis.

4 Aspect extraction

One of the main important steps in sentiments classification is aspect extraction (Rana and Cheah 2016). In this section we categorize current techniques for aspect extraction and selection. As it mentioned, aspect level classification has better performance and a prerequisite for using it is obtaining aspects. Most researches in the field of aspect extraction have been focused on the online reviews (Hu and Liu 2004; Li et al. 2015; Lv et al. 2017). In general,

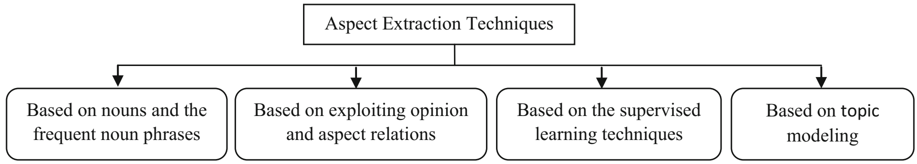

Fig. 4 Classification of aspect extraction techniques

as is shown in Fig. 4, the related techniques can be placed in four categories (Liu 2012): *Extraction based on the frequent noun phrases and nouns* (Jeyapriya and Selvi 2015; Hu and Liu 2004; Li et al. 2015), *Extraction based on exploiting opinion and aspect relations* (Qiu et al. 2011; Wu et al. 2009), *Extraction based on the supervised learning* (Jin et al. 2009; Yu et al. 2011), *Extraction based on topic modeling* (Vulić et al. 2015; Mukherjee and Liu 2012).

4.1 Extraction based on frequency of noun phrases and nouns

This method is known as a simple and effective approach. Generally, when people express their comments about various aspects of a product, they basically use similar words frequently to express their sentiments (Liu 2012). In this method, the nouns and noun phrases are determined by a POS tagger and the names that have been frequently repeated are selected as aspect. POS tags indicate the role of the words in a sentence (Wang et al. 2015a). A list of POS tags has been collected in Table 2 which shows all the POS tags based on (Liu 2012).

Li et al. (2015) suggested a method for improving feature extraction performance by online reviews. Their method which is based on frequent noun and noun phrase is consisted of three important components: frequent based mining and pruning, order based filtering, and similarity based filtering. Their experimental results show that proposed method could be generalized over various domains with different-sized data. Jeyapriya and Selvi (2015) suggested a feature extraction system in product review. They extracted nouns and noun phrase from each review sentence and used minimum support threshold to find frequent features in review sentences. Their accuracy was about 80.32

4.2 Extraction based on relation exploitation between opinion words and aspects

This method uses the existing relationship between the aspects and opinion words in the expressed opinions. Some of the infrequent aspects can be identified with the help of this method. The main idea is that the opinion words can be used to describe the different aspects (Liu and Zhang 2012). Qiu et al. (2011) focused on two fundamental and important issues, opinion lexicon expansion and target extraction, and suggested the double propagation approach. They performed extraction based on syntactic relations that cause link between review words and targets. Relations could be detected according to a dependency parser and then could be used for opinion lexicon expansion and extraction. Results show that their approach is significantly better than existing methods.

4.3 Extraction based on the supervised learning

Supervised learning approaches are promising techniques for aspect extraction which generates a model of aspects by using labeled data. Support vector machine (SVM) (Cortes and Vapnik 1995; Manek et al. 2017), conditional random fields (CRF) (Lafferty et al. 2001), and

Table 2 POS tags (Liu 2012)

Description	Tag	Description	Tag	Description	Tag
Adjective	JJ	Comparative adjectives	JJR	Superlative adjectives	JJS
Adverb	RB	Comparative adverb	RBR	Superlative adverb	RBS
Noun, plural noun, singular	NNS	noun, singular or mass	NN	Comparative adjective	JJR
Coordinating conjunction	CC	Subordination conjunction	IN	Interjection	UH
Determiner	DT	Verb, gerund or present participle	VBG	List item marker	LS
Model verb	MV	Cardinal number	CD	Adjective	JJ
Verb, past participle	VBN	Verb, gerund or present participle	VBG	Particle	RP
Verb, past tense	VBD	Personal pronoun	PP	Possessive ending	POS
Verb, non-3rd-person singular/p	VBP	Verb, 3rd-person singular present	VBZ	Proper noun, plural	NPS
Proper noun, singular	NP	Symbol	SYM	Verb, base form	VB

hidden Markov model (HMM) (Rabiner 1989) are some of the supervised learning methods that can be used to extract the aspects (Liu 2012). In Kobayashi et al. (2007), aspects was extracted from a collection of blog posts using machine learning methods and the results was used as statistical patterns for aspect extraction.

4.3.1 Extraction based on topic modeling

In recent years the statistical topic models are considered as a systematic approach to detect the topics from the text document collections (Vulić et al. 2015; Mukherjee and Liu 2012; Liu 2012). Topic modeling is an unsupervised method for aspect extraction in which it assumes that any document contains k hidden topics. For example in hotel investigation, some standard features such as location, cleanliness and so on are discussed. Now it is possible that there are comments about the quality of internet connection, so we are facing with a hidden topic. In these situations, there is a need for a model to automatically extract relevant aspects without human supervision (Titov and McDonald 2008). Since this approach, uses statistical methods like latent semantic analysis (LSA) (Hofmann 1999) and latent Dirichlet allocation (LDA) (Blei et al. 2003), it is called statistical models too. Also LDA and LSA use the bag of words represented in documents, so they can be used only in document level opinion mining. Ma et al. (2015) proposed an approach of probabilistic topic model based on LDA in order to semantic search over citizens opinions about city issues on online platforms. Their results show that systems based on LDA provide useful information about their staff members. Luo et al. (2015) worked on detection and rating feature in product review which is as important task of opinion mining in aspect level. They presented Quad-tuple PLSA for solving this problem because entity and rating rarely considered in previous researches hence had great performance.

5 Opinion classification techniques

Most important and critical step of opinion mining is selecting an appropriate technique to classify the sentiments. In this section we explain, categorize, summarize and compare proposed techniques in this area. The classification methods which are proposed in the literature can be fall into two groups: *machine learning* and l*exicon-based* approaches. This type of categorization can be seen in some works (Wang et al. 2015a; Petz et al. 2015), but in this paper we address the issue much more comprehensive in more details with better comparison factors and discussions based on the major validated scientific works.

5.1 Machine learning method

According to Fig. 5 three techniques of the machine learning methods are used to classify the sentiments: supervised, semi-supervised, and unsupervised learning methods.

5.1.1 Supervised learning method

In the supervised learning is, the process of learning is carried out using the data of a training set in which, the output value is specified for any input and the system tries to learn a function, by mapping the input to the output, i.e., to guess the relationship between input and output. In this method, the categories are initially specified and any of the training data is assigned to a specific category. In fact in the supervised approach, the classifier categorizes the sentiments

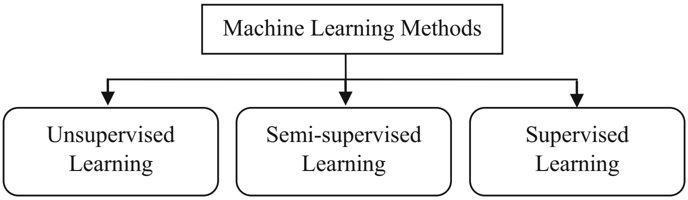

Fig. 5 Machine learning based opinion classification techniques

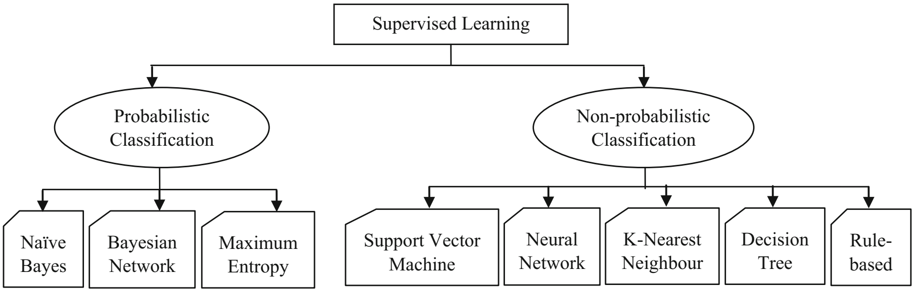

Fig. 6 Supervised learning-based opinion classification methods

using labeled text samples. As shown in Fig. 6, supervised sentiment classification approaches can be divided into two main categories: *Probabilistic Classification* and *Non-probabilistic Classification*.

5.1.1.1. Probabilistic classification Probabilistic classification is one of the popular classifications approaches in the field of the machine learning. These methods are derived from probabilistic models which provide a systematic way for statistical classification in complex domains such as the natural language. Hence, it has an effective application in the opinion mining. *Naive Bayes, Bayesian network* and *maximum entropy* are some of the well-known methods in the field of opinion mining which belong to this kind of classification.

Naive Bayes (NB)

This method is a simple and popular approach in the area of text classification. It is assumed that the existing sentences within the document are subjective which the existence of semantic orientation of words is definitely a final verdict on the subjectivity of the sentences. The features are also selected from a set of words within the documents. It is an approach to text classification that assigns the class $c^* = \text{Arg max}_c\ P(c|d)$, to a given document d. The relation 1 is expressed based on the Bayesian theory (Pang et al. 2002).

$$P(c|d) = \frac{p(c)p(d|c)}{p(d)} \tag{1}$$

where $p(d)$ plays no role in selecting c^*. To estimate the term $p(d|c)$, Naïve Bayes decomposes it by assuming the f_i's are conditionally independent given d's class as in relation 2 (Pang et al. 2002).

$$P_{NB}(c|d) = \frac{P(C)\left(\pi_{i=1}^{m} P\left(f_i|c\right)^{n_i(d)}\right)}{p(d)} \tag{2}$$

where m is the number of features and f is the feature vector.

Bilal et al. (2016) compared efficiencies of three techniques, namely Naive-Bayes, decision tree, and nearest neighbor for classifying Urdu and English opinions in a blog. Their results show that Naïve-Bayes has better performance than two other techniques. Various opinion mining methods have been used naïve Bayes as a probabilistic classifier (Hasan et al. 2015; Parveen and Pandey 2016; Goel et al. 2016; Ramadhani et al. 2016)

Bayesian networks (BN)

The Bayesian network is a probabilistic graphical model representing the relationship between random variables. This model consists of a directed acyclic graph and a set of conditional probability distributions for each of the network variables (Sierra et al. 2009). The network has an extended structure and it is easy to add new variables. In fact, it is a method to describe the joint probability distribution of a set of variables that has described the conditional independence of a set of variables and provides the possibility for combining the prior knowledge with training data about the dependence of variables. In this model the relation 3 (Sierra et al. 2009) is used to calculate the probability distribution of a set of variables. According to this relation variable x is independent from other variables if it has parents:

$$P\,(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(x_i|parents(x_i)) \tag{3}$$

A hierarchical Bayesian network has been used by authors in Ren and Kang (2013) to build a model for the analysis of the human beings emotions. It finds complex emotions in the document by establishing a relationship between the topic modeling and analyzing the emotions. The results indicate that this method is able to operate in complex domains. In Kisioglu and Topcu (2011), a research has also been carried out to find out the most important factors affecting the customers of telecommunications industry considering the benefits of Bayesian network. In this study, 2000 customers' data from Turkish Telecommunication Company has been used as the dataset.

Maximum entropy (ME)

Another probabilistic based approach that is used to sentiment classification is maximum entropy (Pang et al. 2002; Habernal et al. 2015). This method has had an enormous impact in the natural language processing applications (Berger et al. 1996). Since unlike Naïve Bayes, maximum entropy makes no assumptions about the relationships between features, it might potentially perform better when condition independence assumptions are not met (Pang et al. 2002). This method is also known as the exponential classifiers because of having exponential formula shown in relation 4.

$$P_{ME}(c|d) = \frac{1}{z(d)} \exp\left(\sum_i \lambda_{i,c} f_{i,c}(\mathrm{d}, \mathrm{c})\right) \tag{4}$$

In relation 4, $z(d)$ is the normalization function, $\mathrm{f}_{\mathrm{i,c}}$ is a function for the feature f_{i} and the category c, and $\lambda_{i,c}$ is a parameter for the feature weight which has been defined as relation 5 (Pang et al. 2002).

$$f_{i,c}(d, \acute{c}) = \begin{cases} 1, & n_i(d) > 0\, and\, \acute{c} = c \\ 0, & other \end{cases} \tag{5}$$

The model with maximum entropy is the one in the parametric family P($c|d$) maximizing the likelihood (Xia et al. 2011). Numerical methods such as, iterative scaling algorithm and Gaussian prior (Pang et al. 2002) optimization are usually employed to solve the optimization problem. Habernal et al. (2015) performed a deep research on machine learning methods in order to sentiment analysis on social media in Czech. They used two classifiers, maximum entropy and support vector machine, for Facebook dataset; furthermore, for other datasets, maximum entropy is used due to computational possibility for classifying opinions into positive, negative and neutral classes. The research's results show that maximum entropy had better performance rather than support vector machine in most of cases. There are also some other recent researches which is used the maximum entropy foe sentiment analysis (Yan and Huang 2015; Ficamos et al. 2017).

5.1.1.2. Non-probabilistic Classification In some situations, according to the terms of the problem, the probabilistic classifiers cannot be effective. In this case, the other option for doing the classification is using non-probabilistic classifiers. Among all non-probabilistic classifiers, *neural network*, *support vector machine (SVM), nearest neighbor, decision tree* and r*ule-based methods* are widely used in sentiment analysis.

