Chapter 13
Recent Developments in Sentiment Analysis on Social Networks: Techniques, Datasets, and Open Issues

Akrati Saxena, Harita Reddy, and Pratishtha Saxena

Abstract In recent years, sentiment analysis has been highly used on social media datasets to get conclusive information, opinions of users about different topics, such as politics, events, and products, and predicting users' mood or emotions. Marketing companies are very interested in finding out people's opinions and consider that in designing marketing strategies. However, filtering and monitoring the sentiments remain a formidable task due to the proliferation of diverse websites. Researchers have proposed numerous methods for automating the process of filtering and summarizing users' sentiment and opinion on social media. In this chapter, we present a systematic review of sentiment analysis techniques, including machine learning-based, deep learning-based, lexicon-based, graph-based, and hybrid approaches. We further provide a concise list of openly available datasets that have been used for research studies. Finally, we cover future directions and the respective challenges that should be looked at for further enhancing the sentiment analysis for microblog datasets of social networks.

13.1 Introduction

Social media websites furnish a platform for building online social relationships and sharing news, opinions, as well as multimedia to a global audience. With a huge volume of posts containing opinions of people across the world on various topics, the content has been used for insights by both academic researchers and companies

A. Saxena (✉)
Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: a.saxena@tue.nl

H. Reddy
Surat, Gujarat, India

P. Saxena
G. L. Bajaj Institute of Technology and Management, Greater Noida, India
e-mail: pratistha.saxena@glbitm.ac.in

for opinion mining [1]. Sentiment analysis involves analyzing people's sentiment, opinion, or attitudes toward a particular entity, which might include products, services, policies, or celebrities [2]. A common application of sentiment analysis in the industry is to plan marketing strategies based on public sentiment and gauge public sentiment in response to a new product launch. Researchers have shown great interest in proposing novel techniques for classifying the sentiment of online content.

The sentiment classification approaches in previous works can be classified into machine learning-based and lexicon-based methods. However, inadequate contextual information in small paragraphs and individual sentences could pose challenges despite these methods. The availability of abundant data and computational power has paved the way for exploring deep learning solutions that utilize prior information for sentiment analysis. Tasks requiring complex features for sentiment classification and contextual polarity disambiguation in tweets make use of deep convolutional neural networks [3, 4]. In this chapter, we will discuss all these categories of sentiment analysis techniques in depth. In a data-driven model, it becomes imperative to analyze how data is collected, perceived, and communicated. We will also cover various datasets that have been used for training and evaluating sentiment analysis models in the past research. Furthermore, there are multiple directions on which researchers can focus in the future, which we discuss toward the end of the chapter.

The chapter is structured in the following manner. In Sect. 13.2, we present a systematic review of sentiment analysis techniques, including machine learning-based techniques, lexicon-based techniques, hybrid techniques, graph-based techniques, and other techniques. In Sect. 13.3, we list the available datasets. In Sect. 13.4, we discuss the future directions of sentiment analysis in the context of large-scale data-rich social networks. The chapter is concluded in Sect. 13.5.

13.2 Sentiment Analysis Techniques

An important task in sentiment analysis is the selection of the technique for achieving high accuracy in classifying the sentiment of the content. There are three broad categories of techniques, i.e., machine learning (ML)-based, lexicon-based, and hybrid techniques, as shown in Fig. 13.1. Using these approaches, a given text document is assigned a suitable sentiment class: positive, negative, or neutral. ML techniques involve the usage of ML algorithms, for example, Naive Bayes Classifier, k-Nearest Neighbors Classifier, and Support Vector Machines (SVM), along with the linguistic features of the document. This class of techniques can be subdivided into supervised, weakly supervised, semi-supervised, and unsupervised learning techniques. For supervised learning, a dataset containing the text documents annotated with the ground truth sentiment labels is required for training the classification model before the model can identify the sentiment of new documents. Unsupervised learning is used when labeled documents are not available. Lexicon-based techniques involve the use of a sentiment dictionary or lexicon containing opinion words to identify the sentiment toward which a document is oriented. Examples of lexicon-based tech-
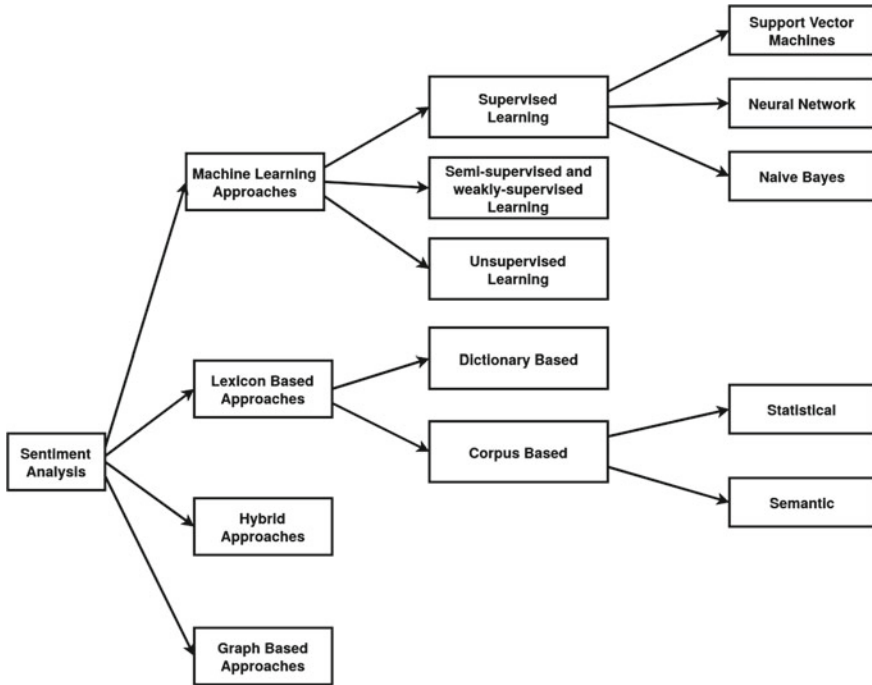
Fig. 13.1 Overview of Sentiment Analysis Techniques

niques are dictionary-based and corpus-based techniques. In the dictionary-based techniques, a pre-existing dictionary containing a sentiment word list is used to find the presence of such words within the text and accordingly determine the sentiment expressed in the document based on counting. The corpus-based approaches additionally rely on contextual information with statistical and semantic methods. Hybrid approaches are usually a fusion of the discussed techniques. Another class of techniques that we discuss is graph-based approaches. In the upcoming sections, we will discuss the sentiment analysis techniques in detail.

13.2.1   *Machine Learning Approaches*

Machine Learning (ML)-based approaches use an existing dataset for learning how to determine the class, with the dataset being split into training and test sets. The training set comprises features of input data samples annotated with appropriate labels indicating the class to which the sample belongs. The classification model is trained using this training set, such that it learns the required features based on the expected output class labels given the feature vectors of the input samples. Based on

the learning on the training set, the model should be able to predict the class labels for new data samples that were not a part of the training set. The performance of the classifier on unseen samples is validated using the test set. The test set samples are given as input, and the classifier predicts the output class, which is verified against the ground truth. For sentiment analysis, the input samples are text documents whose sentiment needs to be determined. ML-based approaches do not rely on a prior lexicon, but leverage the classification algorithms, including the Naive Bayes Classifier (NB) and Support Vector Machine (SVM) [5]. Some of the common features used as an input to ML classifiers are part-of-speech (POS), presence of terms, frequency of terms, n-grams, and negation [6]. In ML approaches, the model learns by using either supervised, unsupervised, or hybrid techniques. These techniques are discussed in the following subsections.

Supervised Learning. Several works in the literature have used supervised learning for sentiment classification. These classifiers are developed with the use of labeled training documents considering sentiment classification as a standard statistical classification. The predictions on new data are made on the basis of the patterns observed in the training set. In the following sections, we will discuss some of the mainly used classification methods.

A. Support Vector Machines (SVM). SVM is an ML technique that performs binary classification by determining the hyper-plane that achieves the best separation of the data that belongs to either of the two classes. The SVM classifier involves the representation of the input samples in the space as points, such that the samples belonging to different classes are separated as much as possible [7]. This algorithm separates text documents into positive and negative sentiment classes. A number of studies on sentiment classification of social media texts with supervised learning have used the SVM algorithm [8–12]. The SVM becomes suitable because the features of a text document are correlated and are linearly separable [13].

Sentiment classification with SVM has been used with a variety of features. Shein et al. [14] proposed a combined ontology-based and Natural Language Processing (NLP)-based approach in which a 3-class classification was done using a linear SVM. Luo et al. [7] compared the results obtained through SVM with the results of naive Bayes (NB) and k-nearest neighbors (KNN) algorithms in the 2-class classification of text documents represented in a vector space model. They noted that SVM outperformed the other two classifiers irrespective of the technique used for feature selection. The one-vs-one (OVO) technique is a common way to convert a multi-class classification problem to multiple binary classification problems. Liu et al. [15] used this strategy along with SVM to obtain better results for multi-class classification compared to the existing techniques.

Balahur [16] proposed a framework using unigram and bigram features from Twitter data with SVM Sequential Minimal Optimization (SVM SMO) [17]. Using SVM SMO helps in overcoming the risk of overfitting on data. The framework was extensible for use in other languages due to less processing of linguistic features and was also suitable for real-time processing. Kiritchenko et al. [18] designed a linear kernel SVM-based architecture to analyze the sentiment of short texts because

the linear-kernel-based SVM performed better relative to the Radial Basis Function (RBF)-based kernel. They extracted sentiment features from the posts containing sentiment-word, hashtags, and emoticons. Their framework obtained the first rank in SemEval-2013 task 2.

The performance of SVM for the sentiment analysis task can be enhanced with the help of the grid search technique for optimizing hyper-parameters [19]. This involves the evaluation of multiple models with varying parameters. Ahmad et al. [19] noted the improved performance of the SVM classifier through this optimization on Twitter and IMDB datasets. Sharma et al. [20] enhanced the performance of SVM by combining it with Boosting algorithm, which is an ensemble algorithm. It can help in improving the ability to generalize due to the selection of the best-performing features at every step. Based on their evaluation of movie and hotel review datasets, it was shown that using an ensemble of boosting along with SVM outperforms an individual SVM classifier.

B. Naive Bayes. The Naive Bayes classifier is frequently applied, especially due to the simplicity in its assumption that the input features are independent. Despite this assumption, this classifier performs reasonably well as compared to other classifiers and has been proved to be useful in use cases, such as text classification [21]. The family of classifiers relies on the Bayes theorem that for a given sample, the class that is most likely to satisfy the input vector is assigned to the sample.

Naive Bayes classifier has been used for real-time sentiment classification of tweets due to its speed [22]. Pak and Paroubek [23] used a multinomial Naive Bayes classifier for classifying Twitter data into three classes: positive, neutral, and negative and it outperformed SVM. They used both n-gram-based features and POS tagging-based features. They also noted that some of the POS tags are good for inferring emotion. Troussas et al. [24] used the Naive Bayes classifier for sentiment analysis of status updates of Facebook users and obtained high accuracy. They showed that Naive Bayes had comparable performance with the Rocchio classifier but with a lower recall. In another work, it was shown that the Naive Bayes classifier outperformed the Maximum Entropy model for sentiment classification of updates on Twitter by users [25]. Kang and Yoo [26] proposed an enhanced Naive Bayes algorithm to address the problem of imbalance between the positive and negative classification accuracy and noted that the algorithm decreased the gap between both accuracies on restaurant review data.

Gamallo et al. [27] proposed two variations of the classifier for their work on tweets, in which the first variant included a classifier trained for three-class classification. The second variant included training the classifier on a corpus after eliminating neutral tweets and thus doing only binary classification. If the tweet has no word that belongs to a polarity lexicon, the tweet was termed as neutral. This binary classifier achieved above 80% precision on the data with two classes.

In another novel approach, Tan et al. [28] adapted Naive Bayes to a new domain for sentiment classification. They used both the previous domain knowledge along with the knowledge from the new domain data, which is unlabeled, thus treating the problem as a domain-transfer problem. They proposed a measure, called Frequently

Co-occurring Entropy, which extracts features occurring commonly in the old as well as the new domain. The features obtained from the old domain are generalizable features. They further constructed an Adapted Naïve Bayes classifier that can leverage the information from a new domain. It increases the weight given to the new domain and reduces the weight given to the old domain during the iteration process. They have shown that this modified classifier outperforms the base classifier and transfer-learning methods.

C. Artificial Neural Networks (ANN). ANNs are somewhat motivated by the neurological connections in the human brain. They contain artificial neurons, which, analogous to a brain neuron, are processing units and are interconnected to each other. The neurons in a particular layer receive the output of the neurons in the previous layer as input, and these inputs carry different weights. After processing the input, the neurons give an output to the neurons in the next layer. The weights associated with the connections among the neurons are fine-tuned based on the data that is being used to train the network. The training dataset, for which the final expected output labels are known, is used for making the network learn the appropriate weights for the connections between the neurons. Neural networks often contain multiple layers, which are also called 'multi-layer' neural networks that consist of hidden intermediate layers which are not directly connected to the input that we provide and the output that the neural network provides. These multi-layer networks are commonly used for non-linear classification problems.

Moraes et al. [29] compared ANNs and SVM for assigning the appropriate sentiment class to a text document. Using the traditional bag-of-words model for word vectorization of text coupled with ANNs resulted in an equal or better performance as compared to SVM when the datasets were not unbalanced. Duncan and Zhang [30] used a feedforward neural network in their work on Twitter sentiment analysis and noted that they faced memory issues when they tried to train the network with large vocabulary derived from a large training set of tweets. They suggested that feature reduction methods can be used for reducing the vocabulary used for training the network. However, advances in computational hardware and the increased availability of large datasets for training ANNs had rekindled the attention given to ANN architectures, especially with a focus on deep neural networks that have multiple layers [2]. Deep networks usually take the vectorized form of the text documents as input, also called word embeddings. These embeddings denote the latent features of the words in the document, and can be learned using ANNs or matrix factorization.

There are several classes of ANNs, a common one being Convolutional Neural Networks (CNN). CNNs are feedforward ANNs that comprise multiple layers with convolving filters that capture the features locally and are heavily used for image classification [31]. Bhargava et al. [32] used ANNs for the sentiment identification of Indian languages' tweets, namely, Hindi, Bengali, and Tamil. Their aim was to assess if neural network architectures outperformed the conventional machine learning models and how the hidden layers affect the performance of neural networks. They created a number of sequential models using a combination of layers based on CNNs, Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM). It was

noted that the increase in the complexity of the language used in the text resulted in reduced accuracy and more hidden layers resulted in enhanced accuracy.

Some works focus on classifying the sentiment of individual sentences, also called sentence-level classification. As in the case of document-level sentiment analysis, ANNs are used to generate representations of sentences. However, since sentences contain lesser information as compared to a document due to their short lengths, additional features such as parse trees and part of speech tags may be incorporated to get good results [2]. Instead of relying on parse trees for syntax and semantics, newer models based on CNNs and RNNs leverage word embeddings that encode this information. Socher et al. [33] used recursive autoencoders based on vector embeddings for variable-length phrases. There are various implementations of ANNs for the sentence-level sentiment classification [33–39].

Short text blogs, which are common on social media, often have limited context and features, which makes it challenging to classify their sentiment. CNNs have been incorporated for sentiment analysis to overcome such challenges. Santos and Gatti [3] used CNNs for obtaining character, word, and subsequently, sentence-based features for movie reviews and tweets. Wang et al. [40] worked on a combined architecture with CNN and RNN for identifying the sentiment in short texts. While CNNs can capture local-level features, RNNs are suitable for capturing long-distance dependencies in data modeled as a sequence. Deep learning methods generally need a large amount of training data for learning accurately. To overcome this limitation, Guan et al. [41] proposed a weak supervised learning method for learning the initial sentence representations followed by the supervised approach based on labeled data for sentiment classification of sentences.

D. Other Supervised Learning Methods. Another class of algorithms is genetic algorithms, which are inspired by the evolution process in biology [42]. Govindarajan proposed an approach using a hybrid machine learning classification model through an ensemble of genetic algorithm and the naive Bayes classifier to obtain a more robust model with a better accuracy [43]. The hybrid classifier gives a better classification accuracy in sentiment analysis on a dataset of movie reviews. The coupling of the methods is done through an arcing classifier. The choice of the two classification techniques was based on their heterogeneity and strengths, and the hybrid approach leveraged the advantages of both the classifiers. While individually, the genetic classifier outperforms Naive Bayes, the combined classifier outperforms the individual classifiers.

Semi-Supervised and Weakly Supervised Learning. In supervised learning techniques, an essential step is the collection of data annotated with their class labels for training. However, for some domains, it is not possible to get an adequate volume of labeled data for training the classifier. In such cases, a practical solution is to use a small set of labeled documents along with unlabeled documents, which may be easily available in large numbers.

Li et al. [44] used a semi-supervised approach to account for use cases where the data is unbalanced between the two sentiment classes. They used iterative under-sampling, and dynamic random subspace generation [45]. Their proposed methods

successfully used the unlabeled data, and dynamically generating the subspaces did better than a static generation. The method also outperformed baseline techniques in data pertaining to four domains. Sintsova et al. [46] used the knowledge obtained from an emotion lexicon to create a classifier for sports-related tweets on Twitter. In the semi-supervised approach, a classifier is initially used to derive pseudo-labels from the unlabeled data, after which the labels are further refined. They used the balanced weighted voting for handling unbalanced labels in the data, and this was shown to enhance the classifier performance.

The Expectation-Maximization (EM) algorithm is also commonly used in semi-supervised learning, especially, for cases where the data has missing labels [47, 48]. Zhai et al. [48] used the EM algorithm for determining the groups to which product features belong. In this approach, the labeled data is first used to train the classifier for obtaining probabilistic labels for the unlabeled samples. Soft-labels created for the set of unlabeled samples using constraints are also used for the initial classifier. After the completion of this step, another classifier is trained with labeled data, as well as unlabeled data with probabilistic labels.

Weak supervision is the broad topic that covers three types of learning: (i) Incomplete supervision, in which only a small number of samples are labeled in the training dataset, (ii) Inexact supervision, in which the labels in the dataset are more abstract or coarse-grained, and (iii) Inaccurate supervision, in which some samples may have wrong labels [49]. He and Zhou [50] proposed a framework in which a pre-existing sentiment lexicon containing words and their respective sentiment is used for training the initial classifier with generalized expectation. The classifier is then applied to unlabeled documents to get sentiment annotations. Even without using labeled data, their proposed approach was shown to perform as good or better compared to the already present weak supervision techniques.

Unsupervised Learning. Unsupervised learning, in which there are no class labels available for input data, has also been adopted for sentiment analysis. Since we do not have pre-existing class labels, the evaluation of unsupervised learning models is relatively difficult, and the decision of choosing a particular model cannot be easily explained. The aim of unsupervised learning techniques is to determine any regularities present in input samples. The prediction is done based on the structures of the input space in which certain patterns occur more often than others. In other words, these techniques utilize sentiment-based patterns to determine the labels for words/phrases.

Hu et al. [51] incorporated emotional signals for unsupervised sentiment analysis. The authors gave examples of ratings and emoticons as emotional signals in posts made by users expressing their sentiment. They work with two types of signals, (i) emotion indication and (ii) correlation. Emotion indication includes signals, such as emoticons, ratings, and stars, which give a strong indication of the conveyed sentiment. The basic idea of emotion correlation is that words often co-occurring together are more probable to have similar emotions. Using such emotional signals, the authors implemented matrix factorization method for unsupervised learn-

ing approach on publicly available Twitter and debate datasets. They noted that their unsupervised framework performed better than the baselines on these two datasets.

13.2.2 Lexicon-Based Approaches

Lexicon-based approaches utilize a sentiment lexicon for sentiment classification. This is usually done by taking a count or weighting the opinion words that are present in the text that is being analyzed [52]. Such techniques rely on the lexical resource and do not require to be trained on a training dataset. A lexicon can contain words, idioms, or phrases annotated with their sentiment polarities. A neutral or objective polarity can also be included in a lexicon. The opinion words lexicon can be constructed using a manual approach but is not time-effective. Usually, manual approaches may be used to complement automated techniques for any correction of mistakes. The common approaches used for getting the opinion words are dictionary-based and corpus-based approaches. The quality of these lexical resources contributes to the effectiveness of the sentiment classification.

Natural languages can be complex, and therefore lexicon-based approach cannot account for some characteristics of language such as slang, sarcasm, and negation [53]. Some of the difficulties of using such approaches are (i) a word's meaning can depend on their application in the sentence, (ii) sentiment words in certain contexts may not express polarity, and (iii) sentences may not include any sentiment words when expressing an opinion [54]. However, the simplicity of the lexicon-based approaches is one advantage. Enumerating the positive and negative words is easy, and the approaches also give the flexibility of adjusting to different languages and giving fast results. Such approaches are suitable when the structure of the data is complex, the data is less, and there is limited time to obtain the results. However, we can enhance the quality of sentiment analysis by using lexicon along with machine learning techniques. One way of creating an opinion lexicon is using a manual approach, but it is inefficient. Therefore, usually, automated approaches are used first, and then manual approaches are used to finally check any mistakes. The automated approaches are described in the upcoming subsections.

Dictionary-Based Approaches. In dictionary-based approaches, a seed list is initially created with opinion words through a manual approach. The sentiment polarity of these words is determined, and these words can be matched against data to be analyzed to find the text polarity [55]. The collection is then extended by extracting antonyms and synonyms of these words using a lexical resource, such as WordNet [56], SentiWordNet, and a thesaurus [57], and appropriately ascertaining the polarity of newly discovered words [58]. The new words are added to the collection through an iterative process, and finally, any errors may need to be corrected manually. An example of lexical resource is SentiWordNet, which has different versions and contains annotations for WordNet synsets based on the polarity [59].

Li et al. [60] proposed a framework related to stock predictions, which included sentiment analysis of news articles for which they used sentiment dictionaries. The sentiment dictionary is used to map the news articles into the suitable sentiment space. First, the news is vectorized in the form of term frequencies. A separate matrix that represents the sentiment dictionary is multiplied with the vectorized news, after which the news is labeled as positive, neutral, or negative based on thresholds. In another work, Filho et al. [61] evaluated the LIWC dictionary for sentiment determination of Brazilian Portuguese texts. They compared the performance against other Portuguese sentiment analysis lexical resources, including SentiLex and Opinion Lexicon and noted that LIWC performed better in detecting positive texts, however, SentiLex worked better with negative ones.

Other works using sentiment dictionary include the one by Qiu and He [62], in which they proposed an approach for advertising that leverages sentiment analysis to determine users' dissatisfaction toward a certain topic. In this approach, a pre-existing dictionary containing positive and negative words has been used to determine opinionated sentences that consist of sentiment words. Based on this, the topic words toward which the users express negative polarity are extracted. This enables us to know the product features of a brand toward which the users' have expressed negative sentiment and then display rival products' advertisements accordingly. One disadvantage of the dictionary-based approach is the inability to capture opinion words that have polarity specific to a certain domain or context.

Corpus-Based Approaches. The collection of positive and negative polarity words can also be generated using a text corpus based on an existing seed list of polarity words and the syntactic and co-occurrence patterns present in the chosen corpus [63]. This approach is helpful in incorporating domain-specific sentiment words and informal slang terms used online but which are not present in a normal dictionary [64]. A statistical or semantic approach is usually implemented with the corpus-based approach. We first discuss the work on corpus-based lexicon creation, followed by a detailed discussion on the semantic and statistical approaches.

Moreno-Ortiz and Fernández-Cruz [65] utilized both the specialized language corpus and the generic corpus to obtain domain-specific opinion words. They gave special attention to those words that have a different semantic orientation in a domain as compared to their orientation in a generic language. With this approach, they were able to obtain specialized terms with domain-specific polarity. One approach to constructing a lexicon is leveraging graph propagation. Velikovich et al. [66] used a graph propagation algorithm in which the graph nodes are the possible candidates for the polarity lexicon and the edges between representing their semantic similarity. The graph is also initialized with a seed list of polarity words with the aim of propagating the information to other nodes from the seeds. The constructed graph was based on n-grams obtained from 4 billion pages on the Web. They evaluated their resultant sentiment lexicon and noted that the lexicon contained phrases, such as spelling variants and slang, and also performed better compared to the existing lexicons for classification of the polarity of a sentence. In another approach, Hatzivassiloglou and McKeown [67] leveraged the conjunctions between the adjectives to extract

semantic orientation of adjectives. Usually, conjunctions connect adjectives with similar orientations, such as 'beautiful and intelligent', except for the case of 'but' where the orientations of the adjectives are dissimilar. They used a Wall Street Journal corpus and log-linear regression to extract the orientation similarity or dissimilarity between the adjectives to construct a graph. Finally, clustering is used to separate the adjectives into the polarity classes.

In a given context, sentences with similar polarities occur near each other. Kanayama and Nasukawa [68] proposed a framework to determine entities that convey the sentiment in a given clause, for which they used intra-sentential context and inter-sentential context on a Japanese corpus. Their framework enabled enlarging of a sentiment lexicon for domain-specific sentiment analysis by identifying polar phrases specific to a domain. Sentiment lexicon can be generated by leveraging social media, which is becoming a common platform where people reveal their sentiments on different topics. Feng et al. [69] compared the usage of Twitter, Wikipedia, and the Web corpora for calculating sentiment similarity, which is a crucial step for determining polarity of words. They noted that Twitter is a good resource for measuring sentiment similarity and that the incorporation of emoticons in the seed list can enhance performance.

The following sections discuss statistical and semantic approaches in detail.

A. Statistical Approaches. Statistical techniques are used to determine co-occurrences of words/phrases. The polarity of a given adjective can depend on the noun it is being used with and the context around it. Instead of fixing the polarities of all the adjectives from the beginning, only certain adjectives are taken with a prior polarity. After this, the targets within the domain are determined. The polarity of other adjectives, also referred to as target-specific adjectives, is computed based on the occurrence of their combination with adjectives and prior polarity in a large corpus [70]. For aspect-level sentiment analysis, after the aspect detection stage, a co-occurrence algorithm [71] is used to determine aspect categories [72]. The algorithm involves a co-occurrence matrix capturing the co-occurrence count between the words. For the sentiment classification part, a sentiment lexicon is created based on the idea that words occurring near aspects with known polarity possibly are of the same polarity.

A large corpus can be used to identify words that commonly occur in positive contexts and the words that commonly occur in negative contexts, based on which the polarity of those opinion words can be determined. In short, words that often appear together in text documents are likely to share similar polarity. The Pointwise Mutual Information (PMI) is used to indicate the probability that two words will co-occur and also to determine how a pair of words are lexically related to each other. The sentiment orientation of a text can be ascertained by measuring the similarity of the words present in the document with the words that have a known polarity [73]. Turney [74] leveraged PMI [75] for determining whether a review recommends a product or does not recommend it. It is based on the idea of extracting phrases containing two words, with one of them being an adjective or an adverb. The overall orientation of an extracted phrase is calculated as the difference of the PMI of the phrase containing

'excellent', a positive word, and the PMI of the phrase with 'poor', a negative word. Finally, the average orientation of all the phrases present in the text is determined to indicate whether the review recommends the product or not. When a significantly large corpus is used for such word similarity-based approaches, sentiment analysis performance across multiple domains becomes more consistent [73].

In a star rating problem, Hogenboom et al. [76] used statistical approaches for sentiment analysis with a five-class classification. They used a vectorized representation of features in the form of sentiment words constructed using a sentiment lexicon. Their basis was that the sentiment expressed by a document is more in terms of the count of distinct words that have a semantically same orientation, and therefore, they used a binary representation. The statistical approaches were then applied to the vectorized representations. For comparing the similarity of documents, they used Jaccard similarity and cosine similarity measures. In an alternate approach, the similarity is measured by modeling classes as probability distributions based on word distributions.

Latent Semantic Analysis (LSA) is a popular technique in information retrieval that is based on the measurement of the semantic closeness of the documents based on the presence of words with similar meanings [77]. The evaluation of the latent semantic structure involves the creation of a term-document matrix, and Singular Value Decomposition (SVD) [78]. In their work on the analysis of hotel reviews given by customers to understand customer satisfaction, Xu et al. [79] used Term Frequency Inverse Document Frequency (TF-IDF) for transforming the frequency matrix and SVD for dimension reduction. In another work, Phu and Tran [80] used the Dennis Coefficient (DNC) with LSA, and the combined LSA and DNC model is used to separate documents into positive or negative classes. Their model can process a high number of documents in a short time and is also applicable to other languages.

Probabilistic LSA has also been used to extract the latent semantic topic in text documents represented using the bag-of-words model [81]. This approach constitutes the computation of the documents' probability distributions using Probabilistic LSA and then utilizes the Z-vector obtained for the sentiment classification. The authors noted that the probabilistic LSA caused a significant improvement in the results of sentiment determination, and it was also suited in the clustering scenario where multiple topics are present in a document. They tested their model on textual data from sources such as Weibo and showed that the model predicted sentiment with greater accuracy compared to the word histogram approach.

B. Semantic Approaches. Semantic approaches to sentiment analysis are proposed on the premise that words that are semantically closer have alike sentiments [82]. If a word's synonyms lean toward a particular polarity, it is likely that the word will be of the same polarity. WordNet is a popular lexical database that explains the semantic relationships between words [56]. Kim and Hovy [83] used WordNet to obtain the synonyms of a word with unknown polarity and determined a word's sentiment class by ascertaining the extent to which the word's synonyms are associated with a particular class. Their work includes the identification of opinion holders pertaining

to specific topics and the classification of their sentiment on the basis of the words expressing polarity presented in the opinionated sentences.

The semantic similarity of words used by users in their posts on social networking websites with the words in a lexical database like WordNet can help us determine the polarity expressed by the users. In sentiment detection for suicide prevention based on Twitter data, WordNet was used to compare the semantic similarity of the words presented in the training data tweets with those in a test tweet [84]. In another approach for classifying tweets, Madani et al. [85] computed the semantic similarity of the tweet with the representative documents of the classes positive, neutral, and negative. The representative documents contain words that belong to a specific sentiment class, and the class with which the tweet has the most semantic similarity is its assigned polarity. Their methods gave positive results in terms of error rate and F1 measure.

Saif et al. [86] performed sentiment classification on the Twitter dataset by incorporating semantic concepts of the entities presented in the text. They evaluated the methods for obtaining entities from tweets and mapping the entities to concepts and noted that AlchemyAPI was more accurate in the mapping of entities to concepts. They proposed three ways to use the semantic information for the classifier, namely, (i) replacement of entities with their corresponding concepts, (ii) semantic augmentation, and (iii) semantic interpolation. They concluded that semantic interpolation outperformed the other two techniques and also showed that semantic techniques are suitable for large datasets spanning multiple topics.

Expert systems to perform automated analysis of weaknesses of a product are very useful as all the reviews given by multiple customers cannot be read manually. Zhang et al. [87] identified polarity associated with each product aspect or attribute for their Weakness Finder. They determined different feature words associated with the aspects of the product. A basic approach to this would be on the premise that explicit feature words belonging to the same category are usually synonyms or antonyms, which can be determined using a dictionary. However, this approach has vocabulary limitations and does not take into consideration domain knowledge. To handle this, they incorporated semantic approaches by taking into consideration the ideographic nature of Chinese characters, which gave them the opportunity to leverage word structures to find words expressing the same concept. They also used Hownet to calculate similarity scores between words [88] and thus expand the set of feature words.

13.2.3 Hybrid Approaches

Using a hybrid of lexicon and machine learning-based approaches enables us to leverage stability of lexicon approaches and the accuracy of machine learning approaches [89]. Mukwazvure and Supreethi [90] used a hybrid approach with the use of sentiment lexicon, followed by training with machine learning algorithms. On a dataset of comments derived from *The Guardian* (www.theguardian.com), the sentiment lex-

icon was used to decide the sentiment annotation, and machine learning classifiers were then trained using this data. The authors also proposed future work in the form of a domain-specific lexicon and the incorporation of opinion rules for getting more robust labels.

Kolchyna et al. [64] showed that the inclusion of lexicon score and some other handcrafted features enhanced the performance of SVM, Naive Bayes, and decision tree classifiers. Feature selection based on information gain showed that the lexicon sentiment score was the top input feature for the classifier. The pSenti, a hybrid of lexicon and learning-based method, is another concept-level system [91]. This approach outperformed systems such as SentiStrength when tested on datasets containing software and movie reviews [92]. The authors highlighted the advantage of their system over an only-lexicon-based approach in terms of sentiment classification accuracy.

13.2.4 Graph-Based Approaches

Approaches involving graph-based techniques have been applied for sentiment analysis. The graph-based techniques are used to obtain the sentiment from a graph-structured framework of opinions with the extracted features. The nodes in the graph are selected based on the features that occurred in the input, such as reviews and stock prices data.

Bordoloi and Biswas [93] created a graph with tokenized words for sentiment analysis of product reviews data, which helped in extracting semantic relationships. The nodes of the graph signify the tokens, and the edges represent the co-occurrence of the corresponding tokens. Main keywords are determined by assigning the degree centrality measure to the nodes in the graph. After polarity is assigned to the keywords, sentiment analysis of the datasets is done, and the model performs better than some of the existing methods. A similar approach was used by Castillo et al. [94] who constructed a co-occurrence graph and used centrality measures [95] for extracting the most important words for supervised learning-based sentiment analysis of documents.

The Potts model is a probability graph model that comprises a network of variables that can take multiple values and the values of the variable are not ordinal [96]. Graph-based approaches can be useful for the polarity classification of individual sentences, especially when the sentence does not have enough features to make a decision about its polarity. By incorporating features outside of the sentence, sentiment classification can be improved. Zhao et al. [97] used features of sentences within the same documents as well as features of sentences in other correlated documents in their graph propagation model. To determine the sentiment of an ambiguous sentence, they incorporated (i) the sentiment features of the neighboring sentences of that sentence and (ii) the sentiment of sentences in other documents that share semantic similarity with that sentence. They constructed a graph in which nodes are the sentences and connections between them represent an intra-document or inter-

document relationship as mentioned above. The probability that a node belongs to a particular sentiment class in the graph is determined using the Potts model [96]. They noted that their proposed model outperformed approaches to sentiment classification that use only features pertaining to that particular sentence. Takamura et al. [98] also used the Potts model in their work on the classification of adjective-noun pairs ('phrases') based on polarity. They constructed a network of words that are interconnected based on the presence of one of the words in another word's gloss. There are two types of connections, the first type indicating the words have similar orientation while the second one indicates dissimilar orientation. They observed that their proposed approach was also successful in classifying phrases containing words that were not known before.

Goldberg and Zhu [99] presented a graph-based technique that leverages a semi-supervised algorithm for inferring the rating of a document using the sentiment analysis, especially focusing on the use case of limited availability of data annotated with ratings. Their undirected graph is comprised of documents as nodes, where an unlabeled document node is connected with the nearest labeled and unlabeled document nodes based on a similarity measure. This approach was shown to be better than the approaches that do not leverage unlabeled data in situations where labeled documents are limited. Wang et al. [100] worked on hashtag-level sentiment classification instead of analyzing the sentiment polarity of each tweet pertaining to a given topic. They improved the hashtag-level sentiment classification by constructing a graph model that captures the co-occurrence relationship between the hashtags, and also used boosting approach to further enhance the results. Through the evaluation of the proposed method on a tweet corpus, they showed that their approach improved the results over the baseline.

13.2.5 Other Approaches

In this section, we discuss the approaches that did not fall under any of the above categories. Formal Concept Analysis (FCA) is a mathematical technique that is used for the representation of knowledge and involves visualizing of concepts and any dependencies between them [101, 102]. The concept is based on lattice and set theory. The input is a matrix with boolean values where a row is an object and the columns refer to attributes. The application of FCA creates a concept lattice with a hierarchy of relationships between concepts, as well as the dependencies between attributes. With FCA allowing the analysis of complicated structures and dependencies, Shein proposed a sentiment classification approach for text documents leveraging domain ontology used with FCA [103]. After the POS tagging and FCA-based ontology steps, SVM was used for sentiment classification.

Fuzzy Formal Concept Analysis (FFCA) is used to create a conceptual hierarchy using fuzzy formal concept [104]. FFCA uses fuzzy logic, and it differs from FCA, in which a binary value is used to indicate the relationship between an object and attribute. In FFCA, the degree of relationship is indicated with a value between 0 and

1 [105, 106]. Li and Tsai [106] in their work on polarity classification used FFCA to extract the internal insights from a corpus and also to understand the abstract concepts of the documents. Their proposed approach is versatile, i.e., it can be used in multiple domains and is also less sensitive to noise. Park et al. [104] used FFCA to get a hierarchy of sentiment words, which was not provided by WordNet [56], and thus provide an extension of hierarchy-based feature-level sentiment analysis.

Cambria [107] explains concept-based sentiment analysis as textual analysis using concepts and affective knowledge through ontology and other semantic networks. This allows an understanding of subtly expressed sentiments by analyzing the interlinking of a concept with another concept that reveals emotions. The Sentic computing approach to sentiment analysis includes the usage of artificial intelligence and semantic web for representing knowledge and allows for analyzing documents at the level of concepts by understanding their dependencies [108]. Cambria et al. [108] proposed SenticNet2, which serves as a resource for all-round analysis of textual language by furnishing both the sentics and the semantics related to more than 14,000 concepts. It can be utilized in real-life applications, such as social data mining, and deal with structured as well as unstructured data.

13.3 Datasets

Here, we discuss publicly available social networking datasets that are annotated for sentiment analysis, extracted from websites such as Twitter and Facebook.

1. CINLP [109]: Tromp and Pechenizkiy collected this Twitter dataset as (i) search tweets having happy smileys like :), :-), :D, etc., for 30 min, (ii) search tweets having sad smileys like :(, :-(, :'(, etc., for another 30 min, and (iii) for neutral messages, tweets are extracted from news pages, for example, BBC, CNN (English), and EenVandaag (Dutch) for 30 min. The dataset contains 11,778 training and 859 testing English tweets, and 4,805 training and 1,057 testing Dutch tweets. This dataset is available at [110].

2. COST [111]: This dataset is extracted from Twitter using positive emoticons, such as :D, :), and XD, and negative emoticons such as :(, :-/, and D:. After preprocessing and removing tweets having emoticons from both the polarities, the dataset contains 17,317 positive and 17,317 negative tweets. A tweet's polarity is set to 1 if it is collected using positive emoticons and 0 if retrieved using negative ones. The dataset can be obtained by emailing the authors.

3. COVID-19 Tweets Dataset [112]: This dataset contains tweets related to the Covid-19 pandemic and was extracted using 54 related keywords. Each tweet is assigned a sentiment score. The dataset is available at [113].

4. EMOT [114]: This English dataset was collected by Go et al. for positive emotion and negative emotion using happy ':)' and sad ':(' emoticons. The authors used 800,000 positive and 800,000 negative emoticon tweets for training. The test data

was manually annotated having 182 positive emoticon tweets and 177 negative emoticon tweets. It is available at [115].

5. EmIroGeFb [116]: This dataset is extracted from Facebook having comments in Spanish on three different topics (i) politics, (ii) football, and (iii) celebrities. The authors selected four Facebook pages on each topic and extracted 400 comments for each topic, 200 from each category, males and females. The dataset is labeled using three kinds of tags, (i) emotions used in the text, (ii) irony usage, and (iii) gender category. It also contains meta-information for each comment like Facebook ID, topic, annotator set, etc. This can be downloaded from http://ow.ly/uQWEs.

6. Facebook Comments v1.0 [117]: Zhang et al. collected 2000 Facebook comments and manually labeled them. The comments are labeled as follows: (i) if a comment has positive and objective sentences, then it is a positive message, (ii) if it has negative and objective sentences, then it is negative, (iii) if it contains only objective sentences, it is objective, (iv) if it consists of more positive sentences than the negative sentences, it is positive mixed, and (v) finally if it has both types of sentences but negative sentences are more, it is negative mixed. The dataset is available at http://cucis.ece.northwestern.edu/projects/Social/sentiment_data.html.

7. FB Sentiment [118]: Tran and Shcherbakov collected this dataset on the 2016 United States presidential elections from two popular news channels CNN and BBC on Facebook. It contains around 100,000 comments on 200 posts. The dataset is available in CSV format at https://raw.githubusercontent.com/saodem74/Sentiment-Analysis/master/Data/comment_data.csv. The CSV file of the dataset has datetime stamp value, topic, post title, comment text, and positive and negative sentiment score (float value).

8. Facebook Status Dataset [119]: The authors collected 62,202 Facebook statuses using *iFeel*, a Facebook Connect application. They collected 25 most recent status updates from a logged-on Facebook user and their friends. After preprocessing, data is labeled in two ways. First, they labeled data with positive and negative labels containing 4,320 statuses. Next, they did multi-class labeling having four labels: unhappy, happy, skeptical, and playful; this dataset has 3,612 labeled status. As we know, the dataset link is not provided by the authors.

9. HASH [120]: This tweet dataset is compiled from the Edinburgh Twitter corpus.[1] It contains 31,861 positive, 64,850 negative, and 125,859 neutral tweets in English.

10. Health Care Reform (HCR) [121]: This dataset was collected in March 2010 by retrieving the tweets having hashtag '#hcr'. It was manually annotated with five labels (neutral, negative, positive, irrelevant, and unsure(other)) and contained 1,381 negative, 541 positive, and 470 neutral tweets.

11. InterTASS corpus [122]: This dataset was released in the 2017 TASS workshop edition. It contains 3,413 tweets (1,008 training, 506 development, and 1,899

---

[1] http://demeter.inf.ed.ac.uk.

testing) in Spanish. Three annotators labeled the tweet polarity as positive, negative, neutral, and none. The dataset is available on the TASS website.

12. ISIEVE [120]: This English tweet dataset was collected and hand-annotated by the iSieve Corporation (www.i-sieve.com). It contains approximately 4,000 tweets annotated with negative, positive, or neutral sentiment related to the topic of the tweet.

13. LIGA_Benelearn11 [123]: This dataset is extracted by collecting tweets in six languages, (i) Dutch, (ii) French, (iii) English, (iv) German, (v) Italian, and (vi) Spanish, and six accounts were selected for each language. This dataset is available at [110].

14. New Delhi political elections [124]: The authors collected Facebook data related to New Delhi political elections from October 4 to December 10, 2013. The dataset contained around 103,000 posts and over 30,000 users that wrote at least 5 posts on around 50 different topics. The dataset is not freely available; authors can be contacted for the details.

15. Obama-McCain Debate (OMD) [125]: It contains 3,238 tweets collected during the first U.S. presidential TV debate that happened in September 2008. This dataset was labeled using Amazon Mechanical Turk (AMT).

16. O'Connor's Corpus (OC) [126]: It consists of one billion tweets collected over from 2008 to 2009 using Twitter API. It was annotated for sentiments using hashtags and emoticons.

17. RepLab 2013 [127]: It contains 142,527 English and Spanish tweets (63,442 positive, 16,415 negative, and 30,493 neutral), out of which 28,983 are in Spanish. The tweets were labeled with the polarity to analyze their potential impact (positive or negative) on the reputation of a company. The dataset was released by RepLab and is available at http://nlp.uned.es/replab2013/, and can be used for testing using the EvAll service.[2]

18. SAB/MAS corpus [128]: This dataset contains 4546 tweets mentioning brands from different sectors, such as automotive, beverages, food, banking, retail, sports, and telecoms. The dataset is in Spanish. The tweets are labeled for four positive ((i) happiness, (ii) trust, (iii) love, and (iv) satisfaction) emotions, and four negative ((i) sadness, (ii) dissatisfaction, (iii) hate, and (iv) fear) emotions. It is available at http://sabcorpus.linkeddata.es/.

19. Sentiment Strength Twitter Dataset (SS-Tweet) [92]: Thelwall et al. collected 4,242 tweets and labeled them with negative and positive sentiment strengths. A negative strength is labeled from -1 (as not negative) to -5 (as highly negative); similarly, positive strength is labeled from 1 (as not positive) to 5 (as highly positive). It is available at [129]. [130] labeled this dataset again with three labels as positive, negative, and neutral.

20. Sanders Twitter Dataset [131]: This dataset contains 5,512 tweets on four topics (i) Apple, (ii) Twitter, (iii) Google, and (iv) Microsoft. It was manually annotated and has 570 positive, 2,503 neutral, 654 negative, and 1,786 irrelevant tweets. It is available at http://www.sananalytics.com/lab.

---

[2] http://www.evall.uned.es/.

21. Tweets v1.0 [117]: This contains 1,000 tweets labeled in the same way as the Facebook Comments v1.0 dataset. The dataset is available at http://cucis.ece.northwestern.edu/projects/Social/sentiment_data.html.
22. Twitter US Airline [132]: This sentiment analysis dataset contains 14,640 tweets related to major US airlines. Tweets are labeled as positive, negative, or neutral.
23. TWITA [133]: The authors collected around 100 million Italian tweets for the time period Feb 2012–Feb 2013. The authors applied a language detection process to make sure that the collected tweets are in Italian. Two subsets having 1,000 tweets are annotated by three independent native-speakers. Each tweet is assigned a label as positive, neutral, or negative.
24. Twitter1 [134]: Agarwal et al. used 11,875 manually annotated tweets in English. The data was acquired from NextGen Invent (NGI) Corporation; please contact the authors for obtaining the dataset.
25. Twitter2 [135]: This is a collection of 667 English tweets. It is available at http://goo.gl/UQvdx.
26. The Dialogue Earth Twitter Corpus [136]: This tweet corpus has three subsets, (i) WA contains 4,490 tweets about weather, (ii) WB contains 8,850 tweets related to weather, and (iii) GASP having 12,770 tweets related to gas prices. The tweets were manually labeled with 5 labels, (i) positive, (ii) negative, (iii) neutral, (iv) not related, and (v) other (cannot tell), with respect to the topic. Refer www.dialogueearth.org for more details.
27. TASS general corpus [137]: This dataset was released in 2012 for the first edition of the TASS challenge. It contains 68,017 tweets in Spanish, collected from Twitter accounts of 154 celebrities, and covers topics such as music, sports, films, politics, entertainment, economy, soccer, technology, literature, and so on. Each tweet is labeled with either of five polarity levels as (i) P+, (ii) P, (iii) NEU, (iv) N, and (v) N+, or with 'NONE' sentiment tag. The dataset is freely available at [138] after signing a non-commercial user agreement.
28. TASS politics corpus [139]: This dataset is also from TASS workshop 2013. It contains 2,500 tweets in Spanish related to four major political parties (IU, PP, UPyD, and PSOE). The tweets are tagged at entity and tweet level with three polarity levels as neutral, positive, negative, and None. The dataset is available at [140].
29. TASS social-TV corpus [141]: This dataset is from TASS 2014 edition and was used in subsequent editions. It contained 2,773 tweets in Spanish related to a football match in 2014.
30. TASS STOMPOL corpus [142]: This dataset was collected during the Spanish general election, 2015. It contains 1,284 Spanish tweets on politics-related topics such as economics, political party, education, and health system from the main political parties (IU, PP, PSOE, Cs, UPyD, and Podemos). Two people annotated the tweets for each user's opinion about each aspect as negative, positive, and neutral; the disagreement was resolved by a third annotator. The dataset is available at [143].

13.4 Future Directions

– Real-time Analysis: In applications of SA such as marketing, fake news detection [144, 145], and opinion mining, the data should be processed in real time to produce and update the results constantly. Real-time analysis is crucial, especially in use cases such as election results prediction, where people's opinions might change quickly. Social media continuously generates a large amount of data, and the information can go viral in just a few seconds. If a person posts a negative review, companies need to act fast to minimize its impact. One can therefore propose fast real-time processing tools for SA [22]. The tools should also consider the data coming from different sources; for example, an election campaign might be running on different websites (Facebook, Twitter) at the same time.

– Building Datasets: Most of the works in sentiment analysis are for English language datasets. There is still a lack of lexicon datasets and benchmark datasets for other languages. Social network data can be collected using API; however, data labeling is a concern. Some of the datasets have been annotated by authors, or they have used crowdsourced portals for the labeling. In crowdsourcing, each text is annotated by 2–3 people, and the final label is decided based on the majority [125]. Though recently researchers have focused on building resources for other languages, the creation of more datasets will be useful to expand sentiment analysis to other languages.

– Preprocessing: Real-world data is noisy and unstructured. It needs preprocessing steps for spelling correction, grammar correction, abbreviation handling, and removal of unnecessary characters and strings, such as URLs from tweets [119]. One can therefore work on designing fast preprocessing steps. There are very few works proposing optimization techniques for feature selection. One can also propose methods for automatic feature selection using machine learning and optimization techniques.

– Domain-Independent Models: The accuracy of most of the proposed models has been verified using domain-specific datasets as the training and testing for a domain-specific corpus gives a higher edge to the accuracy. Cross-domain SA was introduced to reduce the manual effort in training machine learning models, but it is still a major issue. For example, "The screen is curved" is a good review for television but a bad review for a smartphone. Sentiment vocabularies and the strength of sentiment words differ across domains. In some cases, domain-specific models are preferred, such as the review analysis for a certain kind of product, but in some cases, general domain-independent models will be preferred if the incoming text is from different kinds of sources. Based on the demand, we need to build both kinds of models having better accuracy [73].

– Multimodal Reviews: People also post reviews or opinions in the multimedia format on social networking or blogging websites. An example of this is a spoken review of a movie posted on Facebook. Sentiment analysis techniques need to consider features other than text, such as facial expressions, voice base, and emotion,

for multimodal data [146]. The recent trends and availability of multimodal data make it a potential research area.

- Reduce Computational Cost: One major challenge is the reduction of the computational cost of models, especially for real-time analysis [22]. Machine learning methods having lower training costs are in demand. Additionally, the time-taking annotation process causes less availability of datasets. Researchers, therefore, can design semi-supervised or unsupervised methods to learn significant features [46].
- Generalized Models: Researchers have proposed neural network-based models, and their results show higher accuracy for the given dataset. However, the accuracy of different methods in different cases might not be the same, and it will take researchers time and effort to assess the efficiency for the given application dataset. Therefore, more generalized methods might be preferred [147].
- Temporal data analysis: On social networking websites, people constantly keep posting content that represents their opinion, mood, and emotion for a topic. This data can be helpful in the sentiment analysis of future posts of a user. For example, if a person doesn't like a political party, then there is a very high probability that if the user posts about that party it might seem positive but will be a sarcasm. Most of the methods mainly consider lexicons or linguistic features for SA; however, considering the past data can improve the accuracy. Temporal data analysis for opinion mining is still in a fancy state and can be looked at for further research [148].
- Homophily and SA: In real-world networks, the opinion of a user is influenced by its neighbors. It is observed that similar people respond to a situation in a similar manner. One potential direction is to consider the impact of homophily, and spatial-temporal patterns in SA [149].
- Sarcasm and Irony: These are frequently used on social media and it is hard to detect sarcasm in microblog posts. In the last decade, a few researchers have explored this direction [150]. Future research works can consider exploring this further as (i) sarcastic words differ in different languages, and (ii) there is a lack of universal datasets for this type of analysis.
- Recommendation Models: In social networking websites, a user prefers to see the posts confirming their belief and opinion. Hsu et al. [151] have shown that posts containing strong sentiments receive more clicks than the posts containing neutral sentiments, and posts containing negative polarity attract more views than posts containing positive polarity. The feed recommendation model based on SA and the social network of the user will be appreciated [152].
- Bias: Challenges associated with gender, racial, regional, and age-related biases are bound to arise in sentiment analysis systems. Researchers have detected significant biases originating from the language resources in [153]. The designing of fair systems will be highly appreciated and better to apply in real-life applications.

13.5   Conclusion

This chapter has mainly focused on different types of techniques used for sentiment classification of content. These techniques include machine learning-based, lexicon-based, hybrid, and graph-based. Machine learning-based approaches leverage machine learning algorithms to learn from the available data to be able to classify the content as negative, positive, or neutral in nature. Lexicon-based techniques are further categorized as dictionary-based and corpus-based techniques and rely on a sentiment lexicon to determine the polarity of a given document. Hybrid techniques are a combination of both machine learning and lexicon-based techniques. Another novel class of techniques is the graph-based approaches, which leverage the graph-based structures among features or components of the content. Next, we discussed the datasets which are available for the research, extracted from social networking websites. Finally, we discussed open directions for this field of research, ranging from real-time analysis and interdisciplinary research.

References

1. Ya-Han, H., Chen, Y.-L., Chou, H.-L.: Opinion mining from online hotel reviews-a text summarization approach. Inf. Process. Manag. **53**(2), 436–449 (2017)
2. Zhang, L., Wang, S., Liu, B.: Deep learning for sentiment analysis: a survey. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. **8**(4), e1253 (2018)
3. Santos, C.D., Gatti, M.: Deep convolutional neural networks for sentiment analysis of short texts. In: Proceedings of COLING, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 69–78 (2014)
4. Santos, C.N.D.: Think positive: towards twitter sentiment analysis from scratch. In: SemEval@COLING (2014)
5. Vinodhini, G., Chandrasekaran, D.: Sentiment analysis and opinion mining: a survey. Int. J. Adv. Res. Comput. Sci. Technol. **2**, 06 (2012)
6. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), pp. 79–86. Association for Computational Linguistics (2002)
7. Luo, F., Li, C., Cao, Z.: Affective-feature-based sentiment analysis using SVM classifier. In: 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pp. 276–281. IEEE (2016)
8. Wu, C.E., Tsai, R.T.H.: Using relation selection to improve value propagation in a conceptnet-based sentiment dictionary. Knowl.-Based Syst. **69**, 100–107 (2014)
9. Nadia, F., Hruschka, E., Hruschka, E.: Tweet sentiment analysis with classifier ensembles. Decis. Support Syst. **66**, 170–179 (2014)
10. Smailović, J., Grčar, M., Lavrac, , N., Žnidaršič, M.: Stream-based active learning for sentiment analysis in the financial domain. Inf. Sci. **285**, 181–203 (2014)
11. Yan, G., He, W., Shen, J., Tang, C.: A bilingual approach for conducting Chinese and English social media sentiment analysis. Comput. Netw. **75** (2014)
12. Balahur, A., Perea-Ortega, J.: Sentiment analysis system adaptation for multilingual processing: the case of tweets. Inf. Process. Manag. **51**, 547–556 (2015)
13. GMedhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4), 1093–1113 (2014)

14. Shein, K.K.P, Nyunt, T.T.S.: Sentiment classification based on ontology and SVM classifier. In: 2010 Second International Conference on Communication Software and Networks, pp. 169–172. IEEE (2010)

15. Liu, Y., Bi, J.-W., Fan, Z.-P.: A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm. Inf. Sci. **394**, 38–52 (2017)

16. Balahur, A.: Sentiment analysis in social media texts. In: WASSA@NAACL-HLT (2013)

17. Platt, J.: Sequential minimal optimization: a fast algorithm for training support vector machines (1998)

18. Kiritchenko, S., Zhu, X., Mohammad, S.: Sentiment analysis of short informal text. JAIR **50** (2014)

19. Ahmad, M., Aftab, S., Bashir, M.S., Hameed, N., Ali, I., Nawaz, Z.: SVM optimization for sentiment analysis. Int. J. Adv. Comput. Sci. Appl **9**(4), 393–398 (2018)

20. Sharma, A., Dey, S.: A boosted SVM based sentiment analysis approach for online opinionated text. In: Proceedings of the 2013 Research in Adaptive and Convergent Systems, pp. 28–34 (2013)

21. Rish, I. et al.: An empirical study of the naive Bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3, pp. 41–46 (2001)

22. Goel, A., Gautam, J., Kumar, S.: Real time sentiment analysis of tweets using naive Bayes. In: 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), pp. 257–261. IEEE (2016)

23. Pak, A., Paroubek, P.: Twitter as a Corpus for Sentiment Analysis and Opinion Mining, vol. 10 (2010)

24. Troussas, C., Virvou, M., Espinosa, K.J., Llaguno, K., Caro, J.: Sentiment analysis of facebook statuses using naive Bayes classifier for language learning. In: IISA 2013, pp. 1–6. IEEE (2013)

25. Parikh, R., Movassate, M.: Sentiment analysis of user-generated twitter updates using various classification techniques (2009)

26. Kang, H., Yoo, S. J., Han, D.: Senti-lexicon and improved naive Bayes algorithms for sentiment analysis of restaurant reviews. Expert Syst. Appl. **39**, 6000–6010 (2012)

27. Gamallo, P., Garcia, M.: Citius: A naive-Bayes strategy for sentiment analysis on English tweets, pp. 171–175 (2014)

28. Tan, S., Cheng, X., Wang, Y., Xu, H.: Adapting naive Bayes to domain adaptation for sentiment analysis. In: European Conference on Information Retrieval, pp. 337–349. Springer (2009)

29. Moraes, R., Valiati, J., Neto, W.G.:. Document-level sentiment classification: an empirical comparison between SVM and ANN. Expert Syst. Appl. **40**, 621–633 (2013)

30. Duncan, B., Zhang, Y.: Neural networks for sentiment analysis on twitter. In: 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC), pp. 275–278. IEEE (2015)

31. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)

32. Rupal, B., Shivangi, A., Yashvardhan, S.: Neural network-based architecture for sentiment analysis in Indian languages. J. Intell. Syst. **28** (2018)

33. Socher, R., Pennington, J., Huang, E., Ng, A., Manning, C.: Semi-supervised recursive autoencoders for predicting sentiment distributions, pp. 151–161, 01 (2011)

34. Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences. In: 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference, p. 1, 04 (2014)

35. Dos Santos, C., Gatti de Bayser, M.: Deep convolutional neural networks for sentiment analysis of short texts, p. 08 (2014)

36. Poria, S., Cambria, E., Gelbukh, A.: Aspect extraction for opinion mining with a deep convolutional neural network. Knowl.-Based Syst. **108**, 06 (2016)

37. Wang, X., Liu, Y., Sun, C., Wang, B., Wang, X.: Predicting polarities of tweets by composing word embeddings with long short-term memory, pp.1343–1353, 01 (2015)

38. Huang, M., Qian, Q., Zhu, X.: Encoding syntactic knowledge in neural networks for sentiment classification. ACM Trans. Inf. Syst. **35**, 1–27, 06 (2017)
39. Akhtar, M., Kumar, A., Ghosal, D., Ekbal, A. and Bhattacharyya, P.: A multilayer perceptron based ensemble technique for fine-grained financial sentiment analysis, pp. 540–546, 01 (2017)
40. Wang, X., Jiang, W., Luo, Z.: Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 2428–2437. Osaka, Japan, December 2016. The COLING 2016 Organizing Committee
41. Guan, Z., Chen, L., Zhao, W., Zheng, Y., Tan, S., Cai, D.: Weakly-supervised deep learning for customer review sentiment classification. In: IJCAI (2016)
42. Srinivas, M., Patnaik, L.M.: Genetic algorithms: a survey. Computer **27**(6), 17–26 (1994)
43. Govindarajan, M.: Sentiment analysis of movie reviews using hybrid method of naive Bayes and genetic algorithm (2014)
44. Li, S., Wang, Z., Zhou, G., Lee, S.Y.M.: Semi-supervised learning for imbalanced sentiment classification. In: Twenty-Second International Joint Conference on Artificial Intelligence. Citeseer (2011)
45. Tin Kam Ho: The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. **20**(8), 832–844 (1998)
46. Sintsova, V., Musat, C., Pu, P.: Semi-supervised method for multi-category emotion recognition in tweets. In: 2014 IEEE International Conference on Data Mining Workshop, pp. 393–402 (2014)
47. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc.: Ser. B (Methodol.) **39**(1), 1–22 (1977)
48. Zhai, Z., Liu, B., Xu, H., Jia, P.: Grouping product features using semi-supervised learning with soft-constraints. In: Proceedings of the 23rd International Conference on Computational Linguistics, pp. 1272–1280 (2010)
49. Zhou, Z.-H.: A brief introduction to weakly supervised learning. Natl. Sci. Rev. **5**(1), 44–53 (2018)
50. He, Y., Zhou, D.: Self-training from labeled features for sentiment analysis. Inf. Process. Manag. **47**(4), 606–616 (2011)
51. Hu, X., Tang, J., Gao, H., Liu, H.: Unsupervised sentiment analysis with emotional signals. In: Proceedings of the 22nd International Conference on World Wide Web, pp. 607–618 (2013)
52. Hailong, Z., Wenyan, G., Bo, J.: Machine learning and lexicon based methods for sentiment classification: a survey. In: 2014 11th Web Information System and Application Conference, pp. 262–265. IEEE (2014)
53. Khan, T., Durrani, M., Ali, A., Inayat, I., Khalid, S., Khan, K.: Sentiment analysis and the complex natural language. Complex Adapt. Syst. Model. **4**, 02 (2016)
54. Akter, S., Aziz, M.: Sentiment analysis on facebook group using lexicon based approach, pp. 1–4, 09 (2016)
55. Hardeniya, T., Borikar, D.A.: Dictionary based approach to sentiment analysis-a review. Int. J. Adv. Eng. Manag. Sci. **2**(5), 239438 (2016)
56. Miller, G., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.: Wordnet: an on-line lexical database. Commun. ACM **38**, 07 (2008)
57. Mohammad, S., Dunne, C., Dorr, B.: Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus, p. 08 (2009)
58. Hu, M., Liu, B.: Mining and summarizing customer reviews, pp. 168–177, 08 (2004)
59. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Lrec, vol. 10, pp. 2200–2204 (2010)
60. Li, X., Xie, H., Chen, L., Wang, J., Deng, X.: News impact on stock price return via sentiment analysis. Knowl.-Based Syst. **69**, 14–23 (2014)
61. Balage Filho, P., Pardo, T.A.S., Aluísio, S.: An evaluation of the brazilian portuguese liwc dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*, 2013

62. Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J., Chen, C.: DASA: dissatisfaction-oriented advertising based on sentiment analysis. Expert. Syst. Appl. **37**, 6182–6191, 09 (2010)
63. Darwich, M., Mohd Noah, S.A., Omar, N., Osman, N.A.: Corpus-based techniques for sentiment lexicon generation: a review. J. Digit. Inf. Manag. **17**(5), 297 (2019)
64. Kolchyna, O., Souza, T.T.P., Treleaven, P., Aste, T.: Twitter sentiment analysis: Lexicon method, machine learning method and their combination (2015). arXiv:1507.00955
65. Moreno-Ortiz, A., Fernández-Cruz, J.: Identifying polarity in financial texts for sentiment analysis: a corpus-based approach. Procedia - Soc. Behav. Sci. **198**, 330–338 (2015)
66. Velikovich, L., Blair-Goldensohn, S., Hannan, K., McDonald, R.: The viability of web-derived polarity lexicons (2010)
67. Hatzivassiloglou, V., McKeown, K.: Predicting the semantic orientation of adjectives, 05 (2002)
68. Kanayama, H., Nasukawa, T.: Fully automatic lexicon expansion for domain-oriented sentiment analysis. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pp. 355–363 (2006)
69. Feng, S., Zhang, L., Li, B., Wang, D., Yu, G., Wong, K.-F.: Is twitter a better corpus for measuring sentiment similarity? In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 897–902 (2013)
70. Fahrni, A., Klenner, M.: Old wine or warm beer: target-specific sentiment analysis of adjectives, 04 (2008)
71. Schouten, K., Frasincar, F.: Finding implicit features in consumer reviews for sentiment analysis. In: International Conference on Web Engineering, pp. 130–144. Springer (2014)
72. Schouten, K., Frasincar, F., De Jong, F.: Commit-p1wp3: a co-occurrence based approach to aspect-level sentiment analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 203–207 (2014)
73. Read, J., Carroll, J.: Weakly supervised techniques for domain-independent sentiment classification. In: CIKM 2009 (2009)
74. Turney, P.: Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Computing Research Repository - CORR, pp. 417–424, 12 (2002)
75. Turney, P.D.: Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In: European Conference on Machine Learning, pp, 491–502. Springer (2001)
76. Hogenboom, A., Boon, F., Frasincar, F.: A statistical approach to star rating classification of sentiment. In: Management Intelligent Systems, pp. 251–260. Springer (2012)
77. Wang, L., Wan, Y.: Sentiment classification of documents based on latent semantic analysis. In: International Conference on Computer Education, Simulation and Modeling, pp. 356–361. Springer (2011)
78. Deerwester, S., Dumais, S., Landauer, T., Furnas, G., Harshman, R.: Indexing by latent semantic analysis. J. Am. Soc. Inf. Sci. **41**, 391–407 (1990)
79. Xu, X., Wang, X., Li, Y., Haghighi, M.: Business intelligence in online customer textual reviews: understanding consumer perceptions and influential factors. Int. J. Inf. Manag
80. Phu Vo Ngoc: Latent semantic analysis using a dennis coefficient for english sentiment classification in a parallel system. Int. J. Comput. Commun. Control. **13**(3), 408–428 (2018)
81. Ren, W., Han, K.: Sentiment detection of web users using probabilistic latent semantic analysis. J. Multimed. **9**(10), 1194–1200 (2014)
82. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4), 1093–1113 (2014)
83. Kim, S.-M., Hovy, E.: Determining the sentiment of opinions, 01 (2004)
84. Birjali, M., Beni-Hssane, A., Erritali, M.: Machine learning and semantic sentiment analysis based algorithms for suicide sentiment prediction in social networks. Procedia Comput. Sci. **113**, 65–72 (2017)
85. Madani, Y., Erritali, M., Bengourram, J.: Sentiment analysis using semantic similarity and Hadoop MapReduce. Knowl. Inf. Syst. **59**(2), 413–436 (2019)
86. Saif, H., He, Y., Alani, H.: Semantic sentiment analysis of twitter. In: International Semantic Web Conference, pp. 508–524. Springer (2012)

87. Zhang, W., Hua, X., Wan, W.: Weakness finder: find product weakness from Chinese reviews by using aspects based sentiment analysis. Expert. Syst. Appl. **39**(11), 10283–10291 (2012)
88. Dong, Z., Dong, Q., Hao, C.: Hownet and the computation of meaning (2006)
89. El Alaoui, I., Youssef, G., Messoussi, R., Chaabi, Y., Todoskoff, A., Kobi, A.: A novel adaptable approach for sentiment analysis on big social data. J. Big Data **5**, 12 (2018)
90. Mukwazvure, A., Supreethi, K.P.: A hybrid approach to sentiment analysis of news comments. In: 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), pp. 1–6. IEEE (2015)
91. Mudinas, A., Zhang, D., Levene, M.: Combining lexicon and learning based approaches for concept-level sentiment analysis. In: WISDOM '12 (2012)
92. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. J. Am. Soc. Inf. Sci. Technol. **63**(1), 163–173 (2012)
93. Bordoloi, M., Biswas, S.K.: E-commerce sentiment analysis using graph based approach. In: 2017 International Conference on Inventive Computing and Informatics (ICICI), pp. 570–575. IEEE (2017)
94. Castillo, E., Cervantes, O., Vilarino, D., Báez, D., Sánchez, A.: Udlap: sentiment analysis using a graph-based representation. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp. 556–560 (2015)
95. Saxena, A., Iyengar, S.: Centrality measures in complex networks: a survey (2020). arXiv:2011.07190
96. Wu, F.Y.: The potts model. Rev. Mod. Phys. **54**, 235–268 (1982). Jan
97. Zhao, Y., Qin, B., Liu, T.: Integrating intra- and inter-document evidences for improving sentence sentiment classification: integrating intra- and inter-document evidences for improving sentence sentiment classification. Acta Autom. Sin. **36**, 1417–1425 (2010)
98. Takamura,, H., Inui, T., Okumura, M.: Extracting semantic orientations of phrases from dictionary. In: HLT-NAACL (2007)
99. Goldberg, A., Zhu, X.: Seeing stars when there aren't many stars: graph-based semi-supervised learning for sentiment categorization, 01 (2006)
100. Wang, X., Wei, F., Liu, X., Zhou, M., Zhang, M.: Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach, pp. 1031–1040, 10 (2011)
101. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: ICFCA (2009)
102. Škopljanac-Mačina, F., Blašković, B.: Formal concept analysis-overview and applications. Procedia Eng. **69**, 1258–1267 (2014)
103. Shein, K.P.P.: Ontology based combined approach for sentiment classification. In: Proceedings of the 3rd International Conference on Communications and Information Technology, pp. 112–115. World Scientific and Engineering Academy and Society (WSEAS) (2009)
104. Park, S.M., Kim, Y.-G., Baik, D.-K.: Sentiment root cause analysis based on fuzzy formal concept analysis and fuzzy cognitive map. J. Comput. Inf. Sci. Eng. **16**(3) (2016)
105. Quan, T.T., Hui, S.C., Cao, T.H.: A fuzzy FCA-based approach for citation-based document retrieval. In: IEEE Conference on Cybernetics and Intelligent Systems, 2004., vol. 1, pp. 578–583. IEEE (2004)
106. Li, S.-T., Tsai, F.-C.: A fuzzy conceptualization model for text mining with application in opinion polarity classification. Knowl. Based Syst. **39**, 23–33 (2013)
107. Cambria, E.: An introduction to concept-level sentiment analysis. In: Mexican International Conference on Artificial Intelligence, pp. 478–483. Springer (2013)
108. Cambria, E., Havasi, C., Hussain, A.: Senticnet 2: a semantic and affective resource for opinion mining and sentiment analysis. In: FLAIRS Conference (2012)
109. Tromp, E., Pechenizkiy, M.: Rbem: a rule based approach to polarity detection. In: Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining, pp. 1–9 (2013)
110. Databases and hypermedia group, tu/e. https://www.win.tue.nl/~mpechen/projects/smm/#Datasets. Accessed 22 Dec 2020

111. Martínez-Cámara, E., Martín-Valdivia, M.T., Ureña-López, L.A., Mitkov, R.: Polarity classification for Spanish tweets using the cost corpus. J. Inf. Sci. **41**(3), 263–272 (2015)
112. Lamsal, R.: Coronavirus (covid-19) tweets dataset (2020)
113. Coronavirus (covid-19) tweets dataset. https://ieee-dataport.org/open-access/coronavirus-covid-19-tweets-dataset Accessed 22 Dec 2020
114. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224N project report, Stanford **1**(12), 2009 (2009)
115. Training and testing emoticon data. http://cs.stanford.edu/people/alecmgo/trainingandtestdata.zip. Accessed 22 Dec 2020
116. Rangel, F., Hernández, I., Rosso, P., Reyes, A.: Emotions and irony per gender in facebook. In: Proceedings of Workshop ES3LOD, LREC, pp. 1–6 (2014)
117. Zhang, K., Cheng, Y., Xie, Y., Honbo, D., Agrawal, A., Palsetia, D., Lee, K., Liao, W.K., Choudhary, A.: SES: sentiment elicitation system for social media data. In: 2011 IEEE 11th International Conference on Data Mining Workshops, pp. 129–136. IEEE (2011)
118. Tran, H., Shcherbakov, M.: Detection and prediction of users attitude based on real-time and batch sentiment analysis of facebook comments. In: International Conference on Computational Social Networks, pp. 273–284. Springer (2016)
119. Ahkter, J.K., Soria, S.: Sentiment analysis: facebook status messages. Unpublished master's thesis, Stanford, CA, 2010
120. Kouloumpis, E., Wilson, T., Moore, J.: Twitter sentiment analysis: the good the bad and the omg! In: Fifth International AAAI Conference on Weblogs and Social Media (2011)
121. Speriosu, M., Sudan, N., Upadhyay, S., Baldridge, J.: Twitter polarity classification with label propagation over lexical links and the follower graph. In: Proceedings of the First Workshop on Unsupervised Learning in NLP, pp. 53–63 (2011)
122. Martínez-Cámara, E., Díaz-Galiano, M.C., García-Cumbreras, M.A., García-Vega, M., and Villena-Román, J.: Overview of TASS 2017. In: Proceedings of TASS, pp. 13–21 (2017)
123. Tromp, E., Pechenizkiy, M.: Graph-based n-gram language identification on short texts. In: Proceedings of 20th Machine Learning conference of Belgium and The Netherlands, pp. 27–34 (2011)
124. Rastogi, S.S.K., Singhal, R., Kumar, R.: A sentiment analysis based approach to facebook user recommendation. Int. J. Comput. Appl. **90**(16), 21–25 (2014)
125. Diakopoulos, N.A., Shamma, D.A.: Characterizing debate performance via aggregated twitter sentiment. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1195–1198 (2010)
126. O'Connor, B., Balasubramanyan, R., Routledge, B.R., Smith, N.A.: From tweets to polls: linking text sentiment to public opinion time series. In: Fourth International AAAI Conference on Weblogs and Social Media (2010)
127. Amigó, E., De Albornoz, J.C., Chugur, I., Corujo, A., Gonzalo, J., Martín, T., Meij, E., De Rijke, M., Spina, D.: Overview of replab 2013: evaluating online reputation monitoring systems. In: International Conference of the Cross-language Evaluation Forum for European Languages, pp. 333–352. Springer (2013)
128. Navas-Loro, M., Rodrıguez-Doncel, V.: Oeg at TASS 2017: Spanish sentiment analysis of tweets at document level, pp. 43–49 (2017)
129. Index of /documentation. http://sentistrength.wlv.ac.uk/documentation/. Accessed 22 Dec 2020
130. Saif, H., Fernandez, M., He, Y., Alani, H.: Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the STS-gold (2013)
131. Bravo-Marquez, F., Mendoza, M., Poblete, B.: Combining strengths, emotions and polarities for boosting twitter sentiment analysis. In: Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining, pp. 1–9 (2013)
132. Rustam, F., Ashraf, I., Mehmood, A., Ullah, S., Choi, G.S.: Tweets classification on the base of sentiments for us airline companies. Entropy **21**(11), 1078 (2019)
133. Basile, V., Nissim, M.: Sentiment analysis on Italian tweets. In: Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pp. 100–107 (2013)

134. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.J.: Sentiment analysis of twitter data. In: Proceedings of the Workshop on Language in Social Media (LSM 2011), pp. 30–38 (2011)
135. Kontopoulos, E., Berberidis, C., Dergiades, T., Bassiliades, N.: Ontology-based sentiment analysis of twitter posts. Expert. Syst. Appl. **40**(10), 4065–4074 (2013)
136. Asiaee, A., Tepper, T.M., Banerjee, A., Sapiro, G.: If you are happy and you know it... tweet. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 1602–1606 (2012)
137. Román, J.V., Lana Serrano, S., Martínez Cámara, E., González Cristóbal, J.C.: TASS-workshop on sentiment analysis at sepln (2013)
138. TASS: Workshop on semantic analysis at sepln. http://www.sepln.org/workshops/tass/tass_data/download.php. Accessed 22 Dec 2020
139. Villena-Román, J., García-Morera, J., Lana-Serrano, S., González-Cristóbal, J.C.: TASS 2013-a second step in reputation analysis in Spanish. Procesamiento del Lenguaje Natural **52**, 37–44 (2014)
140. Tass 2013 @ sepln. http://www.sepln.org/workshops/tass/2013/corpus.php. Accessed 22 Dec 2020
141. Román, J.V., Cámara, E.M., Morera, J.G., Zafra, S.M.J.: Tass 2014-the challenge of aspect-based sentiment analysis. Procesamiento del Lenguaje Natural **54**, 61–68 (2015)
142. García Cumbreras, M.Á., Martínez Cámara, E., Villena Román, J., García Morera, J.: Tass 2015–the evolution of the Spanish opinion mining systems (2016)
143. Tass 2015. http://www.sepln.org/workshops/tass/2015/tass2015.php#corpus. Accessed 22 Dec 2020
144. Kwon, S., Cha, M., Jung, K.: Rumor detection over varying time windows. PloS one **12**(1), e0168344 (2017)
145. Saxena, A., Hsu, W., Lee, M.L., Leong Chieu, H., Ng, L., Teow, L.N.: Mitigating misinformation in online social network with top-k debunkers and evolving user opinions. In: Companion Proceedings of the Web Conference 2020, pp. 363–370 (2020)
146. Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S.-F., Pantic, M.: A survey of multi-modal sentiment analysis. Image Vis. Comput. **65**, 3–14 (2017)
147. Shelke, N., Deshpande, S., Thakare, V.: Domain independent approach for aspect oriented sentiment analysis for product reviews. In: Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications, pp. 651–659. Springer (2017)
148. Paul, D., Li, F., Teja, M.K., Yu, X., Frost, R.: Compass: spatio temporal sentiment analysis of us election what twitter says! In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1585–1594 (2017)
149. Beigi, G., Hu, X., Maciejewski, R., Liu, H.: An overview of sentiment analysis in social media and its applications in disaster relief. In: Sentiment Analysis and Ontology Engineering, pp. 313–340. Springer (2016)
150. Reyes, A., Rosso, P.: Making objective decisions from subjective data: detecting irony in customer reviews. Decis. Support. Syst. **53**(4), 754–760 (2012)
151. Hsu, P.-Y., Lei, H.-T., Huang, S.-H., Liao, T.H., Lo, Y.-C., Lo, C.-C.: Effects of sentiment on recommendations in social network. Electron. Mark. **29**(2), 253–262 (2019)
152. Aivazoglou, M., Roussos, A.O., Margaris, D., Vassilakis, C., Ioannidis, S., Polakis, J., Spiliotopoulos, D.: A fine-grained social network recommender system. Soc. Netw. Anal. Min. **10**(1), 8 (2020)
153. