Check for
updates

Mining and classifying customer reviews: a survey

L. D. C. S. Subhashini[1,2,3] · **Yuefeng Li**[1] · **Jinglan Zhang**[1] · **Ajantha S. Atukorale**[2] · **Yutong Wu**[1]

Abstract

With the increasing number of customer reviews on the Web, there is a growing need for effective methods to retrieve valuable information hidden in these reviews, as sellers need to gain a deep understanding of customers' preferences in a timely manner. With the continuous enhancement of opinion mining or sentiment analysis research, researchers have proposed many automatic mining and classification methods. However, how to choose a trusted method is a difficult problem for companies, because customer reviews (or opinions) contain a lot of uncertain information and noise. This article reports on a detailed survey of recent opinion mining literature. It also reviews how to extract text features in opinions that may contain noise or uncertainties, how to express knowledge in opinions, and how to classify them. Through this extensive study, this paper discusses open questions and recommends future research directions for building the next generation of opinion mining systems.

Keywords Opinion mining · Feature selection · Knowledge representation · Classification · Sentiment analysis

1 Introduction

E-commerce is growing fast, due to its convenience and reliability. In order to improve customer satisfaction and e-commerce experience, online sellers facilitate reviews whereby customers can express their opinions on products or services that they purchase. As the number of online customers increases, the number of reviews for products or services not only increases rapidly, but also generates many challenging issues in dealing with a large number of text reviews. Because most review collections are very long, it is difficult to monitor each customer opinion separately (Zhang et al. 2010; Cambria 2016). Text mining techniques have been employed to generate useful information or insights from customer reviews for organizations. This exercise is known as sentiment analysis or opinion mining

✉ Yuefeng Li
  y2.li@qut.edu.au

1   Queensland University of Technology, Brisbane, Australia

2   University of Colombo School of Computing, Colombo, Sri Lanka

3   University of Sri Jayewardenepura, Nugegoda, Sri Lanka

(Hu and Liu 2004; Nasukawa and Yi 2003; Cambria et al. 2013b). So far, many models have been proposed in recent years for opinion mining (Quan et al. 2004; Catal and Nangir 2017; Angelpreethi and Kumar 2017; Araújo et al. 2014; Howells and Ertugan 2017; Lau et al. 2009; Poria et al. 2015). The existing opinion mining models provide valuable insights in the e-commerce arena. A lot of business companies can use discovered knowledge in reviews to make strategic plans based on customer preference patterns that have significant impacts on the overall profit of the companies. Also, marketers can accurately evaluate the success of a new product launch using the opinion mining results, as opinion mining models will help them to decide the popularity of products based on market segmentation. If such a system can also cope with uncertainties in opinions, then its users can make more correct decisions. Based on such innovations, new market segments can be created and the profit can be increased. In addition, understanding and considering reviews helps to gain the faith of customers and grow business by providing expansion strategies.

One of the most recent survey papers about opinion mining and sentiment analysis was presented by Piryani et al. (2017), which illustrated analytical mapping of opinion mining and sentiment analysis during the period of 2000 and 2015. The paper focused on identified major publication sources, topics, thematic trends, opinion mining and sentiment analysis levels, data sources and applications. Hence, there is a need to ascertain how opinion mining and sentiment analysis research has been enhanced since 2015. Another survey about how opinion mining had been used in management research was done by Mukhopadhyay (2018). It addressed some major open problems or limitations in the opinion mining area. Most of them have been solved with the recent enhancement of opinion mining research. However, what has happened in this field between 2015 to 2020 has not yet been clearly summarized. We found that one main advance in this period is the ability to handle uncertainties in opinions using fuzzy logic. These fuzzy models were not mentioned in any previous survey papers (Khan et al. 2014; Ravi and Ravi 2015; Cambria et al. 2013b; Kadhim 2019). Further, most survey papers addressed the applications and challenges of opinion mining without identifying how to extract the opinions, how to represent knowledge, and how to classify them (Lo et al. 2017; Hemmatian and Sohrabi 2019; Yousif et al. 2019; Zhou et al. 2012). Taking into account all the above findings, there was a need to survey opinion mining within the time period 2000–2020. This article introduces the latest updates in opinion mining by considering how to extract opinions, how to express knowledge in opinions, and how to classify opinions in a new framework.

The main goal of this survey paper is to introduce and analyze a novel classification that divides opinion mining articles into the following categories: text feature selection, knowledge representation in opinions, and sentiment classification. The paper reviewed one hundred and twenty papers published from 2000 to 2020 using the proposed selection criteria (see Sect. 2.2 for more details). This research points out the direction for the future development in opinion mining and sentiment analysis. It has shown that Support Vector Machine (SVM) and Naive Bayes (NB) are the most widely used supervised approaches for opinion mining. Fuzzy logic and pattern mining have been used in opinion mining to represent knowledge and reasoning mechanisms. To solve the problem of uncertainties in opinion mining, fuzzy logic has been applied by several researchers, but the problem still exists, since it is difficult to identify a clear boundary between the positive and the negative class. By using one or more feature selection methods to select opinion words, feature selection helps sentiment classification. Knowledge discovery also contributed to opinion mining for improving the accuracy of opinion classification. In the Natural Language Processing (NLP) community, some classifiers use features that utilize discovered knowledge, or use both features and discovered knowledge. Using discovered knowledge wisely can

improve the accuracy of classification. Finally, the paper discusses open questions and future research directions.

This paper is organized as follows: Sect. 2 presents an overview of opinion mining, Sect. 3 reviews feature selection in opinion mining, Sect. 4 discusses knowledge representation, and Sect. 5 looks at detail in sentiment classification. In Sect. 6, real world insights, open questions and future trends are discussed, and finally the conclusion is presented.

2 Opinion mining overview

This section first discusses the opinion mining process and outlines each step. It then outlines the review methodology to describe how the research was conducted.

2.1 Process of opinion mining classification

Hu and Liu (2004) defined an opinion in terms of the goal (entity) of the opinion, the attributes of the goal that the opinion is aimed at, and the sentiment (polarity) of a positive, negative or neutral opinion. There are two types of opinions expressed by customers. One type is called regular opinion and the other type is called comparative opinion (Hu and Liu 2004; Nasukawa and Yi 2003; Jindal and Liu 2006). The objective of regular opinions is to give opinion about only one entity with respect to one feature. Conversely, comparative opinion is illustrated for more than one entity with respect to some common features. Most of the research between 2000 and 2020 has been done based on regular opinions, as shown in Fig. 1.

There are mainly two approaches in sentiment analysis: subjectivity and polarity detection. Subjectivity detection is about understanding whether the content contains personal views and opinions as opposed to factual information. Polarity detection is about studying subjectivity with different polarities, intensities or rankings (Lo et al. 2017).

Another way is to classify opinions either explicit or implicit, based on how they have been expressed by customers (Hu and Liu 2004; Nasukawa and Yi 2003). In opinion mining, there are three main classification levels used in the research arena: document level, sentence level, and aspect/feature level.

- Document level classification
  In document level classification, a complete document is considered to be the opinion of the customer review, and extracted opinion is used to decide whether the review expresses a positive or negative opinion about the product (Pang et al. 2002; Turney 2002). Supervised learning methods are most widely used in document level classification (Pang et al. 2002). Unsupervised methods are also tested in document level classification (Turney 2002; Sharma et al. 2014; Moraes et al. 2013; Gao et al. 2015).
- Sentence level classification
  A sentence level classification method assumes that a sentence contains only one single opinion and decides whether the sentence expresses a positive or negative opinion. Most models available in this level use supervised learning methods (Saleh et al. 2011; Chen et al. 2011; Ravi et al. 2017; Afzaal et al. 2016).
- Aspect/feature level classification
  Aspect/feature level opinion mining classifiers identify entities, features and relationships between opinions separately. Supervised learning approaches and lexicon
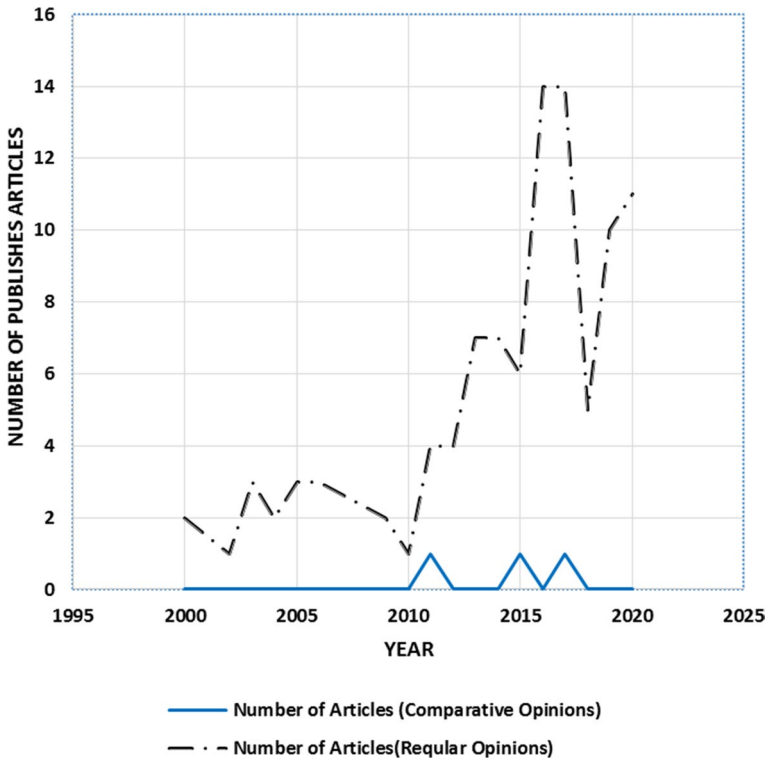
Fig. 1 Published articles: regular versus comparative opinions

based methods are tested in this level (Gu et al. 2017; Jiménez-Zafra et al. 2016; Ravi et al. 2017; Afzaal et al. 2016; Jiménez-Zafra et al. 2016; Spasic et al. 2017). An unsupervised approach was also proposed recently (Jing et al. 2018).

The opinion mining process consists of opinion text pre-processing, entity and feature selection, training set selection, knowledge discovery, classifier design and development and evaluation.

Text pre-processing is the process of removing irrelevant parts of the reviews to improve the performance of classification. Online reviews contain a lot of noise and irrelevant parts. Proper preprocessing of the data can improve the performance of the classifier; it involves tokenization, stop word removal and stemming (Gu et al. 2017). Tokenization is the method of identifying appropriate parts of the text using a parser. For any alphanumeric characters, we use spaces or special characters to terminate them and convert all words to lower case. A stop word removal method will create a list of stop words from high-frequency words, which can be customized according to the application, and then delete these words in the document. Stemming reduces the morphological variations of words using stemmers such as algorithmic (Fautsch and Savoy 2009) or Porter stemmers (Vijayarani et al. 2015). The Porter stemming algorithm has been very widely used for stemming to reduce the text data redundancy (Liu 2010).

Entity and feature selection is the process of reducing the data and identifies the relevant features for the classification process, which makes the classifier perform more effectively. Normally text features are a set of words (or terms). Some classifiers directly use features for the classification process.

Knowledge discovery and presentation is the process of discovering hidden knowledge from the customer review dataset. Knowledge can be represented in term-based, pattern-based, n-grams, part-of-speech (POS) tag-based, ontology-based, semantic aspect-based, rule-based, lexicon sentiment, and other forms. In addition to NLP, fuzzy logic and pattern mining are the latest successful knowledge discovery methods in the field of opinion mining. Compared with a classifier that uses only features, a classifier that uses features and knowledge together performs better.

After obtaining discovered knowledge in textual opinions, another important task is to *develop sentiment classifiers to group opinions effectively* by using features, the discovered knowledge, or the combination of features and knowledge. The sentiment classifiers can be divided into supervised or unsupervised classifiers. Supervised learning approaches are widely used in opinion mining. Their learning methods depend on sets of labelled data (training sets). On the other hand, unsupervised methods are based on unlabelled data sets and grouping reviews into clusters based on similarity measures between these reviews.

Evaluation is also a very important task of the opinion mining process. It is important to use evaluation techniques such as benchmark techniques and benchmark data sets to evaluate sentiment classifiers, such as by comparison with the latest models and popular indicators. Most opinion mining models are evaluated using popular measures of accuracy, precision, recall and F-measure.

2.2 Review methodology

We have applied the review method of Boote and Beile (2005) for this survey. It explores beliefs and topics, and then initiates the search. Then there is storage and organization of information. After that, selecting relevant information is very important. The method also expands the search into several databases and performs an interpretation stage to analyze and synthesize information. Finally, in the communication phase, it gives the results.

The one hundred and twenty articles reviewed in this survey were selected based on the selection criteria (Onwuegbuzie and Frels 2016). Fig. 2 shows the annual distribution of the articles. Most of the selected papers were published between 2016 and 2020. The main criteria for this literature review are:

- Keywords: information retrieval (IR), opinion mining, sentiment analysis, text mining, data mining and artificial intelligence (AI)
- Publication year: 2000–2020
- Citation: citation greater than 10
- Source: the dominant source is journal papers as they provided salient domain knowledge with their research findings. Conference papers and websites are also included as they present the latest or exclusive opinions. Selective reading of textbooks provides practical guidance as well.

The survey is then followed by three steps: feature selection, knowledge representation, and classification of opinion. The main target of this survey is to introduce and analyse a novel opinion mining categorization through these three stages.
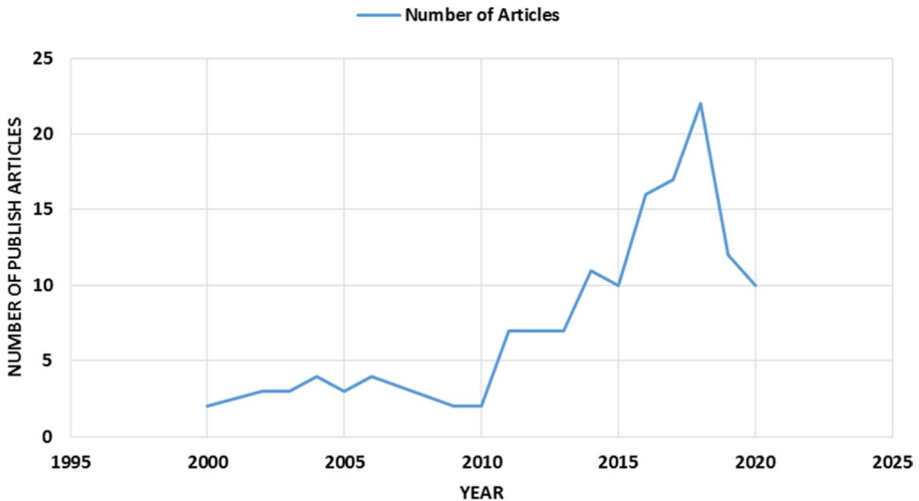
Fig. 2 Annual distribution of papers

3 Feature selection in opinion mining

Opinion mining models have to deal with a huge amount of complex, unstructured review data with numerous features. Aiming to find significant features and reduce the workload of the opinion mining classifiers, feature selection is an important technique in opinion mining. An opinion mining classifier with feature selection has shown significant results with greater accuracy than classification without feature selection. This section explores the different types of features and feature selection methods in opinion mining applications.

3.1 Different features in oinion mining

Features can take two forms, including explicit features (if the feature appears in the review) and implicit features, where the opinion words are considered as feature indicators (if the feature does not appear but is implied in the review) (Noekhah et al. 2017). Valuable meta-data about opinion is explained through their features. Noekhah et al. (2017) identified features in opinion mining, as shown in Fig. 3.

Three forms of data are commonly used in opinion mining applications, including structural (behavioral) information, textual (linguistic) information and relational (network extracted) information. Linguistic features include linguistic or semantic features. These features, such as POS tags, sentiment terms, length and similarity of words, are extracted from reviews and used for the classification process (Martineau et al. 2009; Pasquier et al. 1999; Zimmermann et al. 2015; Li et al. 2010b; Cambria et al. 2012). As unstructured text data are the major parts of customer reviews, the linguistic features became the most valuable features for opinion mining classification. Thus almost all linguistic or semantic features in text classification, such as term frequency, key words, topical words, co-occurrence, similarity, etc., can be used for opinion mining tasks as well. In addition to linguistic features, there are some non-content-related features. Structural (behavioral) features are
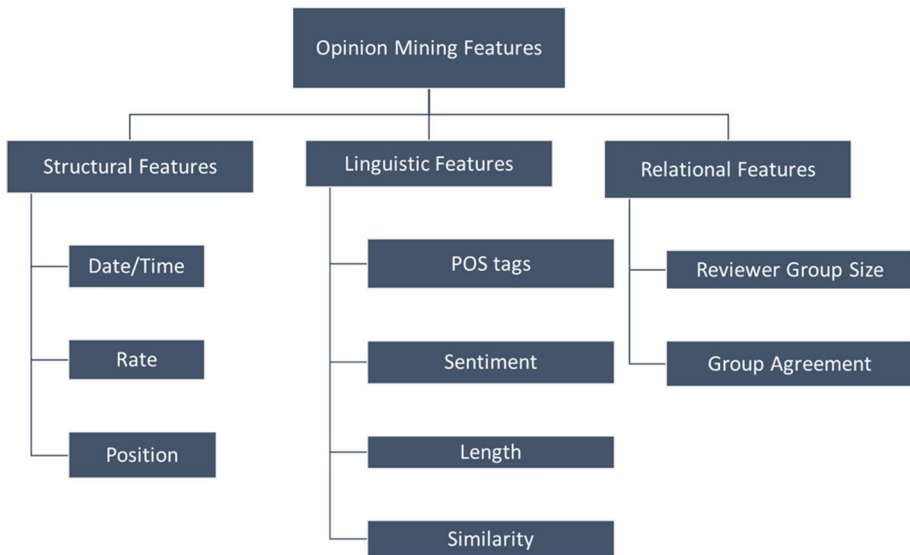
Fig. 3 Opinion mining features

those features which imply the behavior of reviews, reviewers, group of reviewers or targets. They also play a critical role among opinion mining features.

It has been proved that using multiple types of features can improve the accuracy of opinion mining applications (Zhang et al. 2016a). In opinion mining research, selecting the most effective features and combining them to achieve the best performance is a big challenge.

3.2 Different feature selection methods in opinion mining

Most feature selection methods in information retrieval or text classification can be adopted for the task of opinion classification.

The popular and widely used methods in opinion mining applications include Term Frequency Inverse Document Frequency(TF-IDF) (Zheng et al. 2009; Li and Tsai 2013; Moraes et al. 2013; Khairnar and Kinikar 2013; Basari et al. 2013; Martineau et al. 2009), Point-wise Mutual Information (PMI) (Cover and Thomas 2012; Khairnar and Kinikar 2013), Chi-Square (Khairnar and Kinikar 2013; Fan and Chang 2011; Hagenau et al. 2013), Information Gain (IG) (Moraes et al. 2013), Best Matching 25 (BM25) (Vechtomova 2010), Uniformity (Uni) (Li and Tsai 2013), Inverted Conformity Frequency (ICF) (Li and Tsai 2013) and Latent Dirichlet Allocation(LDA). Figure 4 illustrates the usage of different feature selection methods in our selected papers.

One of the most popular algorithms that is widely used in practice for information retrieval tasks as well as opinion mining classification is TF-IDF. The strength of TF-IDF is that it computes the similarity of two documents by extracting most descriptive terms; however, it does not capture co-occurrence in reviews and semantics of opinions. Therefore, the accuracy is low for large datasets. It makes no use of semantic similarities
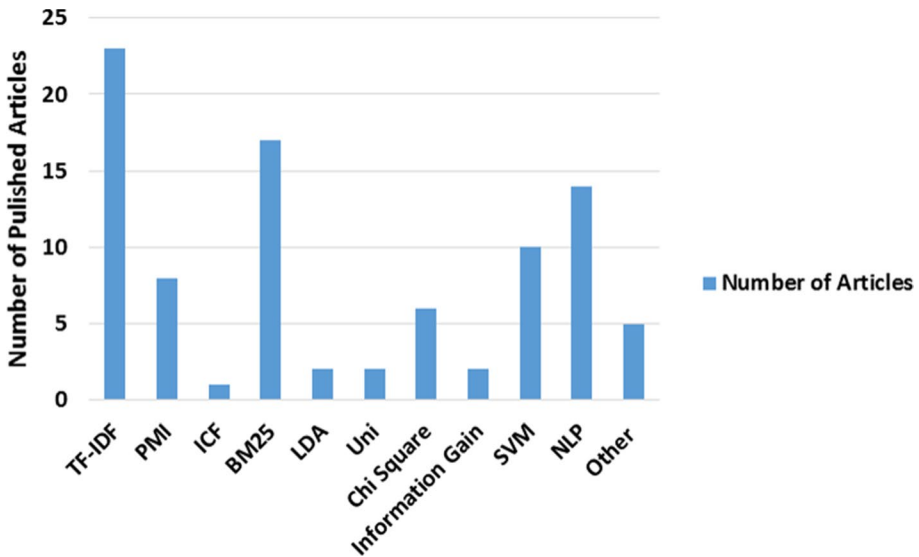
Fig. 4 Number of published papers for each feature selection method

between words. Researchers try to increase the accuracy for large datasets by combining TF-IDF with other feature selection methods such as Uni and ICF (Li and Tsai 2013).

Uni is calculated using Eq. 1.

$$Uni(t_i) = max_j \left[\frac{d_{ij}}{t_{ij}} \times \frac{d_{ij}}{\sum_{j=1}^{k} d_{ij}} \right], k = 2, \tag{1}$$

where t_i is a term i, d_{ij} is the number of documents in which term i appears in category j, t_{ij} indicates the number of times where term i appears in category j. A larger value means that the term is more distinctive in a specific category. Li and Tsai (2013) used Uni > 0.2 as a threshold value for feature selection.

ICF indicates which term should appear frequently in a specific category instead of others. Eq. 2. calculates ICF value:

$$ICF(t_i) = \sum_{j=1}^{k} P_{ij} log_2 P_{ij}, P_{ij} = \frac{d_{ij}}{|j|}, \tag{2}$$

where t_i indicates a term and j indicates a category, d_{ij} is the number of documents which contain term t_i in category j, $|j|$ is the total number of documents contained in category j. The smaller ICF value of a term indicates it appears more frequently in specific categories. Li and Tsai (2013) use $ICF < log(2)$.

Many models try to overcome the limitations in TF-IDF for capturing low-dimensional, latent representations (Yatsko 2013). PMI is good for collection extraction and shows significant results with normalized PMI (Bouma 2009) and the variations of PMI by incorporating significant co-occurrence (Damani 2013) as semantic of opinions are captured. However, Chi-square is better than PMI to represent subjective and vague opinions with its normalized value. This is scale dependent and has great impact on feature selection

with continuous variables as many relevant features may be removed. It is also only applicable for categorical or nominal data, as the relationships between terms are independent. Fan and Chang (2011) used IG, PMI, and Chi-square for feature selection. Three feature selection methods were significantly co-related to each other. When many features were highly redundant to each other, IG reduces the redundancy between features while selecting appropriate features for text categorization (Lee and Lee 2006). Then there is a need of ranking features. Usually a feature with high information gain should be ranked higher than other features because it has stronger power in classifying the data.

BM25 was also proposed for feature ranking purposes. It is a ranking function that was used to rank matched documents according to their relevance to a given query (Robertson et al. 2009; Whissell and Clarke 2011; Vechtomova 2010; Esparza et al. 2012; Luo et al. 2012). Li and Tsai (2013) used BM25 for feature selection in their model with Uni and ICF. Paltoglou and Thelwall (2010) used BM25 and showed that the performance was significantly improved. This makes BM25 very popular because of its efficiency. It performs well in ad-hoc retrieval. The main disadvantage of this model is hard to overcome since it is full of hacks and common problems such as polysemy, synonymy and information overload. Alharbi et al. (2017a) proposed LDA with clustering algorithms to overcome these problems.

Recently, LDA was integrated with clustering algorithms for feature selection (Alharbi et al. 2017a). LDA was used to cluster similar documents to reduce the impact of frequent subjects in the collection during LDA topic extraction. Therefore subjects that are less frequent are not overshadowed by the highly frequent ones. In the clustering stage, a cluster generates a set of semantically related group of words that address one super subject and are highly correlated and redundant. This research showed that the combination of LDA with clustering algorithms can improve the performance of LDA and BM25.

Apart from the above, Apriori Algorithms (Moore et al. 1997) and Latent Semantic Indexing (LSI) (Hofmann 2017) are already used in feature selection. There are other statistical approaches available which are still open for researchers for further research. These include Hidden Markov Model(HMM) (Fine et al. 1998), Genetic Algorithm (GA) (Kristiyanti and Wahyudi 2017), and combinations with other methods such as TF-IDF and HMM. The combination of several feature selection methods can increase the accuracy of opinion mining classifiers and deserves further exploration.

4 Representation of discovered knowledge for opinion mining

Knowledge discovery is the process of discovering hidden knowledge from customer review datasets, which can be useful for solving real world problems. In this section we focus on different representations and discovery methods for hidden knowledge in opinion mining.

4.1 Knowledge representation

Discovered knowledge from opinions can be represented in different forms, such as terms, patterns, phrases, concepts, rules, relations, or ontologies. Some of them are combined with other to increase the accuracy. Most existing popular text mining and classification methods have adopted term-based representation (Zheng et al. 2009; Li and Tsai 2013; Moraes et al. 2013; Khairnar and Kinikar 2013; Basari et al. 2013). They have all suffered

from the problems of polysemy and synonymy (Li et al. 2015b). Li et al. (2015b) have proved that pattern-based methods perform better than term-based ones, as the patterns usually carry more semantic meanings, and semantic relations, and more context information and association rules than single terms. Patterns can be formed by a single word or multiple terms which frequently co-occur in textual data (Shinde and Gill 2014). Patterns can be defined in many forms, such as frequent patterns (Ghorashi et al. 2012), closed patterns (Pasquier et al. 1999) and top-k patterns (Han et al. 2002).

n-grams (or phrases) are more discriminative and carry more semantics than terms in the representation of knowledge (Ifrim et al. 2008). An n-gram is a set of contiguous *n* items with the corresponding frequency (Sun et al. 2017). For opinion mining, unigram and bigram are widely used (Sun et al. 2017). Sharma and Raman (2003) proposed a phrase-based text representation approach that uses rule-based techniques. Extraction of key-phrases from text documents is based on a process of partial parsing. By making the indexing terms more meaningful through reduction of the ambiguity in words considered in isolation, improvement in retrieval effectiveness is sought.

POS tags can be used to represent knowledge as well. Available POS tags, such as adjective or noun, are quite helpful, because opinion words are usually adjectives and opinion targets (i.e., entities and aspects) are nouns or the combination of nouns. Jadav and Vaghela (2016) used POS tagging to represent knowledge and calculated sentiment score with the help of the SentiWordNet dictionary. Recently, a combination of *n*-gram and POS was used for representing knowledge more accurately in opinion mining (Afzaal et al. 2016).

Most recently, some novel methods were applied to knowledge representation, such as ontology-based (Penalver-Martinez et al. 2014), semantic aspect-based (Afzaal et al. 2016; Samha et al. 2014; Rana and Cheah 2017), rule-based (Rana and Cheah 2017; Poria et al. 2014), lexicon sentiment (Taboada et al. 2011; Kang et al. 2012; Cambria et al. 2020) and fuzzy concepts (Zadeh 1996; Quan et al. 2004). These advanced techniques make opinion mining systems easier and more effective in representing discovered knowledge from larger and more complex review datasets.
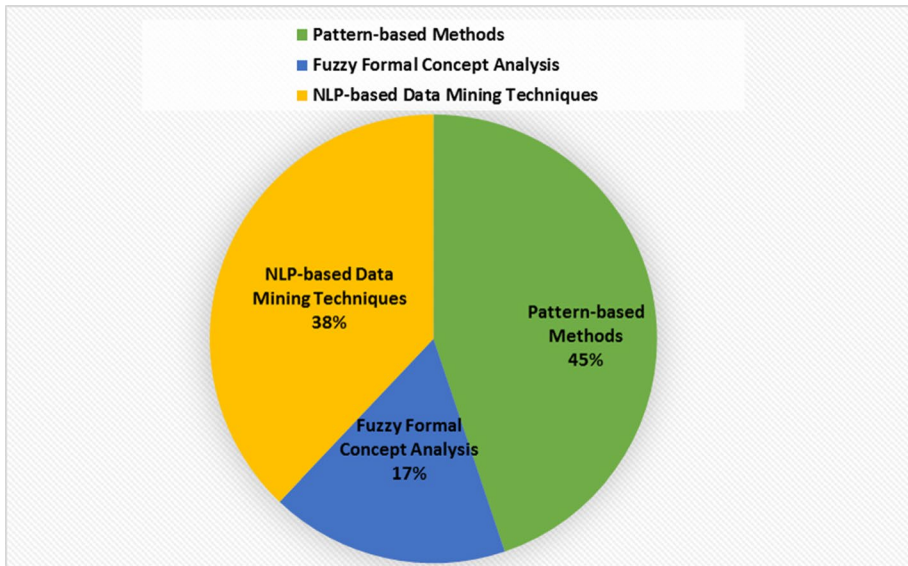
4.2 Knowledge discovery

Different knowledge representations need different suitable discovery methods to achieve the best performance. Fig. 5 shows the percentages of different knowledge discovery methods used in recent published papers according to the above mentioned knowledge representation types.

Table 1 gives details of knowledge discovery methods which were used by each research. The most widely used knowledge discovery method is pattern-based methods which featured in 25 papers out of 120, while NLP based data mining techniques and Fuzzy Formal Concept Analysis (FFCA) feature in two other large portions of published researches. Therefore, we discuss the pattern based, NLP based data mining techniques and FFCA methods, respectively.

4.2.1 Pattern-based methods in opinion mining

In opinion mining, pattern mining can discover sequencing terms that frequently co-occur in a customer review, and such set of terms can represent the knowledge in reviews effectively. Frequent patterns and closed patterns are frequently employed to represent knowledge and trends in a dataset (Li et al. 2015b; Zhong et al. 2012). These trends can be used

Fig. 5 Number of published papers for knowledge discovery method

to make the decisions in a business as well as for customers. Most of the existing pattern-based opinion mining models have used unsupervised approaches (Hu and Liu 2004; Gao et al. 2015; Li et al. 2011). Hu and Liu (2004) presented a model for opinion mining which is based on association rule mining. Its features are noun or noun phrases. Apriori algorithm is used to generate frequent itemsets. After pruning, it removes the features which are not genuine. This algorithm is effective in discovering frequent features. However, when dataset size increases it is required to take many database scans which leads to an increase of the computational time. In the process of generating item sets, the algorithm did not consider the sequence of the items which is vital for discovering frequent patterns. Ghorashi et al. (2012) applied different pattern mining algorithm to enhance the accuracy of Apriori algorithm. They have applied the H-Mine algorithm, a frequent pattern mining algorithm which considers multiple occurrences at the same time for a large dataset. Both of the above algorithms return a large number of patterns, because if a pattern is frequent, each of its child patterns also becomes frequent, which leads to a high computational time (Gao et al. 2015).

Selecting reliable patterns is vital, which enhances the efficiency of generating the frequent itemsets without losing any item (Chee et al. 2019). In order to enhance the efficiency of pattern identification, researchers proposed several techniques: maximum frequent pattern mining (Bayardo 1998), closed frequent pattern mining (Pasquier et al. 1999), and top-k closed pattern mining (Han et al. 2002). Closed patterns (Pasquier et al. 1999) were proposed for handling a large number of frequent patterns. A closed pattern is also a frequent pattern, but it is not included in another sequential pattern that has the exact same support. Therefore, the computational time for finding closed patterns may be reduced and it can also largely reduce the number of frequent patterns. Gao et al. (2015) also introduced a new algorithm called Maximum-matched Pattern Based Topic Modelling (MPBTM) for the above identified limitations of frequent patterns. These advanced pattern mining techniques were already used in text mining but still lack usage in opinion mining.

Table 1 Opinion mining: knowledge discovery methods

References	Year	Knowledge discovery method	Dataset
Chen and Tseng (2011)	2011	SVM	E-commerce web sites
Fan and Chang (2011)	2011	Chi-square test	eBay
Nikfarjam and Gonzalez (2011)	2011	Pattern mining	Drug data
Li et al. (2011)	2011	Pattern mining	Reuters corpus volume 1(RCV1)
Saleh et al. (2011)	2011	TF-IDF	Dataset 1 (Pang et al. 2002)
Yu et al. (2011)	2011	Semantic learning	Online reviews
Eirinaki et al. (2012)	2012	High adjective count algorithm	Dataset 1 (Hu and Liu 2004)
Kang et al. (2012)	2012	Senti-lexicon	Movie review
Nadali and Murad (2012)	2012	Fuzzy logic	Dataset 1 (Hu and Liu 2004)
Zhong et al. (2012)	2012	Pattern mining	RCV1
Ahmad and Doja (2013)	2013	Frequent pattern growth	eBay
Basari et al. (2013)	2013	TF-IDF	Movie review
Li and Tsai (2013)	2013	Fuzzy formal concepts	Ebook, movie review and reauters-21578
Moraes et al. (2013)	2013	TF-IDF	Ebook and movie review
Keshtkar and Inkpen (2013)	2013	POS	Public datasets
Khairnar and Kinikar (2013)	2013	SVM	Twitter data
Wu and Li (2013)	2013	PMI–TFIDF	Public datasets
Chenlo and Losada (2014)	2014	Unigram-bigram	Movie reviews
Bing and Chan (2014)	2014	Fuzzy logic	Twitter
Dalal and Zaveri (2014)	2014	POS	Movie review
Dragoni et al. (2014)	2014	Fuzzy logic	Blitzer dataset
Haque et al. (2014)	2014	Pattern mining	ADR comments
Pimpalkar et al. (2014)	2014	Fuzzy logic	Twitter, Facebook
Priyadarshi et al. (2014)	2014	Neural network	Amazon smartphone review dataset
Sarkar et al. (2014)	2014	NB	Reuters
Sharma et al. (2014)	2014	TF-IDF	Movie review
Quan and Ren (2014)	2014	Fuzzy logic	Facebook

Table 1 (continued)

References	Year	Knowledge discovery method	Dataset
Bravo-Marquez et al. (2014)	2014	Unigrams	Twitter
Amarouche et al. (2015)	2015	Fuzzy logic	Amazon
Bhardwaj et al. (2015)	2015	–	Stock market data
Gao et al. (2015)	2015	Pattern mining	RCV1
Prakash et al. (2015)	2015	Pattern mining	RCV1
Zimmermann et al. (2015)	2015	Fuzzy logic	Dataset 1 (Hu and Liu 2004) and Dataset 2 (Yu et al. 2011)
Xu et al. (2015)	2015	Transfer learning	NTCIR-7 MOAT
Afzaal et al. (2016)	2016	n-gram and POS	Tourist reviews
Ali et al. (2016)	2016	Fuzzy ontology	Online reviews
Darrab and Eregenc (2016)	2016	Frequent pattern mining	Synthetic dataset and four real world datasets
Jadav and Vaghela (2016)	2016	SVM	Twitter and Amazon data
Jinturkar and Gotmare (2016)	2016	POS	Twitter
Jiménez-Zafra et al. (2016)	2016	Senti-lexicon	Movie review
Sadidpour et al. (2016)	2016	Polarity patterns	Political news
Mamgain et al. (2016)	2016	Pattern mining	Twitter
Panchendrarajan et al. (2016)	2016	Pattern mining	Yelp dataset
Pansare (2016)	2016	Frequent pattern mining	Online reviews
Vinodhini and Chandrasekaran (2016)	2016	Pattern mining	Amazon
Zhang et al. (2016a)	2016	Fuzzy rules	Movie review
Poria et al. (2016b)	2016	Lexicon	Semeval 2014
Angelpreethi and Kumar (2017)	2017	Frequent pattern mining	Web options
Bilici and Saygın (2017)	2017	Probabilistic model	Online reviews
Boudia et al. (2017)	2017	Senti-lexicon	Twitter data
Howells and Ertugan (2017)	2017	Frequent patterns	RCV1, Ontology LCSH,R8
Bhardwaj et al. (2015)	2017	Fuzzy rules	Twitter

Table 1 (continued)

References	Year	Knowledge discovery method	Dataset
Catal and Nangir (2017)	2017	Statistically weighted	Movie reviews
Gu et al. (2017)	2017	CNN	Twitter
Hu et al. (2017)	2017	Fuzzy rules	TripAdvisor
Phu et al. (2017)	2017	SVM	Movie review
Ravi et al. (2017)	2017	Fuzzy logic	Bank datasets
Sadhana et al. (2017)	2017	Relation graph	Amazon
Tsirakis et al. (2017)	2017	–	Twitter
Vinodhini and Chandrasekaran (2017)	2017	Sampling method	IBD1 and IBD2
Vo et al. (2017)	2017	Pattern mining	Five real datasets
Jing et al. (2018)	2018	Chi-squre	JD.com
Deshmukh and Tripathy (2018)	2018	Sentiment-lexicon	Amazon
Kang et al. (2018)	2018	Frequent pattern	(Hu and Liu 2004)
Van Hee et al. (2018)	2018	SVM	Twitter
Asgarian et al. (2018)	2018	Sentiment-lexicon	Persian reviews
Saleena et al. (2018)	2018	Novel classification algorithm	Twitter dataset
Ma et al. (2018)	2018	Lexicon	SentiHood
Azizan et al. (2019)	2019	Sentiment-lexicon	Movie review tweets
Jaman and Abdulrohman (2019)	2019	TF-IDF	Twitter
Mishra et al. (2019)	2019	TF-IDF	Flickr photos
Khattak et al. (2020)	2020	Sentiment-lexicon	Drugs, electronics and mobile phone reviews
Park et al. (2020)	2020	Word embedding	Hotel reviews
Jain et al. (2020)	2020	Sentiment-lexicon	Blitzer, acIIMDb, polarity and subjective dataset.
You et al. (2020)	2020	TF, TF-IDF, DROP3 and IB3	HCR, OMD and STS data set.
Li et al. (2020)	2020	Syntactical rules	SemEval-2014 and SemEval-2015 data set.
Da'u et al. (2020a)	2020	NLP	SemEval data set.

Table 1 (continued)

References	Year	Knowledge discovery method	Dataset
Fang et al. (2020)	2020	Feature-based implicit opinion pattern	Car forum reviews

Using patterns can capture a higher level of knowledge (Fang et al. 2020). However, the uncertainty in opinions cannot be dealt with well in some cases. For example, 'camera quality is not bad' is often classified by the machine as a negative review if 'camera bad' is a frequent pattern. Therefore, researchers need to seek other solutions to handle uncertainties more effectively for opinion mining.

4.2.2 NLP-based data mining techniques in opinion mining

NLP is a theoretically motivated range of computational techniques for analysing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications (Liddy 2001).

POS tagging and parsing are techniques that analyse lexical and syntactic information. POS tagging is used to determine the corresponding POS tag for each word in a review sentence, whereas POS parsing extracts the syntax (i.e. the way words are arranged together, and the relationship between them) from the review sentence. Comparing to POS tagging, POS parsing provides richer structured information. As there are similarities and relevances among word segmentation, POS tagging and parsing, some approaches are proposed to deal with these tasks simultaneously (Sun et al. 2017; Azizan et al. 2019; Mishra et al. 2019).

Aspect identification is an important task for knowledge discovery. Recent research work focuses on the POS tagger based NLP technique for aspect identification in reviews. Hu and Liu (2004) proposed a model to extract frequently used words, which used an Apriori algorithm. It firstly extracts all the frequent aspects and then finds the opinion words associated with them. The results showed that noun and noun phrases represent any aspects in a review. However, this work reported that not all the aspects are frequent. In 2005, Liu et al. (2005) used language pattern mining to identify explicit and implicit aspects from negative and positive reviews. The first step involves finding positives or negatives by using Senti-WordNet. Finally, a summary was generated using overall negative and positive aspects. Bafna and Toshniwal (2013) extracted all frequent nouns as aspects and eliminated the nouns that did not represent aspects using a probabilistic power equation. After all the nearest aspects were grouped, adjectives were extracted as opinion words. Poria et al. (2016b) developed a model using LDA, called Sentic LDA, which exploits common-sense reasoning to shift LDA clustering from a syntactic to a semantic level. Sentic LDA leverages on the semantics associated with words and multi-word expressions to improve clustering and, hence, outperform state-of-the-art techniques for aspect extraction. A Hybrid Network for Targeted Aspect-Based Sentiment Analysis model was proposed using an extension of Long Short Term Memory (LSTM), termed Sentic LSTM (Ma et al. 2018). The extended LSTM cell includes a separate output gate that interpolates the token-level memory and the concept-level input. In addition, they propose an extension of Sentic LSTM by creating a hybrid of the LSTM and a recurrent additive network that simulates sentic patterns.

Samha et al. (2014) proposed a model to represent aspect-based summary. The first step was to extract entities that consist of aspects and opinion extraction. Extracted entities were grouped based on synonyms using WordNet dictionary. POS was used to extract aspects and opinions using frequent tags. After aspects were grouped, aspect selection was done using the strength of the sentences; this was calculated based on the weights assigned to tags (adjective, verb or adverb). Finally, based on the weight value, a summary was

generated. Fernandes and GL (2017) developed a rule based classifier. It applied POS tagging rules and developed a classifier using SVM and NB. Once the models are successfully trained, random reviews are collected from web applications or mobile applications and classified according to the predefined classes using NB and SVM.

A feature-based opinion mining system was implemented using ontology (Penalver-Martinez et al. 2014). This model applied NLP techniques for data preprocessing. A domain ontology is used in order to extract the features included in the opinions expressed by users. It also used an existing domain ontology (the Movie Ontology, which is available at http://www.movieontology.org.).

Lexicon based unsupervised sentiment allocation techniques are used to find the opinions from categorized aspects (Kumarasiri and Farook 2018). Lexicon approaches determine the sentiment score of text according to sentiment lexicons in an unsupervised manner. A lexicon is a dictionary of sentiment words and phrases with their polarities and strengths. For each document or sentence, the corresponding polarity is determined by a sentiment score, which is computed by the words or phrases occurring and their sentiment polarities and strengths. However, an unsupervised method always outperformed a supervised method during sentiment allocation. An effective Urdu sentiment analyzer (Mukhtar and Khan 2019) applies rules and make use of this new lexicon to perform Urdu sentiment analysis by classifying sentences as positive, negative or neutral . A model was proposed using lexicon word semantics based on Expected Likelihood Estimate Smoothed Odds Ratio (ELESOR) that were then incorporated with a supervised machine learning based model selection approach (Khan et al. 2017a) .

Feature engineering is very important in machine learning as well as NLP. *n*-gram is the most fundamental feature used in NLP. Syntactic features contain POS tags and syntactic information is used to represent knowledge (Joshi and Penstein-Rosé 2009; Chenlo and Losada 2014; Bravo-Marquez et al. 2014; Cambria et al. 2013a). Semantic features are conjunctions which indicate negation, intensification, and diminution. Negation is important for opinion mining as it reverses the sentiment orientation. Intensification and diminution increase and decrease the strength of sentiments, respectively, and are also useful for opinion mining (Taboada et al. 2011; Cambria et al. 2020). AI and Semantic Web techniques can be used in knowledge discovery. Sentic Computing is a new paradigm for the affective analysis of natural language text, to semantically analyze opinions and exploit different web ontologies to encode the results in a semantically aware format (Cambria et al. 2012). Cambria et al. (2020) integrate top-down and bottom-up learning via an ensemble of symbolic and subsymbolic AI tools, which apply to the interesting problem of polarity detection in a text.

Existing natural language models (Samha et al. 2014; Ma et al. 2018; Cambria et al. 2020; Poria et al. 2014; Morency et al. 2011; Poria et al. 2015) were developed with consideration for the semantic aspects of the reviews. However, ways of coping with uncertainty of opinions are lacking in available models.

4.2.3 Fuzzy formal concept analysis in opinion mining

Fuzzy concepts have been used in opinion mining recently (Li and Tsai 2013; Subhashini et al. 2018) as there are a lot of uncertainties in opinions or reviews. As the methodology is quite different from other methods, in this section, we provide more details for FFCA in order to understand its advantages for opinion mining.

Fuzzy Logic is an approach to handle the concept of partial truth, where the truth value can be any real number between 0 (i.e. completely false) and 1 (i.e. completely true) inclusive (Novák et al. 2012). Fuzzy Logic can be used to extract knowledge from text data where a lot of uncertainties need to be coped with. FFCA (Li and Tsai 2013; Quan et al. 2004) is a theory which combines fuzzy logic and Formal Concept Analysis (FCA) to represent the uncertainties in data. FCA is a technique based on lattice theory (Zadeh 1996; Quan et al. 2004). A formal concept is used to define the relationships between objects and attributes in a domain. Representing uncertainty is very important in opinion mining when describing the relationships. FCA is unable to represent uncertainty of information, but FFCA is handy for that; it is a solution for representing vague information in a knowledge base. Concepts are formulated using attributes (intent) and objects (extent) in a relation which can represent more semantic information.

Definition 1 A fuzzy formal context is a triple context $K = (G, M, I)$, where G is set of objects, M is set of attributes, and I is a fuzzy set domain $G \times M$ and there is a function $\mu : I \rightarrow [0, 1]$ for all relations $(g, m) \in I$ (Zadeh 1996; Quan et al. 2004).

The relationships between the object and the concept should be the intersection of the relationships among them. The object and an attribute are represented as a membership value in a fuzzy formal context. Then the intersection of these membership values should be the minimum of these membership values, according to fuzzy theory (Zadeh 1996). Therefore, the fuzzy formal concept generated from fuzzy formal context can be defined as follows.

Definition 2 Given a fuzzy formal context $K = (G, M, I)$ and a confident threshold T, $A^* = \{m \in M | \forall g \in A : \mu(g, m) \geq T\}$ for $A \subseteq G$ and $B^* = \{g \in G | \forall m \in B : \mu(g, m) \geq T\}$ for $B \subseteq M$ (Zadeh 1996; Quan et al. 2004).

Definition 3 A fuzzy formal concept (fuzzy concept) of fuzzy formal context (G, M, I) for a confident threshold T is a pair (A, B) where $A \subseteq G, B \subseteq M, A^* = B$ and $B^* = A$. For each object $g \in A$ there is a membership μ_g defined as in Eq. (3) (Zadeh 1996; Quan et al. 2004).

$$\mu_g = min_{m \in B}\mu(g, m) \tag{3}$$

where $\mu(g, m)$ is the fuzzy value between object g and attribute m, which is defined on I.

Definition 4 Let (A_1, B_1) and (A_2, B_2) be two fuzzy concepts of a formal context (G, M, I). (A_1, B_1) is the subconcept of (A_2, B_2), denoted as $(A_1, B_1) \leq (A_2, B_2)$, if and only if $A_1 \subseteq A_2$ and $B_2 \subseteq B_1$. Equivalently, (A_2, B_2) is the super concept of (A_1, B_1) (Zadeh 1996; Quan et al. 2004).

Quan et al. (2004) applied FFCA first to the knowledge discovery process. They used ebook reviews for their experiments. FFCA outperformed other state-of-the-art models. These knowledge discovery methods are developed to cope with uncertainties in opinions, but the limitation is that the relationships between attributes are not considered well. Therefore, the concept lattice rejected less semantic information and uncertainty was high. Hence, accuracy cannot be very high. Thus, in consideration of the uncertainty of semantic information of opinions, how to build the relationships between the attributes in the training dataset became challenging. The limitation is that there are fewer relationships

between attributes of the training dataset. One of the latest and most successful research works in fuzzy logic for opinion mining was conducted by Li and Tsai (2013). This work developed a fuzzy opinion mining classification using FFCA to represent the uncertainties in opinions. Feature selection indexes of TF-IDF, ICF and Uni were applied. It also generated fuzzy formal concepts using a term-document matrix, while reducing the dataset by applying a threshold value for each feature index. A concept lattice was generated using the intent and extent of concepts. Normalized TF-IDF was used as the degree of membership for each term in relation to objects. To retrieve the relationships between concepts and categories, the fuzzy composition operation was applied.

To handle uncertainties, it is important to clearly identify the boundary region between positive and negative. The main problem still remains. In the future, researchers can think about handling uncertainties of opinion using the combination of these knowledge discovery methods.

5 Opinion classification

To design and develop an appropriate classifier for sentiments and opinions, classification is the most critical step of the opinion mining process. Therefore, selecting appropriate techniques for opinion classification is an important task in the process. Based on the extensive literature review, sentiment classifiers can be divided into supervised, semi-supervised or unsupervised classifiers. Recently deep learning models (Poria et al. 2016a; Wang et al. 2018) have been applied in opinion mining in supervised, semi-supervised or unsupervised ways. However, researchers have pointed out the disadvantages of existing deep learning models (Li and Chen 2014; Vateekul and Koomsubha 2016) and attention models (Shin et al. 2016; He et al. 2018; Lei et al. 2018; Wang et al. 2016). The main disadvantage is that fixed-length context vector design is incapable of remembering longer sequences. Often it has forgotten the earlier parts of the sequence once it has processed the entire sequence. The attention mechanism was born to resolve this problem.

5.1 Supervised classifiers

Supervised learning is widely used in opinion mining. This kind of learning method depends on labelled data (a training set). There are many kinds of algorithms in literature in relation to supervised learning. Researchers started to work on sentiments and opinion classification with NB and SVM classifiers; then, Artificial Neural Network (ANN) became a significant solution for opinion classification; but they all faced the problem of dealing with uncertainties. Therefore, a fuzzy logic based method for handling uncertainties was proposed (Li and Tsai 2013). The experimental results showed that the method is better than SVM, NB and ANN. Most recently, with extensive experiments, deep learning methods have shown significant improvement, since deep learning models can learn embedded semantic representations of text and provide more complete and comprehensive input features. We will discuss deep learning methods in another sub-section.

5.1.1 NB

The NB classifier is one of the most commonly used classifiers in opinion mining. It classifies a dataset into labels (classes) using calculated probability based on Bayes theorem.

Kang et al. (2012) improved the existing NB classifier for opinion mining to solve the problem of uncertainty between negative and positive classes using the restaurant review dataset. When classifying a new review using the supervised learning mechanism, there is no clear boundary between two classes, which is called conflicts of classes. Their method was capable of reducing the uncertainty to 3.6 %. The modified NB classifier effectively classified opinions when compared to both NB and SVM. Unigrams and bigrams were effective features for opinion mining in their model. This research tried to solve the problem of uncertainty using a supervised method.

Bilal et al. (2016) compared the efficiencies of three techniques, namely NB, decision tree, and nearest neighbor, for classifying Urdu and English opinions in a blog. Their results show that NB has a better performance than the two other techniques. Xia et al. (2015) developed an NB classifier using opinion level features. Observations show that intra-opinion features play an essential role in word sentiment polarity. They help in resolving the polarity of most sentiment words. A Bayesian model that uses both intra-opinion features and inter-opinion features performs better than the term-based Bayesian model and opinion-based Bayesian model that uses only intra-opinion features.

However, there is an obvious need to identify the relationships between terms, and the problem of uncertainty remains the same.

5.1.2 SVM

The SVM approach is the best known supervised classifier for opinion mining (Xu et al. 2015; Phu et al. 2017; Saleh et al. 2011; Moraes et al. 2013; Prakash et al. 2015; Jadav and Vaghela 2016; Jaman and Abdulrohman 2019). SVM uses a kernel function to map an input feature space into a new space where the classes are linearly separable (Vinodhini and Chandrasekaran 2016) and Jadav and Vaghela (2016) compared NB, SVM and optimized SVM on the movie review dataset. Their results indicated that optimized SVM is more accurate than the other two classifiers since it changes the values of kernel parameters. However, it is difficult for this model to use the probabilities of relevance for given classes because of the complex relationship between terms. Therefore, identification of the boundary region is a problem for this model. An ensemble classification system (Saleena et al. 2018) has been developed using a new algorithm which integrates the features of NB, Random Forest classifier and SVM to improve the performance and accuracy of sentiment classification.

5.1.3 ANN

ANN has been applied to opinion classification by several researchers recently (Chen et al. 2011; Moraes et al. 2013; Vinodhini and Chandrasekaran 2016). ANN depends on three aspects, the input and activation function of the unit, the network architecture, and the weight of each input connection. ANN extracts features from a linear combination of the input data, and then models outputs as a non-linear 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: the input layer, output layer and hidden layer.

Vinodhini and Chandrasekaran (2016) implemented a neural network based model for sentiment classification. In their model, features were selected by using natural language techniques such as unigrams, bigrams and trigrams. Thereafter, features were reduced by

using principal component analysis, and feature vectors were produced with unigram and bigram. Based on feature vectors, classification models were implemented. Back propagation neural network was implemented using patterns as the inputs. The architecture was finalized with three layers with 4, 7 and 2 nodes followed by each layer. The logistic function was used as an activation function. Existing ANN models can cope with only trained data and cannot deal with relationship between terms.

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

Neural network requires much more data for the training process and the computation is more expensive than other methods.

5.1.4 Fuzzy logic

For the process of classification, fuzzy composition has been applied recently. This has shown higher accuracy than other existing models (Li and Tsai 2013; Subhashini et al. 2018) in opinion mining. Rule-based fuzzy logic systems were built by several researchers (Afzaal et al. 2016; Bing and Chan 2014). Nadali and Murad (2012) presented a rule-based fuzzy model using adjectives, verbs, adverbs and nouns. Opinion words were identified using a linguistic parser. A triangular membership function was designed to find the membership value. Rules were defined using adjectives, verbs, adverbs and nouns as well. In the defuzzification process Mamdanis defuzzifier was used to convert the fuzzy values into crisp values. This is a basic fuzzy system in which accuracy depends on the number of defined rules. Haque et al. (2014) proposed a fuzzy model to generate sentiment values according to products or service interest, using SentiWordNet, which is a lexical resource for sentiment analysis. With the enhancement of fuzzy logic, researchers attempt to evaluate the fuzzy logic classifiers in opinion mining research (Li and Tsai 2013; Nadali and Murad 2012; Quan et al. 2004; Dalal and Zaveri 2014; Boudia et al. 2017; Pimpalkar et al. 2014; Bing and Chan 2014; Haque et al. 2014; Khattak et al. 2020).

Fuzzy composition is a new approach to procuring classifications for opinion mining. Recently neuro-fuzzy models have been applied in classification (Xing et al. 2017). The reasoning mechanism is based on fuzzy composition (Li and Tsai 2013). The standard composition of two fuzzy relations $P(x, y)$ and $Q(y, z)$ is represented by $P(x, y) \circ Q(y, z)$ which generates $R(x, z)$ on $P \times Q$, where $R(x, z)$ is calculated as in Eq. (4) for $x \in X$ and $z \in Z$ (Li and Tsai 2013).

$$R(x, z) = [P \circ Q](x, z) = max_{y \in Y} min[P(x, y), Q(y, z)] \tag{4}$$

Li and Tsai (2013) used the fuzzy composition to retrieve the relation between concepts and categories. The relation of concept and category ($R_{(C-Catg)}$), was retrieved by using the fuzzy composition on concept-term ($R_{(C-T)}$) and term-category($R_{(T-Catg)}$) relation as in Eq. (5) (Li and Tsai 2013).

$$R_{(C-Catg)} = R_{(C-T)} \circ R_{(T-Catg)} \qquad (5)$$

The relation R_{T-Catg} is also obtained by using the fuzzy composition on the term-review relation R_{T-R} and review-category relation R_{R-Catg} as in Eq.(6) (Li and Tsai 2013), where R_{C-T} was calculated using the generated fuzzy values.

$$R_{T-Catg} = R_{T-R} \circ R_{R-Catg} \qquad (6)$$

However, this model does not clearly indicate how to represent uncertainties in concepts.

To develop a supervised model, users need to provide a large training dataset in order to distinguish different opinion classes accurately. However, it is very hard to collect a large amount of labelled data. Unsupervised models discover the features first and do the clustering, which can adapt to different real time scenarios. The computational time is less and it is easy to collect unlabelled data. Therefore, unsupervised models for opinion mining are also welcomed.

5.2 Unsupervised classifiers

Researchers developed unsupervised models to deal with unlabelled datasets using clusters based on similarities (Turney 2002; Quan and Ren 2014; Jiménez-Zafra et al. 2016; Pang et al. 2008; Recupero et al. 2015). The first unsupervised model for opinion mining was introduced by Turney (2002). The model extracted phases containing adjectives and adverbs. Then, PMI was used to calculate the semantic orientation. Based on the average value of semantic orientation, they classified reviews as recommended or not recommended. This model did not take into consideration the links between the phases and links between phases and reviews. Therefore, the accuracy is low, and it still ignores the representation of uncertain information in opinions.

Quan and Ren (2014) presented an unsupervised algorithm for feature-oriented opinion mining. They used similarity distance of domain vectors, which is formulated using association values between features and the domain. They proposed a novel algorithm, PMI–TFIDF, which used both the association of features and domain entities. This model effectively classifies opinions into three classes: positive, negative, and neutral regions. The model was unable to classify subjective and vague opinions. Most existing models (supervised and unsupervised) are not satisfactory since they cannot cope with uncertainties of opinions. These models focused only on words and their accuracy was low due to the ignored relationship between attributes in training samples. Sentilo (Recupero et al. 2015) combines NLP techniques with knowledge representation using a combined sentiment score in a unsupervised fashion. Relying on a novel lexical resource, it can extract frame-based semantic relations between topics and subtopic. Most recently, a novel query-based unsupervised learning model to represent the implicit relationships in the short text from social media (Albishre et al. 2020) has been proposed to cope with the sparsity problem in social media data analysis, as it focuses on uncertainty reduction from the driven tweets.

Lexicon approaches determine the sentiment score of text according to sentiment lexicons in an unsupervised manner for the classification (Kumarasiri and Farook 2018). Polarity is determined by using a sentiment score. NLP was applied to the classification, which is done based on segmentation units (Li et al. 2015a). Most of the NLP techniques are applied for unsupervised classification (Taboada et al. 2011; Lin and He 2009). LDA is designed to do the classification for annotated data with reduce dependence. Li et al.

([2010a](#)) proposed a Dependency-Sentiment-LDA model which assumes that the sentiments of words form a Markov chain. The sentiment of a word is dependent on previous ones. The transitions of sentiments are determined using two types of conjunctions: related conjunctions and adversative conjunctions. Both of the above models incorporate sentiment lexicons as prior information, which could improve the performance.

5.3 Semi-supervised classifiers

A supervised learning algorithm is a very costly process, especially when dealing with large volumes of data. The most basic disadvantage of any unsupervised learning is that its application spectrum is limited. To counter these disadvantages, the concept of semi-supervised learning was introduced. In this type of learning, the algorithm is trained upon a combination of labeled and unlabeled data (Hussain and Cambria 2018). This combination will contain a very small amount of labeled data and a very large amount of unlabeled data (Miao et al. 2020; Matsuno et al. 2016; Nóra et al. 2010; Lu and Zhai 2008; Khan et al. 2017b; Yu and Kubler 2010).

An aspect based semi-supervised model was developed by Matsuno et al. (2016). They modeled the data into networks to perform the semi-supervised learning and used bipartite networks to represent the data, since network-based approaches have been successfully used to perform semi-supervised learning and the bipartite networks are parameter-free and fast to generate.

The Snippext (Miao et al. 2020), an opinion mining system, was developed over a language model that is fine-tuned through semi-supervised learning with augmented data. This system is for extracting aspect and opinion pairs, and corresponding sentiments from reviews by fine-tuning a language model with very little labeled training data. A semi-supervised topic modelling approach was proposed by integrating opinions scattered around in text articles with those in a well-written expert review for an arbitrary topic (Lu and Zhai 2008).

The problem of exploiting unlabeled samples to perform an emotion recognition task was addressed. They proposed a different regularization procedure which is able to in capsulate an unsupervised pre-training hint in a form of a reference hyperplane(Oneto et al. 2017).

One model was proposed to extract the opinions of the target words and implemented by using the semi-supervised word alignment model (Sadhana et al. 2017). To handle the scenario of insufficient initial labeled data, a novel semi-supervised model based on a dynamic threshold and multi-classifiers was proposed (Han et al. 2019). Their training data were auto-labeled in an iterative way based on the proposed dynamic threshold algorithm, where a dynamic threshold function was proposed to set thresholds for selecting the auto-labeled data.

Word polarity disambiguation is one important part of recent efforts on semi-supervised learning for social data analysis. At the pre-processing stage, a vector of context features is built for each word (w) based on all its occurrences in the positive polarity corpus, and another vector is based on its contexts in the negative polarity corpus. Lexico-syntactic context features are automatically generated from dependency parsed graphs of the sentences containing the word. These two vectors are treated as documents, one with positive and one with negative polarity. To resolve the contextual polarity of a specific instance of the w in a given sentence, its context feature vector is built in the same way, and is treated as the query. Thereafter, an information retrieval (IR)

model is then applied to calculate the similarity of the query to each of the two documents, with the polarity of the best matching document attributed to the query (Vechtomova 2017). This model performed better than the SVM and NB classifiers. Zhao et al. (2012) made use of the dependency relation between words and devised a statistical equation to calculate the probability that the given keyword carries certain sentiment polarity. To enable effective use of opinion-level features, Xia et al. (2015) adopted the Bayesian model to resolve the polarity in a probabilistic manner. Experiments with the opinion corpus demonstrate that opinion-level features can make a significant contribution to word polarity disambiguation in four domains.

5.4 Deep learning

The concept of deep learning was first introduced in opinion mining by Ain et al. (2017); it is capable of providing training to both supervised and unsupervised classification. Deep learning includes networks such as Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN), Recursive Neural Networks and more (Zhang et al. 2016b). Baly et al. (2017) developed a deep learning model (RNTN) and significantly improved the F-measure compared with baseline models. Glorot et al. (2011) proposed a deep learning approach which learns to extract a meaningful representation for each review in an unsupervised fashion using a Stacked Denoising Autoencoder (SDA). This has shown that linear classifiers trained with this higher-level learnt feature representation of domain adaptation for sentiment classification with deep learning outperforms the current state-of-the-art.

Zhou et al. (2010) proposed a novel semi-supervised learning algorithm called Active Deep Networks (ADN) using active learning. The ADN algorithm is constructed by Restricted Boltzmann Machines (RBM) with unsupervised learning using labeled data and abundant unlabeled data. Then the constructed structure is fine-tuned by gradient-descent based supervised learning with an exponential loss function. Second, active learning is applied in the semi-supervised learning framework to identify reviews that should be labeled as training data. The deep learning-based technique can be used for aspect extraction based on the user textual reviews (Da'u et al. 2020a). They have shown how the extracted aspects can be utilized for computing aspect-based ratings, which can finally be integrated into a tensor factorization machine for enhancing the accuracy of the recommendation system.

Recommender systems are utilized for solving information overload problems in areas such as e-commerce, entertainment, and social media (Shokeen and Rana 2020; Batmaz et al. 2019). There are many recommendation systems based on a deep learning approach (Da'u and Salim 2019). Da'u et al. (2020b) proposed a recommendation model based on weighted aspect-based opinion mining, which extracted aspects from the user text review that can be used to generate aspect ratings using a lexicon-based method.

One of the main benefits of deep learning over various machine learning algorithms is its ability to generate new features from a limited series of features located in the training dataset.

In opinion mining, deep learning generates actionable results when solving sentiment classification. While machine learning works only with labeled data, deep learning supports unsupervised learning techniques that allow the system to become smarter on its own (Estrada et al. 2020). The capacity to determine the most important features allows deep learning to efficiently provide data scientists with concise and reliable analysis results.

5.4.1 Convolution neural network(CNN)

CNN includes convolution layers, pooling layers and fully connected layers. Poria et al. (2016a) introduced aspect extraction for opinion mining. They used a seven-layer deep convolution neural network with a set of linguistic patterns. The results show that the model with word embedding and linguistic tags showed more significant results than a word embedding model. The word embedding is an effective tool for opinion reasoning (Cambria et al. 2017; Dragoni and Petrucci 2017).

Another deep learning opinion mining system (Severyn and Moschitti 2015) was proposed using twitter data. The model was initialized using previously embedded words and parameters. To train the network, Stochastic Gradient Descent (SGD) was used and the dropout technique was used to enhance the neural network regularization.

Vateekul and Koomsubha (2016) developed a dynamic neural network which inputs a sentence matrix $\subseteq R^{d\times s}$, where s is the sentence length and d is the word vector length. A filter matrix was used in the convolution layer, and a folding layer was created. Afterwards dynamic k-max pooling layers were used. Finally, the fully connected layer was used with a softmax classification. A significant result with 75.35% accuracy has been achieved with CNN comparing to other baseline models.

CNN was also used to classify medical text data at sentence level (Hughes et al. 2017). After testing several CNN configurations, the model consists of two convolution layers and two pooling layers. The model outperformed existing NLP models by more than 15%.

Recently, stock market analysis was done using a CNN model (Zhang et al. 2018). The method was implemented based on the word vectors. With the increased number of training data, the accuracy achieved was 76 %. Aspect based financial sentiment analysis was developed (Jangid et al. 2018) using multi-channel CNN, which produced an average F1 score of 0.69.

A multi-task ensemble framework that jointly learns multiple related problems was proposed by Akhtar et al. (2019). The ensemble model aims to leverage the learned representations of three deep learning models (i.e., CNN, LSTM and Gated Recurrent Unit [GRU]). Akhtar et al. (2020) proposed a Multi-layer Perceptron (MLP) based ensemble approach to leverage the goodness of various supervised systems. They developed one feature-driven supervised model and three deep neural network architecture based models, viz. LSTM, CNN, and GRU. The classical feature-based system utilizes a diverse set of features to train the model.

5.4.2 Recursive neural network(RNN)

RNN models are usually more efficient for modelling the syntactic structure of the input, as they feed the input sequence of words in a recursive manner.

The most recent model of RNN was introduced to identify top sellers in the underground economy (Li and Chen 2014). This model used the thread classification and snowball sampling. A sentiment tree bank was used and trained using the online corpus. Evaluation was conducted using baseline models of NB, K-Nearest Neighbors (KNN) and SVM. The results show significant outcomes.

A RNN model was proposed for Chinese sentiment classification which achieved higher accuracy than SVM, NB, and Maximum Entropy (ME) (Li et al. 2014). This research used 13550 labelled sentences which were from movie reviews. It firstly introduced the

sentiment tree bank and represented words with 50-dimensional word vectors. Based on recursive deep learning, it then predicted binary sentiment labels for opinions.

5.4.3 Memory network

Memory networks have recently been used by researchers for machine learning tasks. A memory network consists of multiple layers of attention. Wang et al. (2018) developed a memory network to replace the role of a syntactic/dependency parser to capture the relations among words in a sentence for information extraction. In order to model the interactions between aspect terms and opinion terms automatically, they proposed a memory network. In a sentence, they construct a pair of attentions: an aspect attention for aspect terms extraction and an opinion attention for opinion terms extraction. Each of them aims to learn a general prototype vector, a token-level feature vector and a token-level attention score for each word in the sentence. To capture direct relations between aspect and opinion terms, the aspect and opinion attentions are first coupled in learning such that the learning of each attention is affected by the other. This helps to double-propagate information between them. To further capture indirect relations among aspect and opinion terms, they constructed a memory network with multiple layers to update the learned prototype vectors, feature vectors, and attention scores to better propagate label information for aspect and opinion terms co-extraction.

Dyadic Memory Networks (DyMemNN) is proposed for aspect-based sentiment analysis. DyMemNN incorporates composition techniques that model the rich dyadic interactions between aspects and words in a document. To this end, two variations of DyMemNN, namely Tensor DyMemNN and Holo DyMemNN, were proposed, which extended memory networks with neural tensor compositions or holographic compositions (Tay et al. 2017).

5.4.4 Attention models

Existing deep learning models (Li and Chen 2014; Vateekul and Koomsubha 2016) have the disadvantage of fixed-length context vectors as these models are incapable of remembering longer sequences. To overcome the problem of fixed-length context vectors, an attention model was developed.

Yang et al. (2016) developed a Hierarchical Convolutional Attention Networks (HCAN) for text classification. They introduced new self-attention based text classification architecture. The performance of HCANs was compared with the current state-of-the-art, HANS, in four classification tasks: Yelp review sentiment, Amazon review sentiment, Amazon review product category, and Pubmed abstract topic. In all four tasks, HCANs achieved slightly better performance than HANs while it was more than twice as fast to train the learner. Their model uses two levels of attentions: word and sentence. The model consists of

- Word sequence encoder: For a given sentence of words, embed the words to vectors through an embedding matrix.
- Word level attention layer: Not all words contribute equally to the representation of the sentence. Hence, an attention mechanism was introduced to extract such words that are important for describing the meaning of the sentence, and the representation of those informative words is aggregated to form a sentence vector.

- Sentence encoder: For the given sentence vectors, create a document vector in a similar way.
- Sentence level attention layer: To identify a sentence to classify a document, use the attention mechanism, introduce a sentence level context vector, and use the vector to measure the importance of the sentence.

After evaluation on four data sets, the results demonstrated that this model performed significantly better than previous methods.

A novel model was recently implemented to incorporate user and product information for sentiment classification (Wu et al. 2018). The model firstly applied two neural networks to generate two representations. It then used a combining strategy for the training and final prediction, which combined user attention and product attention using the softmax function.

The newest technique applied in attention models is reinforcement-learning, which does the learning based on how best to react to situations through trial and error (Tay et al. 2017). It proposed Gated Multimodal Embedding LSTM with a Temporal Attention model for multimodal sentiment analysis, and performed multimodal fusion at word level. Furthermore, to build a model that is suitable for the complex structure of speech, this paper introduced selective word-level fusion between modalities, using a gating mechanism and reinforcement learning. The attention model is used to divert the focus of the model to important moments in speech.

Aspect-based extraction was done by Li et al. (2018) based on opinion summary and aspect detection history. The model contains two key components, namely Truncated History-Attention (THA) and Selective Transformation Network (STN). All the components in the proposed framework are differentiable. Therefore, the model can be efficiently trained with gradient methods.

5.4.5 Binary code learning

Hashing for collaborative filtering has attracted increasing attention as binary codes can significantly reduce the storage requirement and make similarity calculations efficient. Li et al. (2019) investigated deep collaborative hashing codes on user item ratings using a deep learning framework. In this research, neural networks can learn both user and item representations better and make these close to binary codes such that the quantization loss is minimized.

Generating the Top-N recommendations from a large corpus is computationally expensive to perform at scale. Kang and McAuley (2019) proposed a candidate generation and re-ranking based framework (CIGAR), which first learns a preference-preserving binary embedding for building a hash table to retrieve candidates, and then learns to re-rank the candidates using real-valued ranking models with a candidate-oriented objective. This drew more attention to the candidate generation problem in recommender systems.

Existing hashing methods for collaborative filtering focus on modeling the user-item similarity (a-k-a- preference) but omit the user-user and item-item similarities, which cannot effectively preserve the original geometry in the vector space. Zhang et al. (2019) proposed a method of neural binary representation (NBR) learning approach by combining hashing with a neural network for the large-scale collaborative filtering tasks. The NBR approach also takes the user-user and item-item similarities into account by imposing

anchor smoothing on the binary codes learning, in order to preserve the original geometry in the vector space as much as possible.

5.5 Summary of opinion classification methods

Table 2 summarizes all classification methods in our selected papers. The table shows that most of the available models are supervised models. According to the data shown in Figs. 6 and 7, there are fewer of unsupervised classification models, and most of the models available are binary classification, based on Tables 1 and 2.

The reason for using supervised learning is that it is easy to do the classification using labeled data. However, the decision boundary might be overtrained with the training data, which means that if the training set does not include some examples that you want for a class, the classifier will likely label some data incorrectly. It is also difficult to classify data if there are some uncertainties. Normally, supervised classification needs a large number of samples for the training process. Conversely, unsupervised classification is fairly quick and easy to run. There is no extensive prior knowledge required, but users need to identify and label classes after the learner finishes the job, as the classes are created based purely on spectral information; therefore, they are not as subjective as manual visual interpretation.

It is very difficult for the existing binary classification models to achieve very high accuracy, since opinions contain a lot of uncertainties. The key research issue is how to identify the boundary region explicitly to fully enhance the binary classification. Further research on unsupervised models and three-way classification will significantly enhance binary classification for opinion mining.
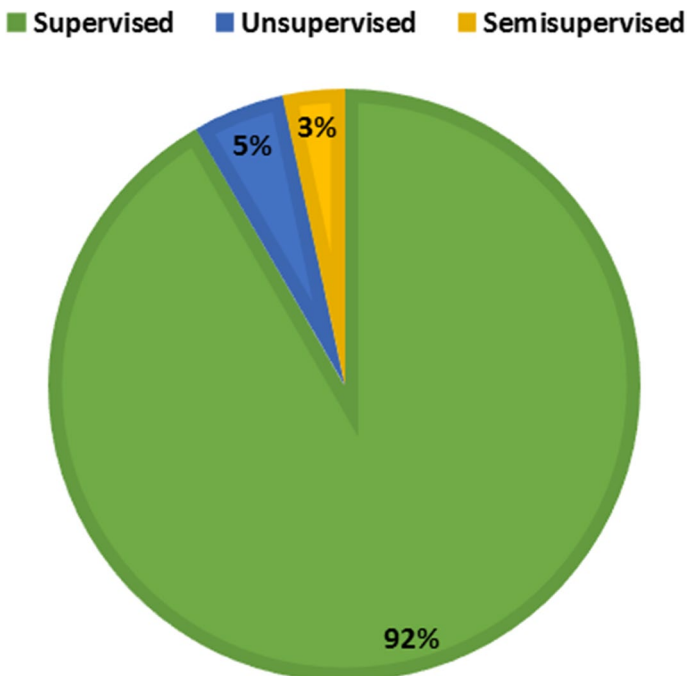

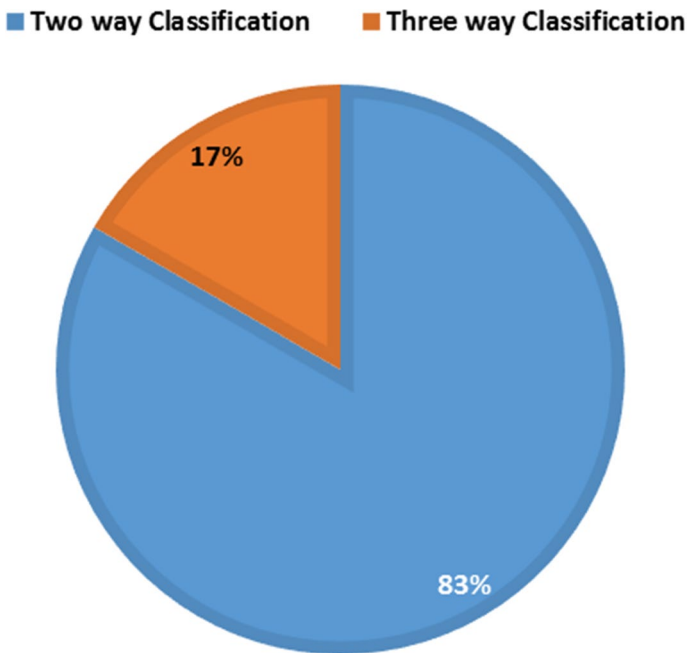**Fig. 6** Published articles: classification methods

Fig. 7 Published articles: number of polarity classification Method

6 Real world insights, open problems and future direction in opinion mining

Today in the real world, opinion mining is being employed to learn the thoughts and feelings of customers and the public, making businesses able to enhance the day-to-day shopping experience of customers, conduct competitive research, and understand opinions. Readers are eager to know what the real-world insights of opinion mining are. In this section, the first subsection will discuss real-world insights of opinion mining. The second subsection discusses open (unanswered) questions to which researchers are still searching for answers in opinion mining models. Some open questions addressed by researchers still remain unanswered. Accordingly, new research topics can go beyond the questions addressed in this survey paper. In the third subsection, the future research directions of opinion mining will be discussed.

6.1 Real-world insights of opinion mining

Most online shopping experiences have been shaped by data mining techniques. Currently Amazon is using basket analysis to predict customer behavior based on past purchases and preferences. For example, when a customer searches for products from Amazon, it suggests that the customer can purchase some products based on previous purchases. Customers can obtain wish-lists, recommendations and search functions for the product facilities.

Table 2 Opinion mining: classification methods

References	Classification Method	Supervised Unsupervised	Polarity	Performance	Dataset
Li et al. (2010b)	RFD (relevance feature discovery)	Supervised	Positive and negative	F-measure of 0.91 %	SemEval14
Chen et al. (2011)	ANN	Supervised	Positive and negative	Highest accuray of 61.1% percentage	Movie review
Saleh et al. (2011)	SVM	Supervised	Positive and negative	The highet accuracy of 0.9115 with TF-IDF feature selection	Dataset 1 (Pang et al. 2002)
Eirinaki et al. (2012)	Maximum opinion score algorithm	Not Given	Rank Opinions	Precision of 0.85 for 50 reviews	Dataset 1 (Hu and Liu 2004)
Kang et al. (2012)	SVM	Supervised	Positive and negative	Precision of results is greater than 0.8 for all datasets	Movie Review
Nadali and Murad (2012)	Fuzzy Logic	Supervised	Very strong, strong, moderate, weak, very weak and average	Accuracy of 74 %	RCV1
Basari et al. (2013)	SVM and particle swarm approach	Supervised	Positive and negative	Accuracy of 77%	Movie Reviews
Li and Tsai (2013)	Fuzzy logic	Supervised	Positive and negative	The highest accuracy reuters-21578. Precision of 0.962, recall of 0.961 and F-measure of 0.967	Ebook, movie reviews and reauters-21578
Moraes et al. (2013)	ANN, SVM and NB	Supervised	Positive and negative	Highest accuracy of 87 % is achieved by NN	Ebook and movie reviews
Khairnar and Kinikar (2013)	SVM	Supervised	Positive and negative	Not given	Twitter data
Bing and Chan (2014)	Fuzzy logic	Supervised	Positive and negative	Not given	Twitter data
Dalal and Zaveri (2014)	Fuzzy logic	Supervised	Positive and negative	Average classification accuracy of 86.47 %	Movie review
Bravo-Marquez et al. (2014)	Lexicon based	Supervised	Positive and negative	Average classification accuracy of 93.7 %	SemEval

Table 2 (continued)

References	Classification Method	Supervised Unsupervised	Polarity	Performance	Dataset
Dragoni et al. (2014)	Fuzzy logic	Supervised	Positive and negative	Precision of 0.8675 and recall of 0.9987	Blitzer dataset
Haque et al. (2014)	Fuzzy logic	Supervised	Positive and negative	Not given	ADR comments
Pimpalkar et al. (2014)	Fuzzy logic	Supervised	Positive and negative	Not given	Twitter and Facebook reviews
Priyadarshi et al. (2014)	Fuzzy logic	Supervised	Positive and negative	Not given	Amazon smart phone review dataset
Sharma et al. (2014)	NLP (POS tagging and word net)	Supervised	Positive, negative and neutral	Accuracy of 63 %	Movie reviews
Quan and Ren (2014)	Feature oriented classification	Unsupervised	Positive, negative and neutral	F-measure is depends on the datasets. Highest F-measure of 87.7 % is for mp3 dataset.	Facebook
Chenlo and Losada (2014)	NLP (POS tagging)	Supervised	Positive and negative	F-measure is 90.31%.	Polarity collection
Prakash et al. (2015)	SVM	Supervised	Positive and negative	Accuracy of 96.2 %	RCV1 dataset
Zimmermann et al. (2015)	Fuzzy logic	Supervised	Positive and negative	With the number of reviews kappa value is higher in the model	Dataset 1 (Hu and Liu 2004) and Dataset 2 (Yu et al. 2011)
Xu et al. (2015)	SVM	Supervised	Positive and negative	Kernel private function showed better results	NTCIR-7 MOAT
Afzaal et al. (2016)	Fuzzy logic	Supervised	Positive and negative	Accuracy of Furia Algorithm is the highest at 90.12 %	Tourist Reviews
Ali et al. (2016)	Fuzzy logic	Supervised	Strongly positive, positive, neutral, negative and strongly negative	Highest accuracy of 85.4% for room dataset	Online reviews
Jadav and Vaghela (2016)	RBL Kernel SVM	Supervised	Positive and negative	For twitter data highest accuracy of 78.18 %	Twitter and Amazon data

Table 2 (continued)

References	Classification Method	Supervised Unsupervised	Polarity	Performance	Dataset
Jiménez-Zafra et al. (2016)	NLP	Unsupervised	Positive and negative	F-measure of 0.808 for restaurant reviews and F-measure of 0.566 for laptops	Movie review
Sadidpour et al. (2016)	Polarity pattern	Supervised	Positive and negative	Precision of 0.91 and recall of 0.48	Political news
Mamgain et al. (2016)	Association rules	Supervised	Positive and negative	For each perfoamance metrix IIT outperforms other institutes	Twitter data
Panchendrarajan et al. (2016)	Rule based	Supervised	Positive and negative	From all the models the highest F-measure is 0.842	Yelp dataset
Vinodhini and Chandrasekaran (2016)	ANN	Supervised	Positive and negative	The highest completeness is 0.93, effectiveness is 91.5 and efficiency is 0.91	Amazon reviews
Zhang et al. (2016a)	Fuzzy logic	Supervised	Positive and negative	Accuracy of the proposed model was 0.9429 for cannon dataset	Movie review
Matsuno et al. (2016)	Bipartite heterogeneous networks	Semi-supervised	Positive and negative	Accuracy of the proposed model was 0.95 for cannon dataset	Reviews of laptops
Angelpreethi and Kumar (2017)	NLP (syntactic dependancy)	Supervised	Positive and negative	Not given	Web options
Sadhana et al. (2017)	NLP (syntactic dependency)	Semi-supervised	Positive and negative	F measure was 0.716 for positive and 0.825 for negative	Amazon dataset
Khan et al. (2017b)	SVM	Semi-supervised	Positive and negative	F measure of 86.50 %	Facebook dataset
Xia et al. (2017)	Life long classifier using NB	Supervised	Positive and negative	F measure of 78.50 %	Twitter dataset and a Chinese Weibo dataset)

Table 2 (continued)

References	Classification Method	Supervised Unsupervised	Polarity	Performance	Dataset
Boudia et al. (2017)	Fuzzy Logic	Supervised	Strongly positive, positive, week positive, week negative, negative and strongly negative	Precision of 0.6746, Recall of 0.6718 and F-measure of 0.6732	Twitter data
Catal and Nangir (2017)	SVM , NB and enhance SVM	Supervised	Positive and negative	Highest accuracy of 87% for Antoloji dataset	Movie reviews
Gu et al. (2017)	ANN	Supervised	Positive and negative	Accuracy of 98.26 %	Twitter data
Hu et al. (2017)	Sentence similarity	Not given	Rank opinions	Not given	TripAdvisor
Howells and Ertugan (2017)	Fuzzy logic	Supervised	Strongly positive, positive, week positive, week negative, negative and strongly negative	Not given	Not given
Rana and Cheah (2017)	NLP (syntactic patterns-based approaches)	Supervised	Positive and negative	Maximum Precision of 0.81 and recall of 0.72.	Dataset 1 (Hu and Liu 2004)
Phu et al. (2017)	SVM	Supervised	Positive and negative	Not given	Movie reviews dataset
Ravi et al. (2017)	Fuzzy logic	Supervised	Positive and negative	Not given	Bank datasets
Sadhana et al. (2017)	NLP (Word alignment model)	Semi-supervised	Positive and negative	For positive reviews F-measure of 0.534 and for negative F-measure of 0.645	Amazon dataSet
Vinodhini and Chandrasekaran (2017)	SVM	Supervised	Positive and negative	Not given	IBD1 and IBD2
Jing et al. (2018)	NLP (lexicon approach)	Unsupervised	Positive and negative	Not given	JD.com
Deshmukh and Tripathy (2018)	Probability based	Supervised	Positive and negative	Accuracy of 88.4 %	Amazon dataset
Kang et al. (2018)	HMM	Supervised	Positive and negative	Average accuracy of 92.5 %	(Hu and Liu 2004)
Van Hee et al. (2018)	SVM	Supervised	Positive and negative	Average Accuracy of 78.1 %	Twitter

Table 2 (continued)

References	Classification Method	Supervised Unsupervised	Polarity	Performance	Dataset
Miao et al. (2020)	NLP (lexicon approach)	Supervised	Positive and negative	F-measure of 0.8539	Yelp dataset reviews
Han et al. (2019)	SVM	Semi-supervised	Positive and negative	F-measure of 0.8064	Movie reviews
Azizan et al. (2019)	Lexicon score	Supervised	Positive and negative	F-measure of 0.8000	Movie reviews
Jaman and Abdulrohman (2019)	SVM	Supervised	Positive and negative	Accuracy of 90.10%	Twitter
Jaman and Abdulrohman (2019)	TF-IDF	Supervised	Positive Negative	Not given	Flickr photos
Khattak et al. (2020)	Fuzzy	Supervised	Positive and negative	F-measure of 0.9000	Drugs, electronics and mobile phone reviews
Jain et al. (2020)	Neutrosophic Set and particle swarm optimization	Supervised	Positive and negative	F-measure of 0.9590 for Blitzer data set	Blitzer, aclIMDb, polarity and subjective dataset.
You et al. (2020)	decision tree, NB, logistic regression and SVM	Supervised	Positive and negative	Highest accuracy is 68.8 % was obtained by SVM	HCR, OMD and STS data set.
Estrada et al. (2020)	CNN, LSTM and ANN	Supervised	Positive and negative	Highest accuracy is 93 % was obtained using deep learning	

Customers always care about the discounts that they gain. Amazon has focused on offering products that are competitive in cost. Currently, this function is available while shopping at Amazon. In the future, this feature can be enhanced by identifying the products that are popular among customers using opinion mining.

Amazon Comprehend (Dale 2018) use machine learning to find the relationship between words in text. This will help content creators and marketers to create personalized recommendations with novel deep learning techniques.

eBay provides a product search facility through matching keywords. Understanding what customers want from keywords that they search for is very important. Product recommendation is the next important part of eBay. This needs to be informed by the product attributes, price ranges, user purchase patterns, and product categories. It has been shown that online companies can improve their profits by using opinion mining.

Opinion mining and sentiment analysis have also been applied on social media data for analysing various socio-economic events, including electoral study and nature disasters management. Predicting political sentiments of voters from twitter data have been developed by Khatua et al. (2020). The model has used a multinomial logistic regression as well as ANN based method to probe the political opinions of users. The model has explored the mix tweets that jointly mention more than one party. This research has elucidated the significance of mix tweets in a multi-party context. Outbreak management applications are recently developed using opinion mining concept. Khatua et al. (2019) developed a outbreak management application for unstructured twitter streams. The application uses word2vec word embedding model for training of deep learning models. Researchers found that the accuracy of word vectors changes in response to input corpora (Twitter-based corpus, model architectures) and hyper-parameter settings (dimension and context window size) in the context of infectious disease outbreaks (Khatua et al. 2019). Finally, this study portrays that during the initial stages of the epidemic, it is hard to collect and aggregate twitter corpus. PubMed is identified as relief option to solve this issue.

6.2 Open questions

We identify and summarize the open questions in opinion mining as follows.

- *Most existing opinion mining models do not focus on dealing with uncertainties in opinions.*

  The problem is that opinions are subjective and vague (uncertain). For example, if an opinion word expresses, not bad, it is likely labelled as positive, as machine learning usually selects features based on frequent words. This problem exists because machine learning algorithms cannot deal with uncertainty as they can only identify words' statistical properties, e.g., frequency. The question is, how can we minimize the uncertain information in opinions. In the research arena, existing supervised classification is used for binary classification and cannot clearly handle uncertain boundaries. That is, it cannot define a clear boundary between positive and negative classes. Unsupervised classification mainly generates clusters with similar functions. The uncertainty problem still remains because of the limitation of feature selection methods. In addition, both supervised and unsupervised models capture less semantic information which reduces the performance of classifiers. With the development of fuzzy models for opinion mining, researchers tried to solve the problem of uncertainty in the representation of knowl-

edge; however, identification of uncertain boundaries between classes is still an open problem.

- *It is a big challenge to identify and extract high quality features and knowledge* (Catal and Nangir 2017; Chen et al. 2011; Gu et al. 2017; Boudia et al. 2017) *in opinion mining.*

   In the field of opinion mining, it is still difficult to identify proper entities and attributes as representative features or valuable knowledge. We also found that the relationship between attributes was not considered well in opinion mining when selecting training data because of uncertainties and the semantic aspects of opinions (Li and Tsai 2013; Moraes et al. 2013; Vinodhini and Chandrasekaran 2017; Howells and Ertugan 2017).

   Researchers applied NLP techniques to handle semantic aspects of opinions (Liu et al. 2005; Jiménez-Zafra et al. 2016). Most models use term-based methods for feature selection (Dalal and Zaveri 2014). Some models directly use terms for the classification (Li et al. 2010b; Chen et al. 2011; Basari et al. 2013), while other methods employ patterns and POS tags to represent knowledge and do the classification (Shinde and Gill 2014; Wu et al. 2004). When we compare the accuracy, the models with higher level features (e.g. topic, concept, knowledge) showed higher accuracy than the other (Li and Tsai 2013). Li et al. (2010b, 2015b) suggested that the combination of low level features (e.g. terms) and high level concepts (patterns) can improve the performance of text mining.

   Most knowledge can be in different forms such as terms, patterns, concepts, etc. Selecting the best form for the classification process is also a research issue. Existing methods such as fuzzy logic, NLP, and pattern mining are used separately for the knowledge discovery process. Researchers can experiment on the combination of different forms to select the best way for knowledge discovery. Pattern mining can identify high level features (knowledge) using patterns, however, deciding the minimum support is a difficult problem. Frequent patterns, and closed patterns have been used to represent knowledge in reviews (Pasquier et al. 1999; Han et al. 2002; Ghorashi et al. 2012). But how to extract only relevant patterns from reviews? This is an open problem for researchers. The combination of several methods such as terms and patterns may increase the accuracy of the classifier (Xu et al. 2015).

- *Most opinion mining systems consider only customer reviews and ignore the manufacturers' perceptions.*

   The available opinion mining systems (Vinodhini and Chandrasekaran 2016; Ahmad and Doja 2013; Amarouche et al. 2015; Quan and Ren 2014; Tsirakis et al. 2017; Sadhana et al. 2017) are mining customer reviews without considering the manufacturer's perspective. Customers are heavily reliant on peer viewpoints when they purchase a new product and product information is essential to support the recommendation. Trusted and reliable product information is available in product documentations such as specifications, user manuals and design documents, where they reflect the correct perception of the manufacturers. The manufacturers' perception includes the product features as declared by the manufacturer, their functionality (overall and individual), associations with other products and features, exceptions under various working conditions, bug fixes, best practices, and recommended user behaviours, etc. Nevertheless, most of the time, the techniques which facilitate finding the other viewpoints, such as opinion mining, and product reputations have only considered the users' perceptions of the product. Their feedback and the comments which lead to the recommendation process can only capture the users' perceptions while giving less attention to the manu-

facturers' perception of that product. Neglect of manufacturers' perceptions may hinder the quality and the accuracy of the recommendation where it is based solely on users' perceptions which are sometimes biased or incomplete (He et al. 2016). Hence, it is essential to obtain both manufacturers' perceptions and users' perceptions in order to get the correct idea of the product. A collection of rich product information will set a strong foundation for an opinion mining system (Fernandes and GL 2017).

- *It is still in doubt which classification level an opinion mining classifier should focus on.*

   As we mentioned before, there are three levels of opinion mining: document level, sentence level and aspect level. In document level, a complete document is considered as an opinion of the customer review and extracted opinions are then used for classification. In sentence level, each sentence gets one opinion; and in aspect level, extract aspects are used for classification, which may be a more accurate way. Document level considers the whole document, therefore, scanning the whole document to compute the opinion words takes more computational time. Sentence level uses an opinion for each sentence, but a sentence cannot describe the user review completely. Aspect level extracts the aspects and relationships between them. Researchers could build a model to get all three level results to decide the best level of opinion mining. However, the best level can be different for different downstream applications.

- *How to deal with biased or malicious reviews for opinion mining classification?*

   Spam and misleading reviews are becoming a critical challenge in the opinion mining community and e-commence industry. We need to develop a detection mechanism to validate reviews using stronger user models (Titov and McDonald 2008; McAuley and Leskovec 2013) to improve confidence in the model as well as in each review. Implementing a model with these features may increase the confidence of customers.

6.3 Future directions

Despite recent advances and research efforts in opinion mining, there is still a plethora of open issues not addressed in this paper that are calling for researchers' attention. For building the next generation of opinion mining systems, we recommend the following research directions based on the recent new developments in text analysis.

- *Develop a detection mechanism to identify biased and malicious reviews.*

   The problem of identifying fake reviews has recently attracted significant interest. Writing fake reviews is a form of cyber-attack as it aims to purposely harm or boost an item reputation. Therefore, it is vital to develop a mechanism to detect these biased and malicious reviews to provide clean information for opinion mining.

- *Take manufactures' perceptions into account for designing opinion mining models.*

   It is desirable to obtain both manufacturers' perceptions and users' perceptions in order to clearly understand user reviews or opinions. Ontology mining will play a major role in the representation of manufacturers' perceptions, including the product information. Therefore, researchers can develop ontology mining models to understand useful features and knowledge within opinions or reviews.

- *Combine multiple feature selections and knowledge discovery techniques.*

   Most opinion mining models use one kind of useful feature. The big problem is how to select a set of relevant features from multiple kinds of useful features. It has been shown that multiple feature selection methods can increase the accuracy of the

classifier (Li and Tsai 2013). The best way forward is the ensemble application of linguistics and knowledge bases, because different approaches of semantic knowledge and machine learning can cover for each other's flaws (Cambria 2016). In the future, researchers can consider more possible combinations of feature selection methods to increase the accuracy. For an example, the combinations of LDA pattern mining with clustering together to select features from large collection of opinions (Alharbi et al. 2017b, 2018, 2017a) can be a promising way in opinion mining. As we mentioned the above, some classifiers use features and some of them use knowledge, or both features and knowledge. It is shown that a model can perform better if it wisely uses knowledge (Li and Tsai 2013; Sadidpour et al. 2016; Ahmad and Doja 2013).

- *Develop a mechanism to handle uncertainties in opinions to enhance sentiment classification.*

  Most of the existing models were developed for binary classification. Three-way decisions were already applied for sentiment classification (Zhang and Wang 2014) to classify opinions into three groups. Three-way decisions are applied as a mathematical tool to handle vagueness and uncertainty of opinions. Recently Three-way decision theory has been extended to finding the boundary region between positive and negative classes to enhance binary classification (Li et al. 2017; Wu et al. 2019) for text classification and document summarization. We think the extended three-way decisions can be used to solve the problem of uncertainties for sentiment classification; this needs further research.

7 Conclusion

This survey paper presents an overview of recent updates on opinion mining in three themes: text feature selection, knowledge representation, and classification of opinions. Papers between 2000 and 2020 were selected based on the proposed selection criteria. After exploring the sentiment analysis process in the relevant literature, it is evident that opinion mining is still an emerging and hot area and there are several open research questions.

Feature selection contributes to sentiment classification to select useful opinion words by using one or multiple feature selection methods. Knowledge representation is used to describe relevant features for a given class and implemented by a corresponding knowledge discovery algorithm. For the NLP community, it has been found that some classifiers use features and some of them use knowledge, or they use both features and knowledge. It appears that wisely employing knowledge can improve the accuracy of opinion mining systems. It is implied that the combination of features and knowledge may perform better than other models for sentiment classification. Fuzzy formal concepts and pattern mining have been used in opinion mining to represent and discover knowledge in reviews.

For the classification task, SVM and NB are the most widely used supervised approaches for opinion mining. To overcome the limitations of SVM and NB, researchers applied deep learning models and attention models to improve the performance. It has shown that deep learning and attention models can produce better results than other state-of-the-art models.

Finally, we have identified five open questions and recommend four future directions for designing and implementing the next generation of sentiment analysis systems.

References

Afzaal M, Usman M, Fong ACM, Fong S, Zhuang Y (2016) Fuzzy aspect based opinion classification system for mining tourist reviews. Adv Fuzzy Syst 2016:2–6

Ahmad T, Doja MN (2013) (2013) Opinion mining using frequent pattern growth method from unstructured text. In: International symposium on computational and business intelligence (ISCBI). IEEE, pp 92–95

Ain QT, Ali M, Riaz A, Noureen A, Kamran M, Hayat B, Rehman A (2017) Sentiment analysis using deep learning techniques: a review. Int J Adv Comput Sci Appl 8(6):424

Akhtar MS, Ekbal A, Cambria E (2020) How intense are you? predicting intensities of emotions and sentiments using stacked ensemble. IEEE Comput Intell Mag 15(1):64–75

Akhtar S, Ghosal D, Ekbal A, Bhattacharyya P, Kurohashi S (2019) All-in-one: emotion, sentiment and intensity prediction using a multi-task ensemble framework. In: IEEE transactions on affective computing, pp 1–1

Albishre K, Li Y, Xu Y, Huang W (2020) Query-based unsupervised learning for improving social media search. World Wide Web 23(3):1791–1809

Alharbi AS, Li Y, Xu Y (2017a) Integrating lda with clustering technique for relevance feature selection. In: Australasian joint conference on artificial intelligence. Springer, pp 274–286

Alharbi AS, Li Y, Xu Y (2017b) Topical term weighting based on extended random sets for relevance feature selection. In: Proceedings of the international conference on web intelligence. ACM, pp 654–661

Alharbi AS, Li Y, Xu Y (2018) An extended random-sets model for fusion-based text feature selection. In: Pacific-Asia conference on knowledge discovery and data mining. Springer, pp 126–138

Ali F, Kwak KS, Kim YG (2016) Opinion mining based on fuzzy domain ontology and support vector machine: a proposal to automate online review classification. Appl Soft Comput 47:235–250

Amarouche K, Benbrahim H, Kassou I (2015) Product opinion mining for competitive intelligence. Procedia Comput Sci 73:358–365

Angelpreethi A, Kumar SBR (2017) An enhanced architecture for feature based opinion mining from product reviews. In: 2017 world congress on computing and communication technologies (WCCCT). IEEE, pp 89–92

Araújo M, Gonçalves P, Cha M, Benevenuto F (2014) ifeel: a system that compares and combines sentiment analysis methods. In: Proceedings of the 23rd international conference on world wide web, pp 75–78

Asgarian E, Kahani M, Sharifi S (2018) The impact of sentiment features on the sentiment polarity classification in persian reviews. Cogn Comput 10(1):117–135

Azizan A, Jamal NNSA, Abdullah MN, Mohamad M, Khairudin N (2019) Lexicon-based sentiment analysis for movie tweets. In: 2019 1st international conference on artificial intelligence and data sciences (AiDAS). IEEE, pp 132–136

Bafna K, Toshniwal D (2013) Feature based summarization of customers reviews of online products. Procedia Comput Sci 22:142–151

Baly R, Hajj H, Habash N, Shaban KB, El-Hajj W (2017) A sentiment treebank and morphologically enriched recursive deep models for effective sentiment analysis in arabic. ACM Trans Asian Low-Resour Lang Inf Process (TALLIP) 16(4):1–21

Basari ASH, Hussin B, Ananta IGP, Zeniarja J (2013) Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization. Procedia Eng 53:453–462

Batmaz Z, Yurekli A, Bilge A, Kaleli C (2019) A review on deep learning for recommender systems: challenges and remedies. Artif Intell Rev 52(1):1–37

Bayardo RJ Jr (1998) Efficiently mining long patterns from databases. ACM Sigmod Rec 27(2):85–93

Bhardwaj A, Narayan Y, Dutta M et al (2015) Sentiment analysis for indian stock market prediction using sensex and nifty. Procedia Comput Sci 70:85–91

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(3):330–344

Bilici E, Saygın Y (2017) Why do people (not) like me?: mining opinion influencing factors from reviews. Expert Syst Appl 68:185–195

Bing L, Chan KC (2014) A fuzzy logic approach for opinion mining on large scale twitter data. In: 2014 IEEE/ACM 7th international conference on utility and cloud computing (UCC). IEEE, pp 652–657

Boote DN, Beile P (2005) Scholars before researchers: on the centrality of the dissertation literature review in research preparation. Edu Res 34(6):3–15

Boudia MA, Hamou RM, Amine A (2017) Fuzzy opinion: detection of opinion based on sentiwordnet dictionary by using fuzzy logic. In: Fuzzy systems: concepts, methodologies, tools, and applications. IGI Global, pp 1576–1595

Bouma G (2009) Normalized (pointwise) mutual information in collocation extraction. In: Proceedings of GSCL, pp 31–40

Bravo-Marquez F, Mendoza M, Poblete B (2014) Meta-level sentiment models for big social data analysis. Knowl-Based Syst 69:86–99

Cambria E (2016) Affective computing and sentiment analysis. IEEE Intell Syst 31(2):102–107

Cambria E, Grassi M, Hussain A, Havasi C (2012) Sentic computing for social media marketing. Multimed Tools Appl 59(2):557–577

Cambria E, Schuller B, Liu B, Wang H, Havasi C (2013a) Knowledge-based approaches to concept-level sentiment analysis. IEEE Intell Syst 28(2):12–14

Cambria E, Schuller B, Xia Y, Havasi C (2013b) New avenues in opinion mining and sentiment analysis. IEEE Intell Syst 28(2):15–21

Cambria E, Hussain A, Vinciarelli A (2017) Affective reasoning for big social data analysis. IEEE Trans Affect Comput 8(4):426–427

Cambria E, Li Y, Xing FZ, Poria S, Kwok K (2020) Senticnet 6: ensemble application of symbolic and subsymbolic ai for sentiment analysis. In: International conference on information and knowledge management, pp 105–114

Catal C, Nangir M (2017) A sentiment classification model based on multiple classifiers. Appl Soft Comput 50:135–141

Chee CH, Jaafar J, Aziz IA, Hasan MH, Yeoh W (2019) Algorithms for frequent itemset mining: a literature review. Artif Intell Rev 52(4):2603–2621

Chen CC, Tseng YD (2011) Quality evaluation of product reviews using an information quality framework. Decis Support Syst 50(4):755–768

Chen LS, Liu CH, Chiu HJ (2011) A neural network based approach for sentiment classification in the blogosphere. J Informetr 5(2):313–322

Chenlo JM, Losada DE (2014) An empirical study of sentence features for subjectivity and polarity classification. Inf Sci 280:275–288

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

Dalal MK, Zaveri MA (2014) Opinion mining from online user reviews using fuzzy linguistic hedges. Appl Comput Intell Soft Comput 2014:2

Dale R (2018) Text analytics apis, part 1: the bigger players. Nat Lang Eng 24(2):317–324

Damani OP (2013) Improving pointwise mutual information (pmi) by incorporating significant co-occurrence. In: Proceedings of the seventeenth conference on computational natural language learning, pp 20–28

Darrab S, Eregenc B (2016) Frequent pattern mining under multiple support thresholds. Wseas Trans Comput Res 10:11

Da'u A, Salim N (2019) Recommendation system based on deep learning methods: a systematic review and new directions. Artif Intell Rev 53:1–40

Da'u A, Salim N, Rabiu I, Osman A (2020a) Recommendation system exploiting aspect-based opinion mining with deep learning method. Inf Sci 512:1279–1292

Da'u A, Salim N, Rabiu I, Osman A (2020b) Weighted aspect-based opinion mining using deep learning for recommender system. Expert Syst Appl 140:112871

Deshmukh JS, Tripathy AK (2018) Entropy based classifier for cross-domain opinion mining. Appl Comput Inf 14(1):55–64

Dragoni M, Petrucci G (2017) A neural word embeddings approach for multi-domain sentiment analysis. IEEE Trans Affect Comput 8(4):457–470

Dragoni M, Tettamanzi AG, da Costa Pereira C (2014) A fuzzy system for concept-level sentiment analysis. In: Semantic web evaluation challenge. Springer, pp 21–27

Eirinaki M, Pisal S, Singh J (2012) Feature-based opinion mining and ranking. J Comput Syst Sci 78(4):1175–1184

Esparza SG, O'Mahony MP, Smyth B (2012) Mining the real-time web: a novel approach to product recommendation. Knowl-Based Syst 29:3–11

Estrada MLB, Cabada RZ, Bustillos RO, Graff M (2020) Opinion mining and emotion recognition applied to learning environments. Expert Syst Appl 150:113265

Fan TK, Chang CH (2011) Blogger-centric contextual advertising. Expert Syst Appl 38(3):1777–1788

Fang Z, Zhang Q, Tang X, Wang A, Baron C (2020) An implicit opinion analysis model based on feature-based implicit opinion patterns. Artif Intell Rev 53:1–28

Fautsch C, Savoy J (2009) Algorithmic stemmers or morphological analysis? an evaluation. J Am Soc Inf Sci Technol 60(8):1616–1624

Fernandes R, Rio D'Souza GL (2017) Semantic analysis of reviews provided by mobile web services using rule based and supervised machine learning techniques. Int J Appl Eng Res 12(22):12637–12644

Fine S, Singer Y, Tishby N (1998) The hierarchical hidden markov model: analysis and applications. Mach Learn 32(1):41–62

Gao Y, Xu Y, Li Y (2015) Pattern-based topics for document modelling in information filtering. IEEE Trans Knowl Data Eng 27(6):1629–1642

Ghorashi SH, Ibrahim R, Noekhah S, Dastjerdi NS (2012) A frequent pattern mining algorithm for feature extraction of customer reviews. IJCSI Int J Comput Sci Issues, Citeseer 9:29–35

Glorot X, Bordes A, Bengio Y (2011) Domain adaptation for large-scale sentiment classification: a deep learning approach. In: International conference on machine learning, pp 513–520

Gu X, Gu Y, Wu H (2017) Cascaded convolutional neural networks for aspect-based opinion summary. Neural Process Lett 46(2):581–594

Hagenau M, Liebmann M, Neumann D (2013) Automated news reading: stock price prediction based on financial news using context-capturing features. Decis Support Syst 55(3):685–697

Han J, Wang J, Lu Y, Tzvetkov P (2002) Mining top-k frequent closed patterns without minimum support. In: 2002 IEEE international conference on data mining. IEEE, pp 211–218

Han Y, Liu Y, Jin Z (2019) Sentiment analysis via semi-supervised learning: a model based on dynamic threshold and multi-classifiers. Neural Comput Appl 32:1–13

Haque M et al (2014) Sentiment analysis by using fuzzy logic. Int J Comput Sci Eng Inf Technol (IJCSEIT) 4:33–48

He C, Parra D, Verbert K (2016) Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. Expert Syst Appl 56:9–27

He R, Lee WS, Ng HT, Dahlmeier D (2018) Effective attention modeling for aspect-level sentiment classification. In: Proceedings of the 27th international conference on computational linguistics, pp 1121–1131

Hemmatian F, Sohrabi MK (2019) A survey on classification techniques for opinion mining and sentiment analysis. Artif Intell Rev 52:1–51

Hofmann T (2017) Probabilistic latent semantic indexing. ACM SIGIR Forum, ACM 51:211–218

Howells K, Ertugan A (2017) Applying fuzzy logic for sentiment analysis of social media network data in marketing. Procedia Comput Sci 120:664–670

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

Hu YH, Chen YL, Chou HL (2017) Opinion mining from online hotel reviews: a text summarization approach. Inf Process Manag 53(2):436–449

Hughes M, Li I, Kotoulas S, Suzumura T (2017) Medical text classification using convolutional neural networks. Stud Health Technol Inform 235:246–50

Hussain A, Cambria E (2018) Semi-supervised learning for big social data analysis. Neurocomputing 275:1662–1673

Ifrim G, Bakir G, Weikum G (2008) Fast logistic regression for text categorization with variable-length n-grams. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp 354–362

Jadav BM, Vaghela VB (2016) Sentiment analysis using support vector machine based on feature selection and semantic analysis. Int J Comput Appl 146(13):26–30

Jain A, Nandi BP, Gupta C, Tayal DK (2020) Senti-nsetpso: large-sized document-level sentiment analysis using neutrosophic set and particle swarm optimization. Soft Comput 24(1):3–15

Jaman JH, Abdulrohman R (2019) Sentiment analysis of customers on utilizing online motorcycle taxi service at twitter with the support vector machine. In: 2019 international conference on electrical engineering and computer science (ICECOS). IEEE, pp 231–234

Jangid H, Singhal S, Shah RR, Zimmermann R (2018) Aspect-based financial sentiment analysis using deep learning. In: Companion of the web conference 2018 on international world wide web conferences steering committee, pp 1961–1966

Jiménez-Zafra SM, Martín-Valdivia MT, Martínez-Cámara E, Ureña-López LA (2016) Combining resources to improve unsupervised sentiment analysis at aspect-level. J Inf Sci 42(2):213–229

Jindal N, Liu B (2006) Mining comparative sentences and relations. AAAI Conf Artif Intell 22:1331–1336

Jing N, Jiang T, Du J, Sugumaran V (2018) Personalized recommendation based on customer preference mining and sentiment assessment from a Chinese e-commerce website. Electron Commer Res 18(1):159–179

Jinturkar M, Gotmare P (2016) Sentiment analysis of customer review data using big data: a survey. Int J Comput Appl ETC 2016:3–8

Joshi M, Penstein-Rosé C (2009) Generalizing dependency features for opinion mining. In: Proceedings of the ACL-IJCNLP 2009 conference short papers, association for computational linguistics, pp 313–316

Kadhim AI (2019) Survey on supervised machine learning techniques for automatic text classification. Artif Intell Rev 52(1):273–292

Kang H, Yoo SJ, Han D (2012) Senti-lexicon and improved naïve bayes algorithms for sentiment analysis of restaurant reviews. Expert Syst Appl 39(5):6000–6010

Kang M, Ahn J, Lee K (2018) Opinion mining using ensemble text hidden Markov models for text classification. Expert Syst Appl 94:218–227

Kang WC, McAuley J (2019) Candidate generation with binary codes for large-scale top-n recommendation. In: Proceedings of the 28th ACM international conference on information and knowledge management, pp 1523–1532

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

Khairnar J, Kinikar M (2013) Machine learning algorithms for opinion mining and sentiment classification. Int J Sci Res Publ 3(6):1–6

Khan FH, Qamar U, Bashir S (2017a) Lexicon based semantic detection of sentiments using expected likelihood estimate smoothed odds ratio. Artif Intell Rev 48(1):113–138

Khan FH, Qamar U, Bashir S (2017b) A semi-supervised approach to sentiment analysis using revised sentiment strength based on sentiwordnet. Knowl Inf Syst 51(3):851–872

Khan K, Baharudin B, Khan A, Ullah A (2014) Mining opinion components from unstructured reviews: a review. J King Saud Univ-Comput Inf Sci 26(3):258–275

Khattak A, Paracha WT, Asghar MZ, Jillani N, Younis U, Saddozai FK, Hameed IA (2020) Fine-grained sentiment analysis for measuring customer satisfaction using an extended set of fuzzy linguistic hedges. Int J Comput Intell Syst 13:744–756

Khatua A, Khatua A, Cambria E (2019) A tale of two epidemics: contextual word2vec for classifying twitter streams during outbreaks. Inf Process Manag 56(1):247–257

Khatua A, Khatua A, Cambria E (2020) Predicting political sentiments of voters from twitter in multi-party contexts. Appl Soft Comput 97:106743

Kristiyanti DA, Wahyudi M (2017) Feature selection based on genetic algorithm, particle swarm optimization and principal component analysis for opinion mining cosmetic product review. In: 2017 5th international conference on cyber and IT service management (CITSM). IEEE, pp 1–6

Kumarasiri C, Farook C (2018) User centric mobile based decision-making system using natural language processing (nlp) and aspect based opinion mining (abom) techniques for restaurant selection. In: Science and information conference. Springer, pp 43–56

Lau RY, Lai CC, Ma J, Li Y (2009) Automatic domain ontology extraction for context-sensitive opinion mining. In: ICIS 2009 proceedings, pp 35–53

Lee C, Lee GG (2006) Information gain and divergence-based feature selection for machine learning-based text categorization. Inf Process Manag 42(1):155–165

Lei Z, Yang Y, Yang M, Liu Y (2018) A multi-sentiment-resource enhanced attention network for sentiment classification. In: Proceedings of the 56th annual meeting of the association for computational linguistics, vol 2, pp 758–763

Li C, Xu B, Wu G, He S, Tian G, Hao H (2014) Recursive deep learning for sentiment analysis over social data. In: Proceedings of the 2014 IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT). IEEE Computer Society, pp 180–185

Li F, Huang M, Zhu X (2010a) Sentiment analysis with global topics and local dependency. AAAI Conf Artif Intell 10:1371–1376

Li H, Chen Z, Mukherjee A, Liu B, Shao J (2015a) Analyzing and detecting opinion spam on a large-scale dataset via temporal and spatial patterns. In: International conference on web and social media, pp 634–637

Li ST, Tsai FC (2013) A fuzzy conceptualization model for text mining with application in opinion polarity classification. Knowl-Based Syst 39:23–33

Li W, Chen H (2014) Identifying top sellers in underground economy using deep learning-based sentiment analysis. In: 2014 IEEE joint intelligence and security informatics conference (JISIC). IEEE, pp 64–67

Li X, Bing L, Li P, Lam W, Yang Z (2018) Aspect term extraction with history attention and selective transformation. In: Proceedings of the twenty-seventh international joint conference on artificial intelligence, pp 4198–4200

Li X, Wang B, Li L, Gao Z, Liu Q, Xu H, Fang L (2020) Deep2s: improving aspect extraction in opinion mining with deep semantic representation. IEEE Access 8:104026–104038

Li Y, Algarni A, Zhong N (2010b) Mining positive and negative patterns for relevance feature discovery. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, pp 753–762

Li Y, Algarni A, Xu Y (2011) A pattern mining approach for information filtering systems. Inf Retr 14(3):237–256

Li Y, Algarni A, Albathan M, Shen Y, Bijaksana MA (2015b) Relevance feature discovery for text mining. IEEE Trans Knowl Data Eng 27(6):1656–1669

Li Y, Zhang L, Xu Y, Yao Y, Lau RYK, Wu Y (2017) Enhancing binary classification by modeling uncertain boundary in three-way decisions. IEEE Trans Knowl Data Eng 29(7):1438–1451

Li Y, Wang S, Pan Q, Peng H, Yang T, Cambria E (2019) Learning binary codes with neural collaborative filtering for efficient recommendation systems. Knowl-Based Syst 172:64–75

Liddy ED (2001) Natural language processing. In: Encyclopedia of library and information science

Lin C, He Y (2009) Joint sentiment/topic model for sentiment analysis. In: Proceedings of the 18th ACM conference on information and knowledge management. ACM, pp 375–384

Liu B (2010) Sentiment Analysis and Subjectivity. In: Handbook of natural language processing, vol 2. pp 627–666

Liu B, Hu M, Cheng J (2005) Opinion observer: analyzing and comparing opinions on the web. In: Proceedings of the 14th international conference on World Wide Web. ACM, pp 342–351

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 Y, Zhai C (2008) Opinion integration through semi-supervised topic modeling. In: Proceedings of the 17th international conference on World Wide Web, pp 121–130

Luo Z, Osborne M, Wang T (2012) Opinionr trieval in twitter. In: Sixth international AAAI conference on weblogs and social media, pp 10–15

Ma Y, Peng H, Khan T, Cambria E, Hussain A (2018) Sentic lstm: a hybrid network for targeted aspect-based sentiment analysis. Cogn Comput 10(4):639–650

Mamgain N, Pant B, Mittal A (2016) Categorical data analysis and pattern mining of top colleges in india by using twitter data. In: 2016 8th international conference on computational intelligence and communication networks (CICN). IEEE, pp 341–345

Martineau J, Finin T et al (2009) Delta tfidf: an improved feature space for sentiment analysis. Int Conf Web Soc Media 9:106

Matsuno IP, Rossi RG, Marcacini RM, Rezende SO (2016) Aspect-based sentiment analysis using semi-supervised learning in bipartite heterogeneous networks. J Inf Data Manag 7(2):141–141

McAuley JJ, Leskovec J (2013) From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews. In: Proceedings of the 22nd international conference on World Wide Web. ACM, pp 897–908

Miao Z, Li Y, Wang X, Tan WC (2020) Snippext: semi-supervised opinion mining with augmented data. Proc Web Conf 2020:617–628

Mishra RK, Urolagin S, et al. (2019) A sentiment analysis-based hotel recommendation using tf-idf approach. In: 2019 international conference on computational intelligence and knowledge economy (ICCIKE). IEEE, pp 811–815

Moore J, Han EH, Boley D, Gini M, Gross R, Hastings K, Karypis G, Kumar V, Mobasher B (1997) Web page categorization and feature selection using association rule and principal component clustering. IBM shared research report/University of Minnesota (Minneapolis, Mn) 98:3

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

Morency LP, Mihalcea R, Doshi P (2011) Towards multimodal sentiment analysis: harvesting opinions from the web. In: Proceedings of the 13th international conference on multimodal interfaces. ACM, pp 169–176

Mukhopadhyay S (2018) Opinion mining in management research: the state of the art and the way forward. J OPSEARCH 55:1–30

Mukhtar N, Khan MA (2019) Effective lexicon-based approach for urdu sentiment analysis. Artif Intell Rev 53:1–28

Nadali S, Murad MA (2012) Fuzzy semantic classifier to determine the strength levels of customer product reviews. Proc Int Conf Adv Comput Sci Appl 2012:60–63

Nasukawa T, Yi J (2003) Sentiment analysis: capturing favorability using natural language processing. In: Proceedings of the 2nd international conference on Knowledge capture. ACM, pp 70–77

Nikfarjam A, Gonzalez GH (2011) Pattern mining for extraction of mentions of adverse drug reactions from user comments. In: AMIA annual symposium proceedings. American Medical Informatics Association, vol 2011, p 1019

Noekhah S, Salim NB, Zakaria NH (2017) A comprehensive study on opinion mining features and their applications. In: International conference of reliable information and communication technology. Springer, pp 78–89

Nóra BM, Lemnaru C, Potolea R (2010) Semi-supervised learning with lexical knowledge for opinion min-
    ing. In: Proceedings of the 2010 IEEE 6th international conference on intelligent computer communi-
    cation and processing. IEEE, pp 19–25

Novák V, Perfilieva I, Mockor J (2012) Mathematical principles of fuzzy logic, vol 517. Springer, Boston,
    MA

Oneto L, Bisio F, Cambria E, Anguita D (2017) Semi-supervised learning for affective common-sense rea-
    soning. Cogn Comput 9(1):18–42

Onwuegbuzie AJ, Frels R (2016) Seven steps to a comprehensive literature review: a multimodal and cul-
    tural approach. SAGE Publications Ltd., Thousand Oaks

Paltoglou G, Thelwall M (2010) A study of information retrieval weighting schemes for sentiment analysis.
    In: Proceedings of the 48th annual meeting of the association for computational linguistics. Associa-
    tion for Computational Linguistics, pp 1386–1395

Panchendrarajan R, Ahamed N, Murugaiah B, Sivakumar P, Ranathunga S, Pemasiri A (2016) Implicit
    aspect detection in restaurant reviews using cooccurence of words. In: Proceedings of the 7th work-
    shop on computational approaches to subjectivity, sentiment and social media analysis, pp 128–136

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up?: sentiment classification using machine learning tech-
    niques. In: Proceedings of the ACL-02 conference on empirical methods in natural language process-
    ing. Association for Computational Linguistics, pp 79–86

Pang B, Lee L et al (2008) Opinion mining and sentiment analysis. Found Trends® Inf Retr 2(1–2):1–135

Pansare V (2016) Effecive pattern identification approach for text mining. Int J Comput Sci Inf Technol
    7:1826–1830

Park SM, Lee SJ, On BW (2020) Topic word embedding-based methods for automatically extracting main
    aspects from product reviews. Appl Sci 10(11):3831

Pasquier N, Bastide Y, Taouil R, Lakhal L (1999) Discovering frequent closed itemsets for association rules.
    In: International conference on database theory. Springer, pp 398–416

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

Phu VN, Chau VTN, Tran VTN (2017) Svm for english semantic classification in parallel environment. Int
    J Speech Technol 20(3):487–508

Pimpalkar A, Wandhe T, Rao MS, Kene M (2014) Review of online product using rule based and fuzzy
    logic with smiley's. Int J Comput Technol 1(1):39–44

Piryani R, Madhavi D, Singh VK (2017) Analytical mapping of opinion mining and sentiment analysis
    research during 2000–2015. Inf Process Manag 53(1):122–150

Poria S, Cambria E, Ku LW, Gui C, Gelbukh A (2014) A rule-based approach to aspect extraction from
    product reviews. In: Proceedings of the second workshop on natural language processing for social
    media (SocialNLP), pp 28–37

Poria S, Cambria E, Gelbukh A, Bisio F, Hussain A (2015) Sentiment data flow analysis by means of
    dynamic linguistic patterns. IEEE Comput Intell Mag 10(4):26–36

Poria S, Cambria E, Gelbukh A (2016a) Aspect extraction for opinion mining with a deep convolutional
    neural network. Knowl-Based Syst 108:42–49

Poria S, Chaturvedi I, Cambria E, Bisio F (2016b) Sentic lda: improving on lda with semantic similarity for
    aspect-based sentiment analysis. In: 2016 international joint conference on neural networks (IJCNN).
    IEEE, pp 4465–4473

Prakash S, Chakravarthy T, Kaveri E (2015) Statistically weighted reviews to enhance sentiment classifica-
    tion. Karbala Int J Mod Sci 1(1):26–31

Priyadarshi R, Rawat S, Kumar P (2014) An implementation of opinion mining using fuzzy inference sys-
    tem. In: Innovative applications of computational intelligence on power, energy and controls with
    their impact on Humanity (CIPECH). IEEE, pp 127–132

Quan C, Ren F (2014) Unsupervised product feature extraction for feature-oriented opinion determination.
    Inf Sci 272:16–28

Quan T, Hui S, Cao T (2004) A fuzzy fca-based approach for citation-based document retrieval. In: 2004
    IEEE conference on cybernetics and intelligent systems, vol 1. IEEE, pp 578–583

Rana TA, Cheah YN (2017) A two-fold rule-based model for aspect extraction. Expert Syst Appl
    89:273–285

Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applica-
    tions. Knowl-Based Syst 89:14–46

Ravi K, Ravi V, Prasad PSRK (2017) Fuzzy formal concept analysis based opinion mining for crm in finan-
    cial services. Appl Soft Comput 60:786–807

Recupero DR, Presutti V, Consoli S, Gangemi A, Nuzzolese AG (2015) Sentilo: frame-based sentiment analysis. Cogn Comput 7(2):211–225

Robertson S, Zaragoza H et al (2009) The probabilistic relevance framework: Bm25 and beyond. Found Trends® Inf Retr 3(4):333–389

Sadhana S, SaiRamesh L, Sabena S, Ganapathy S, Kannan A (2017) Mining target opinions from online reviews using semi-supervised word alignment model. In: 2017 second international conference on recent trends and challenges in computational models (ICRTCCM). IEEE, pp 196–200

Sadidpour S, Shirazi H, Sharef NM, Minaei-Bidgoli B, Sanjaghi ME (2016) Context-sensitive opinion mining using polarity patterns. Int J Adv Comput Sci Appl (IJACSA) 7:146–150

Saleena N et al (2018) An ensemble classification system for twitter sentiment analysis. Procedia Comput Sci 132:937–946

Saleh MR, Martín-Valdivia MT, Montejo-Ráez A, Ureña-López L (2011) Experiments with svm to classify opinions in different domains. Expert Syst Appl 38(12):14799–14804

Samha AK, Li Y, Zhang J (2014) Aspect-based opinion extraction from customer reviews, arXiv preprint arXiv:1404.1982

Sarkar SD, Goswami S, Agarwal A, Aktar J (2014) A novel feature selection technique for text classification using naive bayes. Int Sch Res Not 2014:1–10

Severyn A, Moschitti A (2015) Twitter sentiment analysis with deep convolutional neural networks. In: Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval. ACM, pp 959–962

Sharma R, Raman S (2003) Phrase-based text representation for managing the web documents. In: Proceedings ITCC 2003. International conference on information technology: coding and computing. IEEE, pp 165–169

Sharma R, Nigam S, Jain R (2014) Opinion mining of movie reviews at document level. Int J Inf Theory 3:13–21

Shin B, Lee T, Choi JD (2016) Lexicon integrated cnn models with attention for sentiment analysis. In: Proceedings of the 8th workshop on computational approaches to subjectivity, sentiment and social media analysis

Shinde MR, Gill PC (2014) Pattern discovery techniques for the text mining and its applications. Int J Sci Res (IJSR) 3(5):1660–1664

Shokeen J, Rana C (2020) A study on features of social recommender systems. Artif Intell Rev 53(2):965–988

Spasic I, Williams L, Buerki A (2017) Idiom based features in sentiment analysis: cutting the gordian knot. IEEE Trans Affect Comput 11:189–199

Subhashini L, Li Y, Zhang J, Athukorale A (2018) Opinion classification using pattern mining and fuzzy logic. In: 2018 18th international conference on advances in ICT for emerging regions (ICTer). IEEE, pp 1–7

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

Tang D, Qin B, Liu T, Yang Y (2015) User modeling with neural network for review rating prediction. In: Twenty-fourth international joint conference on artificial intelligence, pp 1340-1346

Tay Y, Tuan LA, Hui SC (2017) Dyadic memory networks for aspect-based sentiment analysis. In: Proceedings of the 2017 ACM on conference on information and knowledge management. ACM, pp 107–116

Titov I, McDonald R (2008) Modeling online reviews with multi-grain topic models. In: Proceedings of the 17th international conference on World Wide Web. ACM, pp 111–120

Tsirakis N, Poulopoulos V, Tsantilas P, Varlamis I (2017) Large scale opinion mining for social, news and blog data. J Syst Softw 127:237–248

Turney PD (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. Association for Computational Linguistics, pp 417–424

Van Hee C, Lefever E, Hoste V (2018) Exploring the fine-grained analysis and automatic detection of irony on twitter. Lang Resour Eval 52:1–25

Vateekul P, Koomsubha T (2016) A study of sentiment analysis using deep learning techniques on thai twitter data. In: 13th international joint conference on computer science and software engineering (JCSSE). IEEE, pp 1–6

Vechtomova O (2010) Facet-based opinion retrieval from blogs. Inf Process Manag 46(1):71–88

Vechtomova O (2017) Disambiguating context-dependent polarity of words: an information retrieval approach. Inf Process Manag 53(5):1062–1079

Vijayarani S, Ilamathi MJ, Nithya M (2015) Preprocessing techniques for text mining-an overview. Int J Comput Sci Commun Netw 5(1):7–16

Vinodhini G, Chandrasekaran R (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

Vinodhini G, Chandrasekaran R (2017) A sampling based sentiment mining approach for e-commerce applications. Inf Process Manag 53(1):223–236

Vo B, Le T, Nguyen G, Hong TP (2017) Efficient algorithms for mining erasable closed patterns from product datasets. IEEE Access 5:3111–3120

Wang W, Pan SJ, Dahlmeier D (2018) Memory networks for fine-grained opinion mining. Artif Intell 265:1–17

Wang Y, Huang M, Zhao L, et al. (2016) Attention-based lstm for aspect-level sentiment classification. In: Proceedings of the 2016 conference on empirical methods in natural language processing, pp 606–615

Whissell JS, Clarke CL (2011) Improving document clustering using okapi bm25 feature weighting. Inf Retr 14(5):466–487

Wu ST, Li Y (2013) Pattern-based web mining using data mining techniques. Int J e-Educ, e-Bus, e-Manag e-Learn 3(2):163

Wu ST, Li Y, Xu Y, Pham B, Chen P (2004) Automatic pattern-taxonomy extraction for web mining. In: Proceedings of the IEEE/WIC/ACM international conference on web intelligence (WI). IEEE, pp 242–248

Wu Y, Li Y, Xu Y (2019) Dual pattern-enhanced representations model for query-focused multi-document summarisation. Knowl-Based Syst 163:736–748

Wu Z, Dai XY, Yin C, Huang S, Chen J (2018) Improving review representations with user attention and product attention for sentiment classification. In: Thirty-second AAAI conference on artificial intelligence, pp 1-9

Xia R, Jiang J, He H (2017) Distantly supervised lifelong learning for large-scale social media sentiment analysis. IEEE Trans Affect Comput 8(4):480–491

Xia Y, Cambria E, Hussain A, Zhao H (2015) Word polarity disambiguation using bayesian model and opinion-level features. Cogn Comput 7(3):369–380

Xing FZ, Cambria E, Zou X (2017) Predicting evolving chaotic time series with fuzzy neural networks. In: 2017 international joint conference on neural networks (IJCNN). IEEE, pp 3176–3183

Xu R, Gui L, Xu J, Lu Q, Wong KF (2015) Cross lingual opinion holder extraction based on multi-kernel svms and transfer learning. World Wide Web 18(2):299–316

Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 conference of the north american chapter of the association for computational linguistics: human language technologies, pp 1480–1489

Yatsko V (2013) Advantages and disadvantages of tf* idf term weighting. World Sci Discov 2(42):10–15

You ZH, Hu YH, Tsai CF, Kuo YM (2020) Integrating feature and instance selection techniques in opinion mining. Int J Data Warehous Min (IJDWM) 16(3):168–182

Yousif A, Niu Z, Tarus JK, Ahmad A (2019) A survey on sentiment analysis of scientific citations. Artif Intell Rev 52(3):1805–1838

Yu J, Zha ZJ, Wang M, Wang K, Chua TS (2011) Domain-assisted product aspect hierarchy generation: towards hierarchical organization of unstructured consumer reviews. In: Proceedings of the conference on empirical methods in natural language processing. Association for Computational Linguistics, pp 140–150

Yu N, Kubler S (2010) Semi-supervised learning for opinion detection. In: 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, vol 3. IEEE, pp 249–252

Zadeh LA (1996) Fuzzy sets. Fuzzy sets, fuzzy logic. And fuzzy systems, Selected papers by Lotfi A Zadeh, World Scientific, pp 394–432

Zhang G, Xu L, Wang L (2018) Sentiments classification in stock network public opinion space based on long-short memory convolution neural network. In: MATEC web of conferences, vol 189. EDP Sciences, p 10010

Zhang H, Sekhari A, Ouzrout Y, Bouras A (2016a) Jointly identifying opinion mining elements and fuzzy measurement of opinion intensity to analyze product features. Eng Appl Artif Intell 47:122–139

Zhang K, Narayanan R, Choudhary AN (2010) Voice of the customers: mining online customer reviews for product feature-based ranking. In: Proceedings of the 3rd conference on online social networks, pp 1–9

Zhang Y, Er MJ, Venkatesan R, Wang N, Pratama M (2016b) Sentiment classification using comprehensive attention recurrent models. In: 2016 international joint conference on neural networks (IJCNN). IEEE, pp 1562–1569

Zhang Y, Wu J, Wang H (2019) Neural binary representation learning for large-scale collaborative filtering. IEEE Access 7:60752–60763

Zhang Z, Wang R (2014) Applying three-way decisions to sentiment classification with sentiment uncertainty. In: International conference on rough sets and knowledge technology. Springer, pp 720–731

Zhao H, Xia Y, Lau RY, Liu Y (2012) Word sentiment polarity disambiguition based on opinion level context. In: 2012 international conference on machine learning and cybernetics, vol 5. IEEE, pp 2007–2012

Zheng S, Zhou Y, Martin T (2009) A new method for fuzzy formal concept analysis. In: Proceedings of the 2009 IEEE/WIC/ACM international joint conference on web intelligence and intelligent agent technology, vol 3. IEEE Computer Society, pp 405–408

Zhong N, Li Y, Wu ST (2012) Effective pattern discovery for text mining. IEEE Trans Knowl Data Eng 24(1):30–44

Zhou S, Chen Q, Wang X (2010) Active deep networks for semi-supervised sentiment classification. In: Coling 2010: posters, pp 1515–1523

Zhou X, Xu Y, Li Y, Josang A, Cox C (2012) The state-of-the-art in personalized recommender systems for social networking. Artif Intell Rev 37(2):119–132

Zimmermann M, Ntoutsi E, Spiliopoulou M (2015) Discovering and monitoring product features and the opinions on them with opinstream. Neurocomputing 150:318–330