Support vector machines (SVM)

SVM (Cortes and Vapnik 1995) is one of the most popular supervised classification methods which has a robust theoretical base and according to the research report (Liu 2007), is likely the most precise method in text classification, which makes it common in sentiment classification (Mullen and Collier 2004; Saleh et al. 2011; Kranjc et al. 2015). Also it shows to be highly effective at traditional text categorization, generally out performing Naïve Bayes (Joachims 1998). SVM finds optimal hyper plane to divide classes (Moraes et al. 2013). The most widely known research to apply the classification on document-level sentiment analysis was conducted by Pang et al. (2002). It used 700 positive and 700 negative labeled documents as training data to build a model with naïve Bayes, maximum entropy and SVM which the best empirical results is obtained by using SVM. Tripathy et al. (2016) used four different machine learning algorithms including stochastic gradient descent, support vector machine, naïve-Bayes and maximum entropy by applying n-gram approach. They reached an acceptable accuracy in their research. Severyn et al. (2016) used SVM for opinion mining on YouTube. Their aim was detecting type and polarity of opinions. Saleh et al. (2011) conducted an experiment on SVM method to classify opinions in various domains by applying some weighting procedure. Their results show that SVM method is a promising and desirable method which can overcome opinion classification.

Artificial neural network (ANN)

Among different machine learning algorithm, ANN has absorbed less attention in this area but recently has attracted more attention and popularity (Tang et al. 2015; Vinodhini and Chandrasekaran 2016). The key idea of ANN is to extract features from linear combination of the input data, and then models output as a nonlinear function of these features. Neural networks are usually displayed as a network diagram which involves nodes connected by links. Nodes are arranged in a layer and the architecture of common neural networks includes three layers: input layer, output layer and a hidden layer. There are two types of neural networks, feed forward and back forward (Moraes et al. 2013). Since in forward network the nodes are only connected in one direction, it is suitable for sentiment classification (Chen et al. 2011). Each connection has a corresponding weight value which is estimated by minimizing a global error function in a gradient descent training process. A neuron is a simple mathematical model which outputs a value in two steps. In first step, the neuron calculates a weighted sum of its input and then obtains its output by applying an activation function to this sum. The activation function is typically a nonlinear function and it ensures that the whole network can estimate a nonlinear function which is previously learned from the input data (Moraes et al. 2013).

In Tang et al. (2015), authors proposed a novel neural network method to investigate review rating prediction regarding user information. This method involves two composition methods: User-Word Composition Vector Model (UWCVM), and Document Composition Vector Model (DCVM). UWCVM modifies the original word vector by user information. Modified word vectors are then entered in DCVM to produce the review representation which is regarded as feature to predict review rating. In order to examine prediction rate, UMCVM is integrated into a feed-forward neural network. The results of DCVM are used as features to make rating predictor without any feature engineering. The neural network parameters are trained in an end-to-end fashion with back propagation. In the area of sentiments analysis, the main deficiency of neural network is that the training time is high.

In Chen et al. (2011), it is suggested to combine the features for building a model based on neural network with a few number of input neuron. In this paper, a neural network approach has been suggested for sentiment classification in the blog sphere. Thus, SO-A, SO-PMI (AND), SO-PMI (Near), and SO-LSI are used as the input neurons of back propagation network. The suggested approach is not only shows higher accuracy in classification, but also is better in training time. Jian et al. (2010) has adopted an ANN-based individual model for simulating the human's judgment on sentiment polarity. Practical results indicate that precision of individual model is higher than support vector machine and hidden Markov model classifiers on movie review corpus. Moraes et al. (2013) has compared two methods of support vector machine and neural network for sentiments analysis on the document level. Their testing results indicate that the neural network has superiority or at least comparable results with SVM. In Duncan and Zhang (2015), authors have examined a feed-forward neural network for sentiment analysis of tweets. Twitter API has been used to collect the training set of tweets with positive and negative keywords. Using the same keywords, testing set of tweets has been collected too. The experimental results show that memory is an important issue in the case that feed-forward pattern network will train with very large vocabularies. Consequently, in this case, there is not enough memory to hold the data structures that are required for training the feed-forward pattern network once the vocabulary is too large.

The computational cost of the neural network is high for training phase in opinion mining and its accuracy is highly dependent on the number of training data. Convolutional neural networks (CNN) are appropriate alternatives which have high performance in the classification of emotions. Chen et al. (2016) has considered temporal relation of reviews and have made a sequential model to improve the performance of the sentiment analysis at the document level. A one-dimensional convolutional neural network has been proposed in this paper to learn the distributed representation of each review. A recurrent neural network was also used then to learn the distributed representation of users and products. Finally, the machine learning has been used to classify the comments. A dynamic convolutional neural network has been exploited in Kalchbrenner et al. (2014) to encounter input sentences that have variable lengths. This method does not require the provision of external properties by the parser. A CNN-based opinion summarization method is proposed Li et al. (2016) to extract a proper set of features for opinion mining on the Chinese micro-blogging systems. Two levels of convolutional neural networks have been combined in Gu et al. (2017) to construct a cascade convolutional neural network (C-CNN). CNNs of the first level do the aspect mapping and the single CNN of the second level specifies polarity of opinions. Some CNN-based opinion mining methods have been also proposed for multimedia sentiment analysis (Cai and Xia 2015).

The first deep learning based approach for feature extraction was represented in Poria et al. (2016). In this paper, a seven-layer architecture deep convolutional neural network has been proposed to tag each word in the sentence as a feature or a non-feature. A multimodal sentiment analysis framework for opinion mining from video contents has been proposed in

Poria et al. (2017), in which audio and textual modalities have been combined by multiple kernel learning using a convolutional neural network. Deep recurrent neural networks (RNN) have been also used for opinion mining. For example a deep RNN in Irsoy and Cardie (2014) has been applied on an opinion mining process which is modeled as a token level sequence labeling task. Some of applications of opinion mining have been also covered by deep convolutional neural networks. For example, Ebrahimi et al. (2016) proposed a deep CNN to detect predatory conversations in social media.

K-nearest neighbor (KNN)
KNN is one of the most popular instances of learning-based methods (Tsagkalidou et al. 2011). In this method, K is the number of considered neighbors which is usually odd, and the distance to these neighbors are determined based on the standard Euclidean intervals (Duwairi and Qarqaz 2014). The underlying assumption in this method is that, all the samples are real points in n-dimensional space. In general, this algorithm is used for two purposes: to estimate the distribution density function of training data as well as to classify testing data based on the training patterns. Suppose $c_1 \ldots c_n$ is a predefined set of categories and any of these categories also includes their own data. According to relation 6, if we want to assign a new data such as x using this method to one of the categories, the nearest category should be selected (Alfaro et al. 2016).

$$d(x, c_j) = \min\{d(x, c_1), \ldots\ldots, d(x, c_n)\} \tag{6}$$

Three methods of machine learning including naive Bayes, SVM, and nearest neighbor are used for classification in Duwairi and Qarqaz (2014) to analyze the Arabic sentiments. The results indicate that SVM has a high precision and the nearest neighbor has a high recall. Alfaro et al. (2016) compared the performance of SVM and the nearest neighbor algorithm for text classification and sentiment analysis by using weblog data as dataset. They show that SVM performs better in two cases: first it does classification with higher precision, second: it has higher computational speed and needs lower training time.

Decision tree (DT)
The decision tree (Quinlan 1986) is one of the most famous inductive learning algorithms that its main goal is to approximate the objective functions with discrete values. It is a noise-resistant method. This tree is known as decision tree because it shows the process of decision making to determine an input sample category. The decision tree can be a good option for opinion mining because it has very good performance against the high-volume data. The decision tree algorithms such as CART, C5.0, C4.0, CHAID, QUEST and also SVM have been used to classify the sentiments in Bastı et al. (2015). Since C5.0, CARD and SVM has more accuracy than the others, these three classification algorithms have been used for the final analysis. The final results indicate that, C5.0 has better accuracy than CHAID and SVM. The overall accuracy of C5.0 is close to 97%. CART and SVM has been estimated approximately to 81 and 72% respectively.

Rule-based method
In the rule-based classification approaches, the produced model is a set of rules. Rule is a knowledge structure which relates known information to other information which is derived from them. A rule consists of antecedent and its associated consequent that have an “if-then” relation. The “if” part, includes the conjunction of conditions and the consequent part, includes the prediction class for target. The target could be a product, film, Business Company or a politician. The general form of a classification rules can be seen in relation 7 (Xia et al. 2016).

$$\{w_1 \wedge w_2 \wedge \ldots \wedge w_n\} \rightarrow \{+|-\} \tag{7}$$

Each word in the preceding rule could be a sentiment showing an antecedent like relation 8.

$$\{\text{good}\} \rightarrow \{+\} \quad \{\text{ugly}\} \rightarrow \{-\} \tag{8}$$

To extract rules from a set of training data, algorithms and methods of generating decision-tree, sequential covering algorithm (SCA), or detecting dependency rules are usually used.

In Gao et al. (2015), in order to explore the emotions in a Chinese micro-blog, a rule-based method has been used. According to the proposed method in this study, a rule-based system underlying the conditions which trigger emotions based on an emotion model and extracts the corresponding cause components in fine-grained emotion. The emotional lexicon is built automatically and manually from the corpus while the proportion of cause compounds are calculated by the influence of a multi-language feature based on Bayesian probability. Their results show that the precision of this method could reach to 82.5%. Also in Wen and Wan (2014), emotion classification issue is addressed in micro-blog data sources. The aim of this study is to classify the text in seven types of emotions including anger, hatred, fear, happiness, love, sadness and surprise. The suggested approach, for each sentence in a micro-blog text, firstly obtains two potential emotion labels by using an emotion lexicon and a machine learning approach using Rule Sequence Class and SVM, and considers each micro-blog text as a data sequence. Then, class sequential rules are mined from the dataset and finally new features are obtained from mined rules for emotion classification. Xia et al. (2016) addressed polarity shift problem in opinion mining at document level. Polarity shift is an important subject and affects machine learning methods performance for opinion classification. According to this, they suggested a three-stage model for detecting polarity shifts and eliminations in which they used rule-based methods and statistical methods in order to detect some polarity shifts, as well as a new algorithm for eliminating polarity shift in negation. Results show that their three-stage model was effective in polarity shift detection problem in opinion classification at document level.

Table 3, summarizes some of the most important recent researches in opinion mining, which are based on the supervised learning. The comparison is made in terms of method type, level, applications, datasets and document language.

We also compare the supervised methods in Table 4. It should be noted that, the value of computation cost and classification error which is from low to high, are qualitative measures and are obtained by studying various articles. Supervised techniques have very high accuracy but in contrast, because of their dependency to labeled training samples they have relatively slow efficiency and high cost. The conducted observations and studies show that the SVM has much attracted the researchers' attention due to the numerous advantages in comparison with other methods, so the use of this method can be seen in the most previous works.

5.1.2 Semi-supervised learning

In this section we focus on semi-supervised sentiment classification (Sindhwani and Melville 2008). Traditional classifiers use only labeled data (feature/label pairs) to train. Labeled samples however are often difficult, expensive, or time consuming to obtain, as they require the efforts of experienced human annotators. In addition, in analyzing the comments, labeling a scenario requires strong domain knowledge. Meanwhile, although unlabeled data may be abundant and easily available on the web and so relatively easy to collect, but there are few ways to use them. Semi-supervised learning addresses this problem by using large amount of

Table 3 Summary of some of the recent articles in supervised learning-based opinion classification

References	Year	Level	Technique used	Application	Dataset	Language
Saleh et al. (2011)	2011	Document level	SVM	Product and service (movies, books, cars, cook wave, phones, hotels, music, camera, computer)	IMDB, epinions, Amazon	English
Liu et al. (2012)	2012	Phrase level	Decision tree, SVM	Restaurant review	Online restaurant evaluation website	English
Alfaro et al. (2016)	2013	Phrase level	SVM, KNN	Electoral campaigns	Personal weblog	Spanish
Ren and Kang (2013)	2013	Document level	Bayesian networks	Predicting the delicate human emotions	weblog	Chinese
Moraes et al. (2013)	2013	Document level	SVM, neural network	Film, product (GPS, BOOK, Camera)	Amazon.com Benchmark dataset	English
Duwairi and Qarqaz (2014)	2014	–	Naïve Bayes, SVM, KNN	Education, sports, political news	Facebook	Arabic
Anjaria and Guddeti (2014)	2014	–	Naïve Bayes, Maximum entropy, Neural network	Us presidential election 2012 and Karnataka state assembly elections (India) 2013	Twitter	English
Habernal et al. (2015)	2014	Document level	Maximum entropy, SVM	Product, movie	Social media	Czech

Table 3 continued

References	Year	Level	Technique used	Application	Dataset	Language
Tang et al. (2015)	2015	Document level	Neural network	Review rating prediction (film and restaurant)	Two benchmark database (rotten tomatoes and yelp)	English
Severyn et al. (2016)	2015	–	SVM	Products (apple ipad, Motorola xoom, fiat500, etc)	YouTube	English and Italian
Jeyapriya and Selvi (2015)	2015	Sentence level	Naïve Bayes	Product review	Amazon, Epoinions, Cnet	English
Gao et al. (2015)	2015	–	Rule-based	Explore the emotion causes effectively	weibo	Chinese
Tripathy et al. (2016)	2016	Document level	SVM, Naïve Bayes, Maximum entropy, Stochastical gradient	Movie review	IMDB	English
Vilares et al. (2017)	2017	Sentence level	Sentiment classifier	Different topics	Twitter	English, Spanish
Pham and Le (2017)	2017	Aspect Level	Neural network	Hotel review	TripAdvisor	English

Table 4 Comparing supervised learning methods in opinion classification

Supervised learning	Advantages	Drawbacks	Assessment
Probabilistic classification			
Naive Bayesian	Very useful for extracting subjective sentence	Difficult implementation	It is an effective technique despite the need for primary knowledge
	Easy interpretatio	Assuming the features are independence	
	Requires low volume of training to start the work		
Bayesian network	Easy to understand in complex domains	Able to cope with a limited number of continuous variables	Achieve good accuracy even with little training samples
	Spend little time and effort to construct the model		Resistant against missing data
Maximum entropy	No restriction to feature space dimensions	Over-fitting Problem	Very desirable performance by increase in the features space
	Be able to combine several sources of knowledge and adding additional knowledge easily		An appropriate method to classify the heterogeneous data on the web
Non-probabilistic classification			
Support vector machine	Relatively easy training	Need to choose an appropriate Kernel function	Very good performance in the experimental result
	A good generalization in theory and practice	Deceleration by increase in the samples	Having the most advantages among the other methods
	Low dependency to the dimensionality of feature space	Problem of interpretation	

Table 4 continued

Supervised learning	Advantages	Drawbacks	Assessment
Neural network	Good performance against noise in data	Difficult implementation and interpretation	It requires more training time compared with other techniques
	Quick execution time	High memory usage	Convolutional neural networks are appropriate alternatives to overcome high computational cost of neural network
K-nearest neighbor	Speed of training time	Sensitive to the type of measurement distance	The classification speed is relatively slow if there is a large training set
	Relatively easy to implement		
Decision tree	Resistant to data noise	It could not be applied in small datasets	Very good performance against large datasets
	Easy to understand and interpret		
Rule-based	Overlap in the decision-making space	Poor performance against noisy data	Limited efficiency in text processing at sentence level
	Easy to understand		

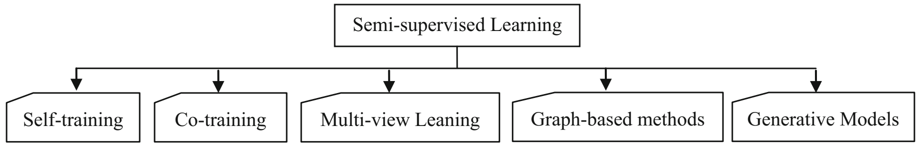

Fig. 7 Classification of semi-supervised learning methods in opinion classification

unlabeled data, together with small amount of labeled data, to build better classifiers. semi-supervised learning approaches requires less human effort and gives higher accuracy, so are of great interest in opinion mining area. Semi-supervised learning methods which proposed in the literature have been categorized in Fig. 7.

5.1.2.1. Self-training This approach is considered as one of the famous and popular methods among the semi-supervised learning methods, which has been used abundantly (Zimmermann et al. 2014; Gao et al. 2014; He and Zhou 2011). In the self-training methods, at first, the classifier should be trained using a small number of labeled training samples, then the trained classifier will be used to classify test (unlabeled) samples. The test samples which have been labeled with the maximum reliability will be added to the set of training samples. The classifier will be trained with the set of all (including new) training samples again and the process will be repeated. In other words the classifier uses its prediction for training. Self-training, generates a model through sentiment labeled data, and uses this model to predict data without sentiment labels. In the result of prediction, data without sentiment labels with high confidence of sentiment level are selected and added to data with sentiment labels with attached sentiment label (Hong et al. 2014). This approach modifies the model, based on its output; so if the model generates wrong output, the model can be modified wrongly. In the other words, the wrong results propagated to next generated models. To alleviate the weak point, Hong et al. (2014) proposed a competitive self-training technique. They create three models based on the output of the model and choose the best. Finally they could improve the performance of sentiment analysis model. Da Silva et al. (2016) proposed a semi-supervised framework to analyze opinions in twitter. Moreover, they used self-training method for a better tweet classification. Their experiment results in real dataset show that proposed framework causes accuracy enhancement in twitter opinion analysis.

5.1.2.2. Co-training Co-training is a semi-supervised learning method, which uses both labeled and unlabelled data. It is assumed that, each example can be explained by two sets of different features which have different and complementary information about every sample. In co-training, two classifiers will be trained separately and the information obtained from this training is shared between each other. Then any of these two classifiers retrain by using training samples which has recently been added by the other classifier. In this way a large set of training data will be formed. This method has been first implemented by Blum et al. (1998). They have used a small set of labeled data and a large set of unlabeled data for the repetition operation to build a more complete classification model. Co-training has been used in a few sentiment analysis tasks. Blum and Mitchell's algorithm is used in Wan (2011) to do sentiment classification on reviews, using Chinese data from one view and English data from another view. Also sentiment classification of tweets which are offered in Liu et al. (2013a, b) used co-training approach.

5.1.2.3. Multi view learning In this method, the main assumption is that the collection of hypotheses is compatible together. In this method, the final goal is to produce k models based on k views. By utilizing this agreement, a different representation of the problem is used to improve the overall performance of classification. One of the applications of multi view learning is in cross-lingual sentiment analysis. Cross-lingual sentiment analysis is considered critical for classifying the sentiments and it has been widely investigated in recent years (Hajmohammadi et al. 2014). The objective of cross-lingual study is to use the labeled data of the source language (mostly English) to compensate the lack of labeled data in the target language. Multi view learning is a very effective method in such matters because it has several views. In fact, different views and different source languages can complement each other to cover the sentimental terms of test data. For example, Mihalcea et al. (2007) generates subjectivity analysis resources in a new language from English sentiment resources by using a bilingual dictionary. Hajmohammadi et al. (2014) is one of the research samples proving the efficiency of the mentioned method in the context of cross-lingual sentiment analysis. In the above-mentioned research, the semi-supervised multi view learning method has been used to classify the sentiments by using data collection of book review in 4 different languages.

5.1.2.4. Graph-based methods Graph-based learning has been focused by researchers in the last decade (Wang et al. 2015b; Lu 2015) and has been confirmed to be effective for many NLP tasks (Xing et al. 2018). Graph-based models have been used for sentiment classification (Lu 2015), automatic creating of sentiment lexicons (Hassan and Radev 2010), cross-lingual sentiment analysis (Hajmohammadi et al. 2015), and social media analysis (Speriosu et al. 2011). Graph-based learning represents data as a weighted graph in which vertices represent instances and edges reflect instances similarities. The assumption is that tightly connected instances are likely belonging to same class. In sentiment classification, instances are documents. Document similarity is used to show the adjacency of document sentiments. Three vital questions must be answered when graph-based learning applies to sentiment classification (Ponomareva 2014):

1. How a sentiment graph is constructed? The answer to this question is a key point for having successful performance.
2. Another question is raised according to the similarity measure between documents which state the adjacency of document sentiments rather than content. Sentiment classification requires a similarity metric that assigns values to a pair of documents on the basis of their sentiment in a way that documents which have same sentiment, have high level of similarity and vice versa.
3. What kind of algorithm should we use in this method? Graph Mincut (Blum and Chawla 2001), label propagation (LP) (Khan et al. 2014), spectral graph transducer (Joachims 2003), manifold regularization (Belkin et al. 2006), modified adsorption (Talukdar and Crammer 2009) and measure propagation (Subramanya and Bilmes 2011) are the most popular algorithms that have been used in this area.

Wang et al. (2015b) applied a graph-based semi-supervised method for classifying twits into six different classes. Its experimental results show desirable performance of graph method. Lu (2015), proposed a model based on social relations and text similarity has been suggested in micro-blog sentiment analysis. It used micro-blog/micro-blog relations to build a graph-based semi-supervised classifier. Experiments on two real-world datasets show that Its graph-based semi-supervised model outperforms the existing state-of-the-art models.

5.1.2.5. Generative models Generative model defines a distribution on the inputs and the training is conducted for each class. Then the Bayes rule can be used to predict the belonging of the sample to the class at the time of testing phase. According to the dataset, we consider

$\{x^{(i)}, y^{(i)}\}$ the pair where $x^{(i)}$ is the ith document of the training set, $y^{(i)} \in \{-1, +1\}$ and N is the number of training samples. There will be two training models (Mesnil et al. 2015):

$p^{+}(x|y{=}{+}1)$ Means $\{x^{(i)}\text{with condition } y^{(i)} = +1\}$
$p^{-}(x|y{=}{+}1)$ Means $\{x^{(i)}\text{with condition } y^{(i)} = -1\}$

According to Bayes rule at the test time, for input (x) we have relation 9 (Mesnil et al. 2015):

$$R = \frac{p^{+}(xly = +1)}{p^{-}(xly = -1)} \times \frac{p(y = 1)}{p(y = -1)} \tag{9}$$

In relation 9, if $R > 1$, x belongs to the positive class, otherwise it will belong to the negative class. In Mesnil et al. (2015) authors proposed a powerful and at the same time very simple ensemble system for sentiment analysis. Three conceptually different baseline models are combined in this system: one based on a generative approach (language models), one based on sentences' continuous representation, and one based on a clever TF-IDF reweighing of document's bag-of-word representation. Each of these three models contributes to the success of overall system, lead to achieving high performance on IMDB[9] movie review dataset.

Semi-supervised techniques to classify the sentiments have been summarized in Tables 5 and 6. Table 5 contains the latest researches conducted in this field and we discuss in Table 6, the total assessment of these methods and the expression of their advantages as well as their drawbacks.

5.1.3 Unsupervised learning (clustering)

In unsupervised learning methods, a set of training samples is considered which only the input value is specified for them and the accurate information about the output is not available. The clustering-based approaches are able to produce moderately accurate analysis results without any human participation, linguist knowledge or training time (Li and Liu 2014). Clustering is considered as an unsupervised method which aims to find a structure in data collections that have not been classified yet and no pre-defined classes already exists. In the other words, clustering is putting data in different groups in which, the members of each group are similar to each other from a particular point of view. Thus, the data of a cluster have the maximum similarity and the data of different clusters have minimum similarity. The similarity criterion is the distance i.e. the samples being closer to each other are placed in the same cluster. For example, in documents clustering, the closeness of two sample documents can be determined based on the number of common words of the two documents.

The application of clustering on sentiments analysis was introduced in Li and Liu (2012), where feature extraction was done by calculating the TF-IDF criterion. Based on this criterion, the terms are chosen as feature which has more TF-IDF value using relation 10 (Li and Liu 2012). In this relation, TF is proportional to the frequency of a term in the document and IDF is considered as a weighting factor which inversely represents the dispersion amount of a term in different documents.

$$\mathrm{TF-IDF} = \mathrm{tf}_i^{*}\mathrm{idf}_i = \frac{n_i}{\sum_i n_i} * log\frac{M}{df_i} \tag{10}$$

where n is the number of repetition of word i in the text, m is the total number of sentences and *idf* is the total number of sentences including the word i.

[9] www.IMDB.com.

Table 5 Summary of some of the recent articles in semi-supervised learning based opinion classification

References	Year	Level	Technique used	Application	Data base	Language
Li et al. (2012)	2012	–	Graph-based, multi-view learning	Product review (cell phone)	Forums	Chinese
Hong et al. (2014)	2014	–	Self-training	Analyze users emotion	Twitter	English
Zimmermann et al. (2014)	2014		Self-training	Product review and movies	Twitter	English
Hajmohammadi et al. (2014)	2014	–	Multi-view learning	Book review	Amazon	English, French, German, Japanese
Wang et al. (2015b)	2015	–	Graph-based	Intent tweets into six categories namely (food and drink, travel, career and education, good and services, event and activites and trifle)	Twitter	English
Carter and Inkpen (2015)	2015	–	Co-training	Product (lap taps) and services (hotel)	SemEval-2014	English
Wang et al. (2015c)	2015	–	Co-training	Predict the user gender	Weibo	Chinese

Table 5 continued

References	Year	Level	Technique used	Application	Data base	Language
Claypo and Jaiyen (2015)	2015	–	Graph-based	Health care reform and politics (American presidential debate on September 26, 2008)	Twitter	English
Mesnil et al. (2015)	2015	–	Generative models	Film review	IMDB	English
Khan et al. (2016)	2016	–	Graph-based	Film review	Cornell movie review	English
Iosifidis and Ntutsi (2017)	2017	–	Self-training, Co-training	Stream data	TSentiment15	English
Rout et al. (2017)	2017	–	Co-training	Hotel review	gold standard	English

Table 6 Comparing semi-supervised learning methods in opinion classification

Semi-supervised learning	Advantages	Disadvantages	Assessment
Self-training	Simplicity of the method	If there is an error in the input sample there is a possibility to strengthen it	Traditional self-training method has poor performance
	No dependence to classification model	Sensitive to outlier	
Co-training	It reaches to high accuracy in classification with very limited number of labeled data	Poor performance in the datasets which only have a unique view point	Very sensitive to data
		Many features must be available to achieve an optimal performance	Different accuracy in simple and complex domains
Graph-based	No need for deep linguistic analysis and manual attempts	Sensitive to noise	Good performance in many natural language processing problems
	Be able to cope with binary and multi-class classifications	Its performance is sensitive to graph structure	It is an alternative and competitive method against other semi-supervised methods
			Same performance in simple and complex domains
Multi-view learning	Efficient in cross-lingual issues and various linguistic resources	A failure will arise if you choose inappropriate language with low vocabulary as a source language	Using several different viewpoints rather than one viewpoint, leads to good performance
Generative models	When the set of labeled training samples are too low, it can reach to high accuracy	Low efficiency	In the most cases do not perform well
		Inflexibility	

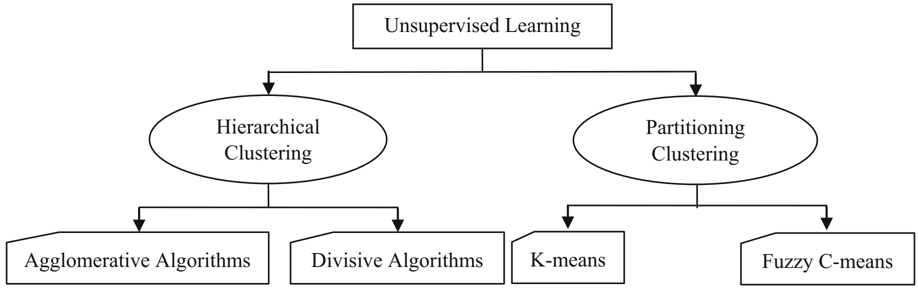

Fig. 8 Classification of clustering algorithm in opinion classification

Since then, many algorithms have been proposed for data clustering in the opinion mining area. As is shown in Fig. 8, these algorithms can be categorized into two main classes (Li and Liu 2014): *partition clustering* and *hierarchical clustering*.

5.1.3.1. Partition clustering In the partitioning clustering approach, we deal with a collection of clusters that do not have overlap and any object will be assigned only to one cluster. The objective of partition clustering is to divide data in such a way that data within a cluster have the most similarity and on the other hand have the most distance with the data of other clusters (Li and Liu 2014). The similarity criterion is the Euclidean distance and eventually it can create a super spherical environment. Sensitivity to the noise is the main disadvantage of this method.

K-means

One of typical algorithms in clustering is K-mean. In this algorithm, firstly K centers are defined randomly for each of K clusters. In next step, each of the input data set is linked to the closest center. After the end of this step, new centers are calculated again for batches obtained from previous step. In next step, a connection is established between each set data and the nearest obtained center (Claypo and Jaiyen 2015). Hence, the value of K is changed in each repeating round of this loop. This algorithm will be finished when no change occur in the value of K. The aim of running this algorithm is to minimize the target function. Target function is defined as relation 11 (Claypo and Jaiyen 2015):

$$\mathrm{J} = \sum_{i=1}^{n} \sum_{j=1}^{k} x_i - c_j^2 \tag{11}$$

According to relation 11, $\|x_j - c_j\|^2$ calculates the distance of x_i to cluster center c_j, usually using Euclidean distance of relation 12 (Claypo and Jaiyen 2015).

$$\mathrm{d(x, y)} = \sqrt{\sum_{i=1}^{n} (x_i - y_j)} \tag{12}$$

where d denotes distance and n is the number of data.

K-means is an appropriate algorithm for clustering-based sentiment analysis. An example of clustering-based sentiment analysis can be seen in Li and Liu (2014) which have introduced some new methods to expand the existing capabilities of clustering and pursues two following steps: In the first step, the contents will be initially processed in order to identify concrete expressions which will increase the accuracy. In the next step sentiment analysis will be conducted through voting mechanism and distance measuring method. Its experimental results show that the sentiment analysis using clustering has acceptable quality and is a very good option to recognize the neutral viewpoints. In Claypo and Jaiyen (2015), authors have proposed an opinion mining on Thai restaurants review using K-means clustering and MRF feature selection technique to select the relevant features. It reduced the number of features and the computational times. Then, K-means is adapted for clustering the Thai restaurants review into two positive and negative groups. The paper concluded that K-means clustering is compatible with MRF features selection since it can achieve the best performance in the clustering. Although this method is appropriate for big datasets but it could terminates to a local optimum and its usage is limited to situations where we can define the centroids for data. Another weakness of this algorithm is that it is highly sensitive against the noises and also the choice of initial centroids (Li and Liu 2014).

Fuzzy C-means

The number of clusters (C) is previously known in this algorithm. The target function for this algorithm is as relation 13 (Acampora and Cosma 2015).

$$\mathrm{J} = \sum_{i=1}^{d} \sum_{j=1}^{c} \mu_{ij}^{m} \| x_i - v_j^2 \| \tag{13}$$

where m > 1 is a real number which is considered 2 by default. x_i is the ith sample and v_j is center of cluster j. μ_{ij} shows the measure of dependency of the ith sample in the jth cluster. $|| * ||$ indicates sample distance to the center of the cluster in which, Euclidian distance is used. Matrix μ can be defined from μ_{ij} which is a value between 0 and 1. The order of the membership and the centers of clusters can be calculated from relations 14 and 15 (Acampora and Cosma 2015):

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (x_i - v_j^2 / x_i - v_k^2)^{2 \big/ m-1}} \tag{14}$$

$$v_j = \frac{\sum_{i=1}^{c} u_{ik}^{m} x_i}{\sum_{i=1}^{d} u_{ij}^{m}} \tag{15}$$

Zimmermann et al. (2016) presented a framework to extract implicit features from given comments about products and to identify their polarity. Their proposed method has a mechanism which merges clusters having common features. Fuzzy C-means method is also used in this paper. Acampora and Cosma (2015) performed a comparison of fuzzy methods performance in predicting rating of products on 6 datasets. They used an effective computational intelligence framework based on genetic algorithm to minimize data dimensions. Their research results show that the proposed framework is useful for predicting rating.

5.1.3.2. Hierarchical clustering Hierarchical clustering is actually a set of nested clusters which organized as a tree (Li and Liu 2014). In this method, clusters or groups are allowed to have sub-groups. Hierarchical clustering methods can be divided into two main categories based on their structure.

Divisive methods

In divisive (also called top-down) method, first, all data are considered as one cluster and then those data which have less similarity to each other will be broken to separate clusters through an iterative process until the clusters become single-member.

Bottom-up or agglomerative
In agglomerative (also called bottom-up) method, first, each data is considered as a separate cluster and, in any step, the data that are more similar to each other are combined until we eventually achieve to one or a specified number of clusters.

An effective method for unsupervised clustering for collecting opinions from idea portals is used in De and Kopparapu (2013). Their clustering technique is based on agglomerative algorithm. In Shi and Chang (2008), the authors proposed a method to automatically construct a product hierarchical concept model based on the online review of the products. Their three-phase method can be summarized as identifying nouns from the review dataset, creating a contextual representation vector for every nouns, and top-down noun clustering in order to create a concept hierarchy. The results of the experiments conducted on three sets of online reviews (in China and English) illustrated the qualitative and quantitative effectiveness and robustness of this approach. The summary of unsupervised techniques to classify the sentiments is represented in Tables 7 and 8.

5.2 Lexicon-based approach

The words containing sentiment are used to express the positive or negative feelings. For example, words such as "good", "beautiful" and "amazing" induce positive feeling to human and words like "bad", "ugly" and "scary" are the words with negative polarity (Liu 2015, 2012). By word's polarity, we mean the feeling and assessment that the word brings in the mind. It should be noted that, the words that bear the sentiment are mostly adjectives and adverbs but some names like "trash" and "garbage" and verbs like "hate" and "love" have sentiment as well (Liu 2012). In the lexicon of opinion mining, each word that contains the sentiment has come along with the polarity of that word. This lexicon may be weighted or un-weighted. Generally, weight is a number or a probability that is considered for a word to show the level of positivity or negativity. In fact, lexicon based methods are used to calculate the orientation of documents according to the semantic orientation of words and phrases within documents. This method focuses on individual words and doesn't consider the words having comparative sense such as: "the best", "worse", etc (Liu 2012). In this method, a lexicon of sentiments can be found by using the text analysis. Also in some cases, natural language processing is used to find the syntactic structure and help to find the semantic relationships (Moreo et al. 2012; Caro and Grella 2013). The reviews classification using semantic orientation applied to unsupervised classification in Turney (2002) for the first time that was based on syntactic patterns and POS tags. This method has three main steps:

1. Extract phrases containing adjectives or adverbs,
2. Estimate the semantic orientation of each phrase (which uses PMI-IR to calculate semantic orientation), the point-wise mutual information formula is defined by relation 16 (Cover and Thomas 2012).

$$M_i(w) = \log\left(\frac{f(w).p_i(w)}{f(w).p_i}\right) = \log\left(\frac{p_{i(w)}}{p_i}\right) \tag{16}$$

   $M_i(w)$ has been defined between the word w and the class i. w has a positive relationship to class i, if M_i (w) is greater than zero and has a negative relationship, if M_i (w) is smaller than zero. The expected co-occurrence of class I and word w, on the basis of mutual independence, is given by $(w) \cdot p_i$, and the true co-occurrence is given by $f(w) \cdot p_i(w)$.
3. Classify the review based on the average semantic orientation of the phrases.

Lexicon based methods can be classified into two main groups (Liu 2012) as shown in Fig. 9. Researchers have used different methods to compile the sentiment words by creating

Table 7 Summary of some of the recent articles on unsupervised learning based opinion classification

References	Year	Level	Technique used	Application	Data base	Language
Shi and Chang (2008)	2008	–	Divisive algorithm	Product (cell phone,mp3 player) and services (hotel)	Ctrip, Amazon	English, Chinese
Tsagkalidou et al. (2011)	2011	–	K-means	People emotion laden reactions and attitude	Twitter	English
Li and Liu (2012)	2012	Document level	K-means	Film review	IMDB	English
Vakali and Kafetsios (2012)	2012	–	Divisive algorithm	Effect-aware community detection (cultural, social, economical, political events)	Twitter	English
De and Kopparapu (2013)	2013	Document level	Agglomerative algorithm	Company ideas portal site	Company ideas portal site	English
Archambault et al. (2013)	2013	Document level	Agglomerative algorithm	Twitter (us cities and election 2012 dataset)	Twitter (us cities and election 2012 dataset)	English
Claypo and Jaiyen (2015)	2015	Document level	K-means	Restaurant review	Tripadvisor	English

Table 7 continued

References	Year	Level	Technique used	Application	Data base	Language
Gupta et al. (2015)	2015	Document level	K-means	Users mood swing analyzer	Facebook	English
Gupta et al. (2015)	2015	Document level	K-means	Users mood swing analyzer	Facebook	English
Phu et al. (2017)	2017	Document level	Fuzzy C-means	Movie reviews	Facebook	English
Huang et al. (2017)	2017	Document level	Latent Dirichlet allocation (LDA)	Different topic	Weibo	English, Chinese
Garcia-Pablos et al. (2017)	2017	Aspect level	Latent Dirichlet allocation (LDA)	Hotels, restaurants, electronic devices	SemEval-2016	English, Spanish, French and Dutch
Pandey et al. (2017)	2017	Document level	K-means cuckoo search method (CSK)	Different topic	Twitter	English

Table 8 Comparing unsupervised learning (clustering) methods in opinion classification

Unsupervised learning	Advantages	Disadvantages	Assessment
Partitioning clustering			
K-means	These algorithms perform complexity in a linear time to make them proper for large dataset	It doesn't have enough accuracy if there are ambiguities	Low-cost, efficient and very convenient method to analyze the sentiments
	The algorithm does not need to pre-know the class of a document	The problem of these algorithms is their sensitivity to the initial center points and the assumption of knowing number of clusters	The result is predictably poor in terms of accuracy and stability
	Does not need a training process	These algorithms cannot handle outliers and noise well and cannot perform on non-convex clusters	
	This means that it is free from human participation	Clustering results are unstable due to the random selection of centroids in k-means	
	Low amount of memory required		
Fuzzy c-means	It always converges	Computation time is high	FCM is slower than k-means
		It is sensitive to primary guesses and may stop at local minimums	
		It is noise-sensitive	
Hierarchical clustering			
Agglomerative algorithm	Do not dependent to the number of clusters and center of gravity	The time complexity in hierarchical algorithms is high, o(n^2) for single-link algorithms, and (n^2 logn) for complete-link algorithms	Its high cost is a limitation in large scale dataset applications although their good cluster qualities
	Good performance against the noise		Has more challenges than divisive method
Divisive algorithm	Good performance against the noise	Require large amount of memory	Even though getting efficient accumulating qualities, the expenditure to be allocated to this project can be considered a restricting factor for its large scale application
	Do not relying on the number of clusters and center of gravity	Nonlinear time complexity	

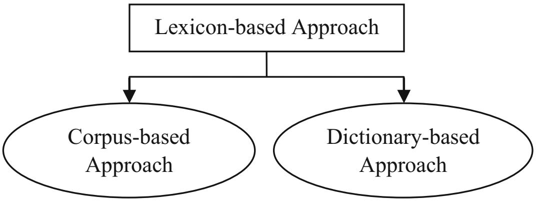

Fig. 9 Classification of lexicon-based approaches in opinion classification

lexicon. Lexicon can be created manually, which is a very difficult and time consuming task, or can be created automatically, that is a few words are used as seed to extend the lexicon lists. Dictionary and corpus based methods include the automatic methods for creating the lexicon which we will describe them in this section (Liu 2012).

5.2.1 Dictionary-based approach

The use of dictionary method to compile the sentiments is common. Since lexicons basically contain lists of synonyms. This method is, in fact, based on the bootstrapping technique. The procedure in this method is as follows: First, in order to create a small set, a few sentiment words will be identified manually which have positive or negative semantic orientation. Then the algorithm helps to grow this collection by searching in WORD NET[10] and other online dictionaries to find synonyms and antonyms (Liu 2012). This process continues until no new words can be detected. In Qiu et al. (2011) and Xia et al. (2015) some sample of bootstrapping methods can be seen. Qiu et al. (2011) suggested supportive results by bootstrapping propagation and a few aspects as seeds. However this method has three main issues:

1. In bootstrapping, aspects are used directly as seeds. Hence, specific aspects (e.g. motion image quality) restrict the propagation ability. Accordingly, in this approach, manpower should be assigned carefully for choosing general aspects (e.g. quality). In fact, general aspects can be extracted by a generalization treatment on specific aspects.
2. Propagation is applied on the test data directly. However the test data are often very small which leads to further reduction in propagation ability.
3. Although compiling dependency rules is difficult and laborious, these rules are very important for the process of propagation. There are interesting patterns for combining general aspects and a strict condition could be assigned for controlling these patterns' quality.

The lexicon method has been used in order to analyze Twitter sentiments in Pandarachalil et al. (2015). Twitter sentiment analysis consists of two steps, in pre-processing step, the polarity of sentiments is determined and then sentiment analysis is performed. The tweets' polarity is estimated by three sentiment lexicons as SENTIWORDNET, SENTISLANGNET, and SENTICNET.

- **SENTICNET** It is a public word source made by sentic computing. The polarity score for each concept is calculated in $[1, -1]$ interval.
- **SENTIWORDNET** It is a common word source used to analyze the sentiments and it is derived from WORD NET and the polarity score is calculated in $[0, 1]$ interval. There are around 117,569 available sentisynsets which are all unigrams, so, SENTIWORDNET

[10] http://wordnet.princeton.edu/.

provide the sentiment score only in syntactic level. This method fails in providing the sentiment scores for phrases like “no” and “very good”.

- **SENTISLANGNET** This lexicon derived from SENTIWORDNET and SENTICNET, which created for slangs. Since tweets are rich in slangs, abbreviations and emoticons, SENTISLANGNET can be so useful for sentiment analysis in Twitter.

Saif et al. (2016) introduced a lexicon-based approach called Senticircle for analyzing Twitter opinions, which has desirable performance in determining particular background opinion orientation.

5.2.2 Corpus-based approach

As is well-known in the sentiment analysis area, the semantic orientation of a word is domain-dependent. The compilation of a set of polar words, the most suitable for obtaining domain-dependent opinion word is that known as the corpus-based approach (Molina-González et al. 2015). This method relies on the syntactic rules. The main idea of this method has been proposed in Hatzivassiloglou and McKeown (1997), in which, a list of adjectives are made containing the sentiments and then new adjectives along with their orientation are determined using some other linguistic constraints. A few rules are also applied on the connective terms, such as: ‘and’, ‘but’, ‘or’ etc. called sentiment consistency. For example, one of these rules is on the conjunction word “and” in this way that, joining the adjectives together do not change in the initial orientation. Consider the sentence “this film is good and attractive”, the conjunction word “and” has caused two adjectives of “good” and “attractive” join to each other which according to this rule when the word “good” is known as a positive, “attractive” is also considered positive (Liu 2012). According to the lack of linguistic sources except for English which is one of important issues in opinion mining field, developing lexicons are critical for classification in various languages. Molina-González et al. (2015) addressed developing opinion lexicon and incorporating domain information for Spanish language classification systems. Their results indicate the integrity of new lexicon. Liao et al. (2016) proposed a new approach which incorporates domain lexicon with groups of feature using syntax and semantic. Their experiment results in real dataset showed that their proposed approach outperformed other baselines of opinion target extraction. The corpus-based approach helps to solve the problem of finding sentiment words with specific orientation context, but its performance varies in different domains, because a word may be positive in one domain and negative in another domain. It has been proved that the dictionary based method is more useful because it is difficult anyway to prepare large entries that can cover all English words (Liu and Zhang 2012).

Table 9 shows a number of recent studies concentrated on this area.

5.3 Comparing techniques of sentiment classification

The comparison between two main techniques of opinion mining is expressed in Table 10. In this table, two machine learning and lexicon technique have been assessed in terms of efficiency, accuracy, strengths and weaknesses. Although the accuracy rate of the supervised learning approach is good, their cost is high for several reasons and also slow efficiency (slow on training and fast on testing). On the other hand, Lexicon methods, do not demand human involvement, their accuracy is limited and have very fast efficiency. The accuracy rate of clustering-based approaches is moderate and they have fast efficiency. Figure 10 also shows an overall view of several classification techniques which have been used in opinion mining and sentiment analysis.

Table 9 Summary of some of the recent articles on lexicon approach based opinion classification

References	Year	Level	Technique used	Application	Data base	language
Taboada et al. (2011)	2011	Sentence level	Dictionary and corpus based	Product and services, film	Epinions, Rotten tomatoes, IMDB	English
Molina-González et al. (2014)	2014	–	Corpus based	Hotel reviews	Tripadvisor	Spanish
Sharma et al. (2014)	2014	Document level	Dictionary based	Film reviews	Rotten tomatoes, IMDB	English
Rao et al. (2014)	2014	Word level	Dictionary based	Identifying social emotion on certain entities and news events	News sites (sine society) and SemEval	English Chinese
Feng et al. (2015)	2014	–	Dictionary based	Identifying social emotion	Weibo	Chinese
Molina-González et al. (2015)	2014	–	Corpus based	Product, hotel, movie	Epinions	spanish
Jiménez-Zafra et al. (2015)	2015	Aspect level	Dictionary and corpus based	Product (laptop) and service (restaurant)	SemEval 2014	English and Spanish
Chinsha and Joseph (2015)	2015	Aspect level	Dictionary based	Restaurant review	TripAdvisor	English
Rathan et al. (2017)	2017	Aspect level	Dictionary based	Mobile phone reviews	Twitter	English
Zhou et al. (2017)	2017	–	Corpus based	Product (tablet)	Kindle Fire HD tablets	English
Chao and Yang (2018)	2018	Aspect level	Corpus based	Restaurant reviews	IPEEN, TripAdvisor	Chinese

Table 10 Comparing machine learning techniques and lexicon approach in opinion classification

Sentiment classification techniques	Efficiency	Accuracy	Strengths	Weaknesses
Machine learning				
Supervised learning	Slow	Very high	Having the ability to analyze numerous categories	Dependence on the labeled training documents
			Effectiveness in the discovery of subjectivity issue	Requires the presence of human effort and linguistic knowledge
			Noise-resistant	High cost
				Require time to train, and for high dimensional data, the process is highly time consuming
				Reply on human participation
Semi-supervised learning	Medium	High	Good performance if there are ambiguities in the review	If there is noise in the unlabeled samples, the classification faces with trouble
			Could help to achieve the highest accuracy using as little human annotation effort as possible	
Unsupervised learning (clustering)	Fast	Medium	It does not require much human participation	Its ability to analyze multiple categories has not still been proven
				Not resistant against the noise
			It needs to be efficient and widely applicable	The number of clusters in most cases is unknown
				The accuracy can sometimes be relatively low
				Instability of results

Table 10 continued

Sentiment classification techniques	Efficiency	Accuracy	Strengths	Weaknesses
Lexicon-based				
Dictionary-based	Very fast	Relatively low	Does not require any labeled training samples	Unable to find the opinion words with the specific content orientation domain which is not in the lexicon
			Easy access to words lexicon and their orientation	Discomfit in the texts with a certain semantic dependency
			Provide better results for less banded domain	less accurate during consideration of different domains
Corpus-based	Very fast	Relatively low	The ability to find the opinion words with the specific content orientation	Variable performance due to the extent domain of lexicon
			Provides better results when domains are different	The difficulty to provide large texts with the ability to cover all the text words
				Can not be used alone

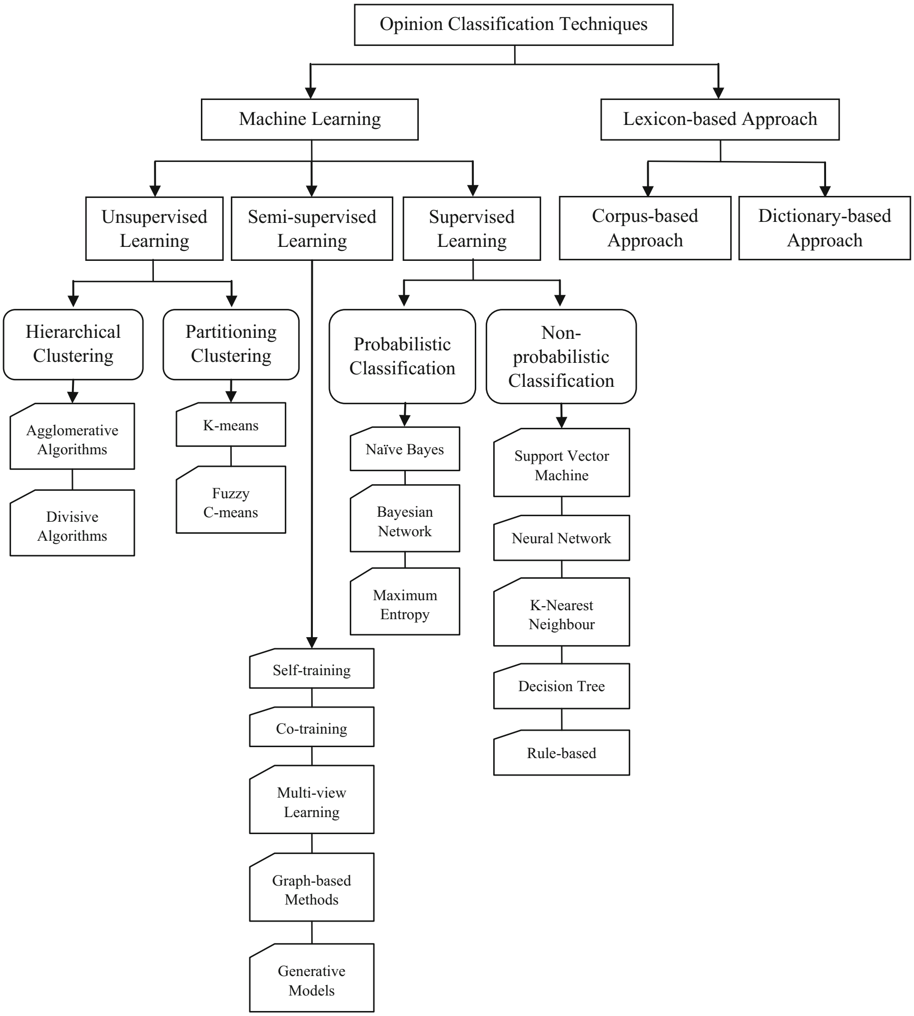

Fig. 10 An overall view of several classification techniques in opinion mining

6 Evaluation criteria

To evaluate the performance of different methods for text classification and sentiment analysis, four criteria are considered from information retrieval (Baeza-Yates and Ribeiro-Neto 1999; Olson and Delen 2008) including accuracy, precision, recall and f1-score which are defined by Eqs. 17–20 (Alfaro et al. 2016). In these criteria T_p, F_n, F_p *and* T_n respectively refer to the number of correct results that have been diagnosed correctly, the number of incorrect results that have been diagnosed incorrectly, the number of incorrect results that have been diagnosed correctly and the number of correct results that have been diagnosed incorrectly. In all the above criteria, the result is ranged between 0 and 1; whatever it is closer to 1, the performed operation has better results.

- **Accuracy criterion** Accuracy is the proportion of appropriate classified results to the entire population.

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{17}$$

- **Precision criterion** Precision is the ratio of the correctness of the results within the system outputs. This criterion indicates that the more the number of correct results and the less the number of incorrect results is, the more is the precision of the performed operation.

$$\text{Precision} = \frac{T_p}{T_{p+} F_p} \tag{18}$$

- **Recall criterion** Recall is the ratio of correct results that the system assigned compared to the base of all results. In fact this criterion indicates that the higher the number of correct results and the less the number of correct results that has been diagnosed incorrectly is, the more is the recall power of performed operation.

$$\text{Recall} = \frac{T_p}{T_{p+} F_n} \tag{19}$$

- **F1-score criterion** Another popular index is F1 whose formula comes from the combination of precision and recall. This criterion indicates the amount of precision and recall power of mentioned operation.

$$\text{F1} = 2 * \frac{precision \cdot recall}{precision + recall} \tag{20}$$

- **Kappa criterion** Kappa is also a well-known criterion which measures the agreement between two classifiers and is calculated using relation 21 (Ahmed and Danti 2016).

$$\text{Kappa} = \frac{P_0 - P_e}{1 - P_e} \tag{21}$$

where P_0 is the relative agreement of classifiers and P_e is the hypothetical probability of chance agreement.

Moreover, there are some of more application-specific criteria which are used for evaluating opinion mining and sentiment analysis. For example, Jiang et al. (2017a) represents a new evaluation metric to investigate the performance of its proposed method. It uses emoticons as the benchmark of its approach as relation 22.

$$e_b^d = \frac{1}{N_e^d} \sum_{i=1}^{N_e^d} e_{e_i} \tag{22}$$

where N_e^d and ei are the number of emoticons and one emoticon of micro-blog d respectively. e_{e_i} is the emotion vector of an emoticon, and e_b^d is the emotion vector of micro-blog d. Area under curve (AUC) is another evaluation metric which has been used in Lv et al. (2017) along with another well-known criteria such as precision, recall, and F1 score. The receiver operating characteristic (ROC) curve has been quantified as AUC in the mentioned paper.

7 Future direction

There are several open issues in opinion mining and sentiment analysis which should be more addressed in the future. Some of the most important open issues in different aspects of sentiment analysis can be listed as follows.

- Opinion mining can help discover and extract useful and profound knowledge resources using the concept level sentiment analysis. The analysis of emotions at the level of the concept, because of its action beyond the other three levels, has recently attracted the attention of researchers. Using different approaches and combining them at the conceptual level can be further investigated in the future (Bisio et al. 2017; Bajpai et al. 2017; Peng et al. 2017; Cambria 2016; Jain and Jain 2017; Qazi et al. 2016).
- The tendency of the convolution neural network method is very high in the new articles. In the future, it can be more focused on identifying neutral comments and improving the performance of the models by using the convolution neural network method on large corpus and domain dependent corpora (Xu et al. 2017; Majumder et al. 2017; Poria et al. 2016, 2017; Wehrmann et al. 2017; Chen et al. 2016; Jiang et al. 2017b).

8 Conclusions

During the last decade we have seen an increasing interest in processing and analyzing unstructured data especially web-based data. Opinion mining is a research subject that has had a significant growth and it has attracted the attention of many researchers from the computer to management sciences in recent years. In this paper, we have briefly examined the opinion mining area and its related classification techniques. The aim of this paper is to develop a proper survey by examining the well-known existing methods and their challenges. Supervised methods have been stated as an appropriate model to classify the comments having high accuracy and validity but because of relying on tagged training documents, they seek a very slow and expensive trend. In addition, studies on semi-supervised techniques are increasing and considering the popularity of micro-blogs such as Twitter, semi-supervised techniques are very good options for micro-blogs. In this paper, it has been pointed out to the popularity of clustering and lexicon-based methods among the researchers that have taken a growing trend. Although the techniques and algorithms used to analyze the comments are rapidly progressing, many of the issues in this field of study still needs further work and remains unsolved.

References

Abdul-Mageed M, Diab M, Kübler S (2014) SAMAR: subjectivity and sentiment analysis for arabic social media. Comput Speech Lang 28(1):20–37

Acampora G, Cosma G (2015) A comparison of fuzzy approaches to E-commerce review rating prediction. In: 2015 conference of the international fuzzy systems association and the European society for fuzzy logic and technology (IFSA-EUSFLAT-15). Atlantis Press

Ahmed S, Danti A (2016) Effective sentimental analysis and opinion mining of web reviews using rule based classifiers. In: Behera H, Mohapatra D (eds) Computational intelligence in data mining—volume 1. Advances in intelligent systems and computing, vol 410. Springer, New Delhi

Alfaro C, Cano-Montero J, Gómez J, Moguerza JM, Ortega F (2016) A multi-stage method for content classification and opinion mining on weblog comments. Ann Oper Res 236(1):197–213

Anjaria M, Guddeti RMR (2014) A novel sentiment analysis of social networks using supervised learning. Soc Netw Anal Min 4(1):1–15

Appel O, Chicalana F, Carter J, Fujita H (2016) A hybrid approach to the sentiment analysis problem at the sentence level. Knowl Based Syst 108:110–124
Arab M, Sohrabi MK (2017) Proposing a new clustering method to detect phishing websites. Turk J Electr Eng Comput Sci. https://doi.org/10.3906/elk-1612-279
Archambault D, Greene D, Cunningham P (2013) Twittercrowds: techniques for exploring topic and sentiment in microblogging data. Preprint. arXiv:1306.3839
Baeza-Yates R, Ribeiro-Neto B (1999) Modern information retrieval, vol 463. ACM Press, New York
Bajpai R, Poria S, Ho D, Cambria E (2017) Developing a concept-level knowledge base for sentiment analysis in Singlish. Preprint. arXiv:1707.04408
Balahur A, Perea-Ortega JM (2015) Sentiment analysis system adaptation for multilingual processing: the case of tweets. Inf Process Manag 51(4):547–556
Balahur A, Hermida JM, Montoyo A (2012) Detecting implicit expressions of emotion in text: a comparative analysis. Decis Support Syst 53(4):742–753
Balazs JA, Velasquez JD (2016) Opinion mining and information fusion: a survey. Inf Fusion 27:95–110
Bastı E, Kuzey C, Delen D (2015) Analyzing initial public offerings' short-term performance using decision trees and SVMs. Decis Support Syst 73:15–27
Belkin M, Niyogi P, Sindhwani V (2006) Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J Mach Learn Res 7:2399–2434
Berger AL, Pietra VJD, Pietra SAD (1996) A maximum entropy approach to natural language processing. Comput Linguist 22(1):39–71
Bilal M, Israr H, Shahid M, Khan A (2016) Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, decision tree and KNN classification techniques. J King Saud Univ Comput Inf Sci 28:330–344
Bing L, Chan KC, Ou C (2014) Public sentiment analysis in Twitter data for prediction of a company's stock price movements. In: IEEE 11th international conference on e-business engineering (ICEBE), pp 232–239
Bisio F, Meda C, Gastaldo P, Zunino R, Cambria E (2017) Concept-level sentiment analysis with SenticNet. In: Cambria E, Das D, Bandyopadhyay S, Feraco A (eds) A practical guide to sentiment analysis. Socio-affective computing, vol 5. Springer, Cham, pp 173–188
Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
Blitzer J, Dredze M, Pereira F (2007) Biographies, bollywood, boom-boxes and blenders: domain adaptation for sentiment classification. In: ACL, vol 7, pp 440–447
Blum A, Chawla S (2001) Learning from labeled and unlabeled data using graph mincuts. In: Brodley CE, Danyluk AP (eds) Proceedings of the eighteenth international conference on machine learning, pp 19–26
Blum A, Mitchell T (1998) Combining labeled and unlabeled data with co-training. In: Proceedings of the eleventh annual conference on computational learning theory. ACM, pp 92–100
Boiy E, Moens MF (2009) A machine learning approach to sentiment analysis in multilingual Web texts. Inf Retr 12(5):526–558
Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. J Comput Sci 2(1):1–8
Bouadjenek MR, Hacid H, Bouzeghoub M (2016) Social networks and information retrieval, how are they converging? A survey, a taxonomy and an analysis of social information retrieval approaches and platforms. Inf Syst 56:1–18
Bravo-Marquez F, Frank E, Pfahringer B (2016) Building a Twitter opinion lexicon from automatically-annotated tweets. Knowl Based Syst 108:65–78
Cai C, Xia B (2015) Convolutional neural networks for multimedia sentiment analysis. In: 4th Springer conference on natural language processing and Chinese computing, pp 159–167
Cambria E (2013) An introduction to concept-level sentiment analysis. In: MICAI 2013: Advances in soft computing and its applications Mexican international conference on artificial intelligence, pp 478–483
Cambria E (2016) Affective computing and sentiment analysis. IEEE Intell Syst 31(2):102–107
Cambria E, Schuller B, Xia Y, Havasi C (2013) New avenues in opinion mining and sentiment analysis. IEEE Intell Syst 28(2):15–21
Carter D, Inkpen D (2015) Inferring aspect-specific opinion structure in product reviews using co-training. In: Gelbukh A (ed) Computational linguistics and intelligent text processing. CICLing 2015. Lecture notes in computer science, vol 9042. Springer, Cham, pp 225–240
Chao AFY, Yang H (2018) Using Chinese radical parts for sentiment analysis and domain-dependent seed set extraction. Comput Speech Lang 47:194–213
Chen LS, Liu CH, Chiu HJ (2011) A neural network based approach for sentiment classification in the blogosphere. J Informetr 5(2):313–322
Chen L, Wang F, Qi L, Liang F (2014) Experiment on sentiment embedded comparison interface. Knowl Based Syst 64:44–58

Chen T, Xu R, He Y, Xia Y, Wang X (2016) Learning user and product distributed representations using a sequence model for sentiment analysis. IEEE Comput Intell Mag 11(3):34–44

Chinsha TC, Joseph S (2015) A syntactic approach for aspect based opinion mining. In: IEEE international conference on semantic computing (ICSC), pp 24–31

Claypo N, Jaiyen S (2015) Opinion mining for thai restaurant reviews using K-means clustering and MRF feature selection. In: 7th international conference on knowledge and smart technology (KST), pp 105–108

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Cover TM, Thomas JA (2012) Elements of information theory. Wiley, London

Da Silva NFF, Coletta LF, Hruschka ER, Hruschka ER Jr (2016) Using unsupervised information to improve semi-supervised tweet sentiment classification. Inf Sci 355:348–365

Dang Y, Zhang Y, Chen H (2010) A lexicon-enhanced method for sentiment classification: an experiment on online product reviews. IEEE Intell Syst 25(4):46–53

Daud A, Khan W, Che D (2017) Urdu language processing: a survey. Artif Intell Rev 47(3):279–311

Dave K, Lawrence S, Pennock DM (2003) Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In: Proceedings of the 12th international ACM conference on World Wide Web, pp 519–528

De A, Kopparapu SK (2013) Unsupervised clustering technique to harness ideas from an Ideas Portal. In: International IEEE conference on advances in computing, communications and informatics (ICACCI), pp 1563–1568

De Fortuny EJ, De Smedt T, Martens D, Daelemans W (2014) Evaluating and understanding text-based stock price prediction models. Inf Process Manag 50(2):426–441

Di Caro L, Grella M (2013) Sentiment analysis via dependency parsing. Comput Stand Interfaces 35(5):442–453

Duncan B, Zhang Y (2015) Neural networks for sentiment analysis on Twitter. In: IEEE 14th international conference on cognitive informatics & cognitive computing (ICCICC), pp 275–278

Duwairi RM, Qarqaz I (2014) Arabic sentiment analysis using supervised classification. In: International IEEE conference on future internet of things and cloud (FiCloud), pp 579–583

Ebrahimi M, Suen CY, Ormandjieva O (2016) Detecting predatory conversations in social media by deep convolutional neural networks. Digit Investig 18:33–49

Farra N, Challita E, Assi RA, Hajj H (2010) Sentence-level and document-level sentiment mining for arabic texts. In: Proceedings of IEEE international conference on data mining workshops, pp 1114–1119

Feng S, Song K, Wang D, Yu G (2015) A word-emoticon mutual reinforcement ranking model for building sentiment lexicon from massive collection of microblogs. World Wide Web 18(4):949–967

Fernández-Gavilanes M, Álvarez-López T, Juncal-Martínez J, Costa-Montenegro E, González-Castaño FJ (2016) Unsupervised method for sentiment analysis in online texts. Expert Syst Appl 58:57–75

Fersini E, Messina E, Pozzi FA (2016) Expressive signals in social media languages to improve polarity detection. Inf Process Manag 52(1):20–35

Ficamos P, Liu Y, Chen W (2017) A Naive Bayes and maximum entropy approach to sentiment analysis: capturing domain-specific data in Weibo. In: IEEE international conference on big data and smart computing (BigComp), pp 336–339

Gao W, Li S, Xue Y, Wang M, Zhou G (2014) Semi-supervised sentiment classification with self-training on feature subspaces. In: Su X, He T (eds) Chinese Lexical Semantics. CLSW 2014. Lecture notes in computer science, vol 8922. Springer, Cham, pp 231–239

Gao K, Xu H, Wang J (2015) A rule-based approach to emotion cause detection for Chinese micro-blogs. Expert Syst Appl 42(9):4517–4528

Garcia-Pablos A, Guadros M, Rigau G (2017) W2VLDA: almost unsupervised system for aspect based sentiment analysis. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2017.08.049

Goel A, Gautam J, Kumar S (2016) Real time sentiment analysis of tweets using Naive Bayes. In: 2nd international conference on next generation computing technologies (NGCT), pp 257–261

Grefenstette G, Qu Y, Shanahan JG, Evans DA (2004) Coupling niche browsers and affect analysis for an opinion mining application. In: Proceedings of RIAO '04 Coupling approaches, coupling media and coupling languages for information retrieval, pp 186–194

Gu X, Gu Y, Wu H (2017) Cascaded convolutional neural networks for aspect-based opinion summary. Neural Process Lett 46:1–20

Gupta E, Rathee G, Kumar P, Chauhan DS (2015) Mood swing analyser: a dynamic sentiment detection approach. Proc Natl Acad Sci India Sect A Phys Sci 85(1):149–157

Habernal I, Ptáček T, Steinberger J (2015) Supervised sentiment analysis in Czech social media. Inf Process Manag 51(4):532–546

Hajmohammadi MS, Ibrahim R, Selamat A (2014) Cross-lingual sentiment classification using multiple source languages in multi-view semi-supervised learning. Eng Appl Artif Intell 36:195–203

Hajmohammadi MS, Ibrahim R, Selamat A (2015) Graph-based semi-supervised learning for cross-lingual sentiment classification. In: Guyen N, Trawiński B, Kosala R (eds) Intelligent Information and Database Systems. ACIIDS 2015. Lecture notes in computer science, vol 9011. Springer, Cham, pp 97–106

Hasan KMA, Sabuj MS, Afrin Z (2015) Opinion mining using Naïve Bayes. In: IEEE International WIE conference on electrical and computer engineering (WIECON-ECE)

Hassan A, Radev D (2010) Identifying text polarity using random walks. In: Proceedings of the 48th annual meeting of the association for computational linguistics. Association for Computational Linguistics, pp 395–403

Hatzivassiloglou V, McKeown KR (1997) 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. Association for Computational Linguistics, pp 174–181

He Y, Zhou D (2011) Self-training from labeled features for sentiment analysis. Inf Process Manag 47(4):606–616

Hofmann T (1999) Probabilistic latent semantic indexing. In: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval. ACM, pp 50–57

Hong S, Lee J, Lee JH (2014) Competitive self-training technique for sentiment analysis in mass social media. In: 15th international symposium on soft computing and intelligent systems (SCIS), 2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS). IEEE, pp 9–12

Huang F, Zhang S, Zhang J, Yu G (2017) Multimodal learning for topic sentiment analysis in microblogging. Neurocomputing 253:144–153

Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 168–177

Iosifidis V, Ntutsi E (2017) Large scale sentiment learning with limited labels. In: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, pp 1823–1832

Irsoy O, Cardie C (2014) Opinion mining with deep recurrent neural networks. In: Proceedings of the conference on empirical methods in natural language processing (EMNLP)

Jain A, Jain M (2017) Location based Twitter opinion mining using common-sense information. Glob J Enterp Inf Syst 9(2):28–32

Jeyapriya A, Selvi K (2015) Extracting aspects and mining opinions in product reviews using supervised learning algorithm. In: 2015 2nd international conference on electronics and communication systems (ICECS). IEEE, pp 548–552

Jian Z, Chen X, Wang HS (2010) Sentiment classification using the theory of ANNs. J China Univ Posts Telecommun 17:58–62

Jiang D, Luo X, Xuan J, Xu Z (2017a) Sentiment computing for the news event based on the social media big data. IEEE Access 5:2373–2382

Jiang M, Wang J, Lan M, Wu Y (2017b) An effective gated and attention-based neural network model for fine-grained financial target-dependent sentiment analysis. In: Springer international conference on knowledge science, engineering and management, pp 42–54

Jiménez-Zafra SM, Martín-Valdivia MT, Martínez-Cámara E, Ureña-López LA (2015) Combining resources to improve unsupervised sentiment analysis at aspect-level. J Inf Sci 42:213–229

Jin W, Ho HH, Srihari RK (2009) 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

Joachims T (1998) Text categorization with support vector machines: learning with many relevant features. Springer, Berlin, pp 137–142

Joachims T (2003) Transductive learning via spectral graph partitioning. In: ICML, vol 3, pp 290–297

Kagan V, Stevens A, Subrahmanian VS (2015) Using twitter sentiment to forecast the 2013 pakistani election and the 2014 indian election. IEEE Intell Syst 1:2–5

Kalchbrenner N, Grefenstette E, Blunsom P (2014) A convolutional neural network for modelling sentences. In: Proceedings of the 52nd annual meeting of the association for computational linguistics

Kanayama H, Nasukawa T (2006) Fully automatic lexicon expansion for domain oriented sentiment analysis. In: Proceedings of the conference on empirical methods in natural language processing, Association for Computational Linguistics, pp 355–363

Keshavarz H, Abadeh MS (2017) ALGA: adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs. Knowl Based Syst 122:1–16

Keshtkar F, Inkpen D (2013) A bootstrapping method for extracting paraphrases of emotion expressions from texts. Comput Intell 29(3):417–435

Khan FH, Bashir S, Qamar U (2014) TOM: Twitter opinion mining framework using hybrid classification scheme. Decis Support Syst 57:245–257

Khan FH, Qamar U, Bashir S (2016) Multi-objective model selection (MOMS)-based semi-supervised framework for sentiment analysis. Cogn Comput 8(4):614–628
Khan FH, Qamar U, Bashir S (2017) Lexicon based semantic detection of sentiments using expected likelihood estimate smoothed odds ratio. Artif Intell Rev 48(1):113–138
Kisioglu P, Topcu YI (2011) Applying Bayesian belief network approach to customer churn analysis: a case study on the telecom industry of Turkey. Expert Syst Appl 38(6):7151–7157
Kobayashi N, Inui K, Matsumoto Y (2007) Extracting aspect-evaluation and aspect-of relations in opinion mining. In: EMNLP-CoNLL, vol 7, pp 1065–1074
Kranjc J, Smailović J, Podpečan V, Grčar M, Žnidaršič M, Lavrač N (2015) Active learning for sentiment analysis on data streams: methodology and workflow implementation in the ClowdFlows platform. Inf Process Manag 51(2):187–203
Kumar S, Morstatter F, Liu H (2014) Twitter data analytics. Springer, Berlin
Lafferty J, McCallum A, Pereira FC (2001) Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the eighteenth international conference on machine learning, pp 282–289
Li G, Liu F (2012) Application of a clustering method on sentiment analysis. J Inf Sci 38(2):127–139
Li G, Liu F (2014) Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. Appl Intell 40(3):441–452
Li G, Chang K, Hoi SC (2012) Multiview semi-supervised learning with consensus. IEEE Trans Knowl Data Eng 24(11):2040–2051
Li S, Zhou L, Li Y (2015) Improving aspect extraction by augmenting a frequency-based method with web-based similarity measures. Inf Process Manag 51(1):58–67
Li Q, Jin Z, Wang C, Zeng DD (2016) Mining opinion summarizations using convolutional neural networks in Chinese microblogging systems. Knowl Based Syst 107:289–300
Li Q, Guo X, Bai X (2017) Weekdays or weekends: exploring the impacts of microblog posting patterns on gratification and addiction. Inf Manag 54(5):613–624
Liao C, Feng C, Yang S, Huang H (2016) A hybrid method of domain lexicon construction for opinion targets extraction using syntax and semantics. J Comput Sci Technol 31:595–603
Liu B (2007) Web data mining: exploring hyperlinks, contents, and usage data. Springer, Berlin
Liu B (2012) Sentiment analysis and opinion mining. Synthesis lectures on human language technologies. Morgan & Calypool Publishers, pp 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016
Liu B (2015) Sentiment analysis: mining opinions, sentiments, and emotions. Cambridge University Press, Cambridge
Liu B, Zhang L (2012) A Survey of Opinion Mining and Sentiment Analysis. In: Aggarwal C., Zhai C. (eds) Mining text data. Springer, Boston, MA, pp 415–463
Liu J, Seneff S, Zue V (2012) Harvesting and summarizing user-generated content for advanced speech-based HCI. IEEE J Sel Top Signal Process 6(8):982–992
Liu S, Li F, Li F, Cheng X, Shen H (2013a) Adaptive co-training SVM for sentiment classification on tweets. In: Proceedings of the 22nd ACM international conference on conference on information & knowledge management. ACM, pp 2079–2088
Liu S, Zhu W, Xu N, Li F, Cheng XQ, Liu Y, Wang Y (2013b) Co-training and visualizing sentiment evolvement for tweet events. In: Proceedings of the 22nd international conference on World Wide Web companion. International World Wide Web Conferences Steering Committee, pp 105–106
Lo SL, Cambria E, Chiong R, Cornforth D (2017) Multilingual sentiment analysis: from formal to informal and scarce resource languages. Artif Intell Rev 48(4):499–527
Lu TJ (2015) Semi-supervised microblog sentiment analysis using social relation and text similarity. In: 2015 international conference on big data and smart computing (BigComp). IEEE, pp 194–201
Luo W, Zhuang F, Zhao W, He Q, Shi Z (2015) QPLSA: utilizing quad-tuples for aspect identification and rating. Inf Process Manag 51(1):25–41
Lv Y, Liu J, Chen H, Mi J, Liu M, Zheng Q (2017) Opinioned post detection in Sina Weibo. IEEE Access 5:7263–7271
Ma B, Zhang N, Liu G, Li L, Yuan H (2015) Semantic search for public opinions on urban affairs: a probabilistic topic modeling-based approach. Inf Process Manag 52:430
Ma H, Jia M, Zhang D, Lin X (2017) Combining tag correlation and user social relation for microblog recommendation. Inf Sci 385–386:325–337
Majumder N, Poria S, Gelbukh A, Cambria E (2017) Deep learning-based document modeling for personality detection from text. IEEE Intell Syst 32(2):74–79
Manek AS, Shenoy PD, Mohan MC, Venougopal KR (2017) Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. World Wide Web 20(2):135–154

Marcheggiani D, Täckström O, Esuli A, Sebastiani F (2014) Hierarchical multi-label conditional random fields for aspect-oriented opinion mining. In: de Rijke M et al (eds) Advances in information retrieval. ECIR 2014. Lecture notes in computer science, vol 8416. Springer, Cham, pp 273–285

Marrese-Taylor E, Velásquez JD, Bravo-Marquez F (2014) A novel deterministic approach for aspect-based opinion mining in tourism products reviews. Expert Syst Appl 41(17):7764–7775

Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J 5(4):1093–1113

Mele I (2013) Web usage mining for enhancing search-result delivery and helping users to find interesting web content. In: Proceedings of the sixth ACM international conference on Web search and data mining. ACM, pp 765–770

Mesnil G, Mikolov T, Ranzato MA, Bengio Y (2015) Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews. Preprint. arXiv:1412.5335

Mihalcea R, Banea C, Wiebe JM (2007) Learning multilingual subjective language via cross-lingual projections. In: Proceedings of the Association for Computational Linguistics (ACL 2007), Prague

Mohammad SM, Zhu X, Kiritchenko S, Martin J (2015) Sentiment, emotion, purpose, and style in electoral tweets. Inf Process Manag 51(4):480–499

Molina-González MD, Martínez-Cámara E, Martín-Valdivia MT, Urena-López LA (2014) Cross-domain sentiment analysis using Spanish opinionated words. In: Métais E, Roche M, Teisseire M (eds) Natural language processing and information systems. NLDB 2014. Lecture notes in computer science, vol 8455. Springer, Cham, pp 214–219

Molina-González MD, Martínez-Cámara E, Martín-Valdivia MT, Ureña-López LA (2015) A Spanish semantic orientation approach to domain adaptation for polarity classification. Inf Process Manag 51(4):520–531

Moraes R, Valiati JF, Neto WPG (2013) Document-level sentiment classification: an empirical comparison between SVM and ANN. Expert Syst Appl 40(2):621–633

Moreo A, Romero M, Castro JL, Zurita JM (2012) Lexicon-based comments-oriented news sentiment analyzer system. Expert Syst Appl 39(10):9166–9180

Mudinas A, Zhang D, Levene M (2012) Combining lexicon and learning based approaches for concept-level sentiment analysis. In: Proceedings of the first international workshop on issues of sentiment discovery and opinion mining

Muhammad A, Wiratunga M, Lothian R (2016) Contextual sentiment analysis for social media genres. Knowl Based Syst 108:92–101

Mukherjee A, Liu B (2012) Aspect extraction through semi-supervised modeling. In: Proceedings of the 50th annual meeting of the association for computational linguistics: long papers—volume 1. Association for Computational Linguistics, pp 339–348

Mullen T, Collier N (2004) Sentiment analysis using support vector machines with diverse information sources. In: EMNLP, vol 4, pp 412–418

Nofer M, Hinz O (2015) Using Twitter to predict the stock market. Bus Inf Syst Eng 57(4):229–242

Olson DL, Delen D (2008) Advanced data mining techniques. Springer, Berlin

Pandarachalil R, Sendhilkumar S, Mahalakshmi GS (2015) Twitter sentiment analysis for large-scale data: an unsupervised approach. Cogn Comput 7(2):254–262

Pandey AC, Rajpoot DS, Saraswat M (2017) Twitter sentiment analysis using hybrid cuckoo search method. Inf Process Manag 53(4):764–779

Pang B, Lee L (2004) A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics

Pang B, Lee L, Vaithyanathan S (2002) 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

Parveen H, Pandey S (2016) Sentiment analysis on Twitter data-set using Naive Bayes algorithm. In: 2nd international conference on applied and theoretical computing and communication technology (iCATccT), pp 416–419

Penalver-Martinez I, Garcia-Sanchez F, Valencia-Garcia R, Rodriguez-Garcia MA, Moreno V, Fraga A, Sanchez-Cervantes JL (2014) Feature-based opinion mining through ontologies. Expert Syst Appl 41(13):5995–6008

Peng H, Cambria E, Hussain A (2017) A review of sentiment analysis research in Chinese language. Cogn Comput 9(4):423–435

Petz G, Karpowicz M, Fürschuß H, Auinger A, Stříteský V, Holzinger A (2015) Computational approaches for mining user's opinions on the Web 2.0. Inf Process Manag 51(4):510–519

Pham D, Le A (2017) Learning multiple layers of knowledge representation for aspect based sentiment analysis. Data Knowl Eng. https://doi.org/10.1016/j.datak.2017.06.001

Phu VN, Dat ND, Tran VTN, Chau VTN, Nguyen TA (2017) Fuzzy C-means for english sentiment classification in a distributed system. Appl Intell 46(3):717–738
Ponomareva N (2014) Graph-based approaches for semi-supervised and cross-domain sentiment analysis. PhD Thesis, University of Wolverhampton
Poria S, Gelbukh A, Hussain A, Howard N, Das D, Bandyopadhyay S (2013) Enhanced SenticNet with affective labels for concept-based opinion mining. IEEE Intell Syt 28(2):31–38
Poria S, Cambria E, Winterstein G, Huang GB (2014) Sentic patterns: dependency-based rules for concept-level sentiment analysis. Knowl Based Syst 69:45–63
Poria A, Cambria E, Gelbukh A (2016) Aspect extraction for opinion mining with a deep convolutional neural network. Knowl Based Syst 108:42–49
Poria S, Peng H, Hussan A, Howard N, Cambria E (2017) Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis. Neurocomputing 261:217–230
Qazi A, Syed KBS, Raj RG, Cambria E, Tahir M, Alghazzawi D (2016) A concept-level approach to the analysis of online review helpfulness. Comput Hum Behav 58:75–81
Qiu G, Liu B, Bu J, Chen C (2011) Opinion word expansion and target extraction through double propagation. Comput Linguist 37(1):9–27
Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106
Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286
Ramadhani RA, Indirani F, Nugrahadi DT (2016) Comparison of Naive Bayes smoothing methods for Twitter sentiment analysis. In: International conference on advanced computer science and information systems (ICACSIS), pp 287–292
Rana TA, Cheah Y (2016) Aspect extraction in sentiment analysis: comparative analysis and survey. Artif Intell Rev 46(4):459–483
Rao Y, Lei J, Wenyin L, Li Q, Chen M (2014) Building emotional dictionary for sentiment analysis of online news. World Wide Web 17(4):723–742
Rathan M, Hulipalled VR, Venugopal KR, Patnaik LM (2017) Consumer insight mining: aspect based Twitter opinion mining of mobile phone reviews. Appl Soft Comput. https://doi.org/10.1016/j.asoc.2017.07.056
Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowl Based Syst 89:14–46
Ren F, Kang X (2013) Employing hierarchical Bayesian networks in simple and complex emotion topic analysis. Comput Speech Lang 27(4):943–968
Riaz S, Fatima M, Kamran M, Nasir MW (2017) Opinion mining on large scale data using sentiment analysis and k-means clustering. Clust Comput 20:1–16
Rout JK, Dalima A, Choo KR, Bakshi S, Jena SK (2017) Revisiting semi-supervised learning for online deceptive review detection. IEEE Access 5:1319–1327
Saif H, He Y, Fernandez M, Alani H (2016) Contextual semantics for sentiment analysis of Twitter. Inf Process Manag 52(1):5–19
Saleh MR, Martín-Valdivia MT, Montejo-Ráez A, Ureña-López LA (2011) Experiments with SVM to classify opinions in different domains. Expert Syst Appl 38(12):14799–14804
Scholer F, Kelly D, Carterette B (2016) Information retrieval evaluation using test collections. Inf Retr J 19(3):225–229
Severyn A, Moschitti A, Uryupina O, Plank B, Filippova K (2016) Multi-lingual opinion mining on youtube. Inf Process Manag 52(1):46–60
Shah RR, Yu Y, Verma A, Tang S, Shaik AD, Zimmermann R (2016) Leveraging multimodal information for event summarization and concept-level sentiment analysis. Knowl Based Syst 108:102–109
Sharma R, Nigam S, Jain R (2014) Opinion mining of movie reviews at document level. Preprint. arXiv:1408.3829
Shi B, Chang K (2008) Generating a concept hierarchy for sentiment analysis. In: IEEE international conference on systems, man and cybernetics, SMC 2008. IEEE, pp 312–317
Sierra B, Lazkano E, Jauregi E, Irigoien I (2009) Histogram distance-based bayesian network structure learning: a supervised classification specific approach. Decis Support Syst 48(1):180–190
Sindhwani V, Melville P (2008) Document-word co-regularization for semi-supervised sentiment analysis. In: 8th IEEE international conference on data mining, pp 1025–1030
Singh J, Gupta V (2017) A systematic review of text stemming techniques. Artif Intell Rev 48(2):157–217
Sisodia DS, Verma S (2012) Web usage pattern analysis through web logs: a review. In: 2012 international joint conference on computer science and software engineering (JCSSE). IEEE, pp 49–53
Sohrabi MK (2018) A gossip-based information fusion protocol for distributed frequent itemset mining. Enterp Inf Syst. https://doi.org/10.1080/17517575.2017.1405286

Sohrabi MK, Akbari S (2016) A comprehensive study on the effects of using data mining techniques to predict tie strength. Comput Hum Behav 60:534–541
Sohrabi MK, Azgomi H (2017a) Parallel set similarity join on big data based on locality-sensitive hashing. Sci Comput Program 145:1–12
Sohrabi MK, Azgomi H (2017b) TSGV: a table-like structure based greedy method for materialized view selection in data warehouse. Turk J Electr Eng Comput Sci 25(4):3175–3187
Sohrabi MK, Barforoush AA (2012) Efficient colossal pattern mining in high dimensional datasets. Knowl Based Syst 33:41–52
Sohrabi MK, Barforoush AA (2013) Parallel frequent itemset mining using systolic arrays. Knowl Based Syst 37:462–471
Sohrabi MK, Ghods V (2014) Top-down vertical itemset mining. In: Proceedings of the SPIE 9443 sixth international conference on graphic and image processing
Sohrabi MK, Ghods V (2015) Top- materialized view selection for a data warehouse using frequent itemset mining. In: Proceedings of the ICACTE conference, Berlin, Germany
Sohrabi MK, Ghods V (2016) CUSE: a novel cube-based approach for sequential pattern mining. In: Proceedings of the IEEE international symposium on computational business intelligence, Olten, Switzerland
Sohrabi MK, Karimi F (2018) Feature selection approach to detect spam in the Facebook social network. Arab J Sci Eng. https://doi.org/10.1007/s13369-017-2855-x
Sohrabi MK, Marzooni HH (2016) Association rule mining using new FP-linked list algorithm. J Adv Comput Res 7(01):23–34
Sohrabi MK, Roshani R (2017) Frequent itemset mining using cellular learning automata. Comput Hum Behav 68:244–253
Sohrabi MK, Tajik A (2017) Multi-objective feature selection for warfarin dose prediction. Comput Biol Chem 69:126–133
Speriosu M, Sudan N, Upadhyay S, Baldridge J (2011) Twitter polarity classification with label propagation over lexical links and the follower graph. In: Proceedings of the first workshop on unsupervised learning in NLP. Association for Computational Linguistics, pp 53–63
Subrahmanian VS, Reforgiato D (2008) AVA: adjective-verb-adverb combinations for sentiment analysis. IEEE Intell Syst 23(4):43–50
Subramanya A, Bilmes J (2011) Semi-supervised learning with measure propagation. J Mach Learn Res 12:3311–3370
Sun J, Wang G, Cheng X, Fu Y (2015) Mining affective text to improve social media item recommendation. Inf Process Manag 51(4):444–457
Sun S, Luo C, Chen J (2017) A review of natural language processing techniques for opinion mining systems. Inf Fusion 36:10–25
Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. Comput Linguist 37(2):267–307
Talukdar PP, Crammer K (2009) New regularized algorithms for transductive learning. In: Buntine W, Grobelnik M, Mladenić D, Shawe-Taylor J (eds) Machine learning and knowledge discovery in databases. ECML PKDD 2009. Lecture notes in computer science, vol 5782. Springer, Berlin, pp 442–457
Tang D, Qin B, Liu T, Yang Y (2015) User modeling with neural network for review rating prediction. In: Proceedings of IJCAI, pp 1340–1346
Tang H, Tan S, Cheng X (2009) A survey on sentiment detection of reviews. Expert Syst Appl 36(7):10760–10773
Tripathy A, Agrawal A, Rath SK (2016) Classification of sentiment reviews using n-gram machine learning approach. Expert Syst Appl 57:117–126
Titov I, McDonald R (2008) Modeling online reviews with multi-grain topic models. In: Proceedings of the 17th international conference on World Wide Web, pp 111–120
Tsagkalidou K, Koutsonikola V, Vakali A, Kafetsios K (2011) Emotional aware clustering on micro-blogging sources. In: D'Mello S, Graesser A, Schuller B, Martin JC (eds) Affective computing and intelligent interaction. ACII 2011. Lecture notes in computer science, vol 6974. Springer, Berlin, pp 387–396
Tsai AC, Wu C, Tsai RT, Hsu JY (2013) Building a concept-level sentiment dictionary based on commonsense knowledge. IEEE Intell Syt 28(2):22–30
Tsakalidis A, Papadopoulos S, Cristea AI, Kompatsiaris Y (2015) Predicting elections for multiple countries using Twitter and polls. IEEE Intell Syst 30(2):10–17
Turney P (2002) 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 ACL'02, Association for Computational Linguistics, pp 417–424

Unankard S, Li X, Sharaf M, Zhong J, Li X (2014) Predicting elections from social networks based on sub-event detection and sentiment analysis. In: Web information systems engineering—WISE 2014. Springer, Berlin, pp 1–16
Vakali A, Kafetsios K (2012) Emotion aware clustering analysis as a tool for Web 2.0 communities detection: implications for curriculum development. In: World Wide Web Conference. WWW
Velásquez JD (2013) Combining eye-tracking technologies with web usage mining for identifying Website Keyobjects. Eng Appl Artif Intell 26(5):1469–1478
Vilares D, Alonso MA, Gómez-Rodríguez C (2017) Supervised sentiment analysis in multilingual environments. Inf Process Manag 53(3):595–607
Vinodhini G, Chandrasekaran RM (2016) A comparative performance evaluation of neural network based approach for sentiment classification of online reviews. J King Saud Univ Comput Inf Sci 28(1):2–12
Vulić I, De Smet W, Tang J, Moens MF (2015) Probabilistic topic modeling in multilingual settings: an overview of its methodology and applications. Inf Process Manag 51(1):111–147
Wan X (2011) Bilingual co-training for sentiment classification of Chinese product reviews. Comput Linguist 37(3):587–616
Wang G, Zhang Z, Sun J, Yang S, Larson CA (2015a) POS-RS: a random subspace method for sentiment classification based on part-of-speech analysis. Inf Process Manag 51(4):458–479
Wang J, Cong G, Zhao XW, Li X (2015b) Mining user intents in twitter: a semi-supervised approach to inferring intent categories for tweets. In: Twentyninth AAAI conference on artificial intelligence
Wang J, Xue Y, Li S, Zhou G (2015c) Leveraging interactive knowledge and unlabeled data in gender classification with co-training. In: Liu A, Ishikawa Y, Qian T, Nutanong S, Cheema M (eds) Database Systems for Advanced Applications. DASFAA 2015. Lecture notes in computer science, vol 9052. Springer, Cham, pp 246–251
Wang G, Zheng D, Yang S (2017a) FCE-SVM: a new cluster based ensemble method for opinion mining from social media. Inf Syst e-Bus Manag 15:1–22
Wang W, Tan G, Wang H (2017b) Cross-domain comparison of algorithm performance in extracting aspect-based opinions from Chinese online reviews. Int J Mach Learn Cybern 8(3):1053–1070
Wehrmann J, Becker W, Cagnini HE, Barros RC (2017) A character-based convolutional neural network for language-agnostic Twitter sentiment analysis. In: IEEE international joint conference on neural networks (IJCNN), pp 2384–2391
Wen S, Wan X (2014) Emotion classification in microblog texts using class sequential rules. In: Twentyeighth AAAI conference on artificial intelligence
Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase level sentiment analysis. In: Proceedings of HLT/EMNLP-05
Wu Y, Zhang Q, Huang X, Wu L (2009) Phrase dependency parsing for opinion mining. In: Proceedings of the 2009 conference on empirical methods in natural language processing: volume 3. Association for Computational Linguistics, pp 1533–1541
Wu F, Song Y, Huang Y (2016) Microblog sentiment classification with heterogeneous sentiment knowledge. Inf Sci 373:149–164
Xia R, Zong C, Li S (2011) Ensemble of feature sets and classification algorithms for sentiment classification. Inf Sci 181(6):1138–1152
Xia Y, Cambria E, Hussain A (2015) AspNet: aspect extraction by bootstrapping generalization and propagation using an aspect network. Cogn Comput 7(2):241–253
Xia R, Xu F, Yu J, Qi Y, Cambria E (2016) Polarity shift detection, elimination and ensemble: a three-stage model for document-level sentiment analysis. Inf Process Manag 52(1):36–45
Xing FZ, Cambria E, Welsch RE (2018) Natural language based financial forecasting: a survey. Artif Intell Rev. https://doi.org/10.1007/s10462-017-9588-9
Xu L, Lin J, Wang L, Yin C, Wang J (2017) Deep convolutional neural network based approach for aspect-based sentiment analysis. Adv Sci Technol Lett 143:199–204
Yan X, Huang T (2015) Tibetan sentence sentiment analysis based on the maximum entropy model. In: 10th international conference on broadband and wireless computing, communication and applications (BWCCA), pp 594–597
Yan Z, Jiang X, Pedryc W (2017) Fusing and mining opinions for reputation generation. Inf Fusion 36:172–184
Yang B, Cardie C (2014) Context-aware learning for sentence-level sentiment analysis with posterior regularization. In: ACL, no 1, pp 325–335
Yin PY, Guo YM (2013) Optimization of multi-criteria website structure based on enhanced tabu search and web usage mining. Appl Math Comput 219(24):11082–11095
Yu J, Zha ZJ, Wang M, Chua TS (2011) Aspect ranking: identifying important product aspects from online consumer reviews. In: Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, vol 1, pp 1496–1505

Zhang X, Gong W, Kawamura Y (2004) Customer behavior pattern discovering with web mining. In: Yu JX, Lin X, Lu H, Zhang Y (eds) Advanced web technologies and applications. APWeb, Lecture notes in computer science, vol 3007. Springer, Berlin, pp 844–853

Zhou F, Jiao JR, Yang XJ, Lei B (2017) Augmenting feature model through customer preference mining by hybrid sentiment analysis. Expert Syst Appl 89:306–317

Zhuang L, Jing F, Zhu XY (2006) Movie review mining and summarization. In: Proceedings of the 15th ACM international conference on information and knowledge management, pp 43–50

Zimmermann M, Ntoutsi E, Spiliopoulou M (2016) Extracting opinionated (sub) features from a stream of product reviews using accumulated novelty and internal re-organization. Inf Sci 329:876–899

Zimmermann M, Ntoutsi E, Spiliopoulou M (2014) A semi-supervised self-adaptive classifier over opinionated streams. In: 2014 IEEE international conference on data mining workshop (ICDMW). IEEE, pp 425–432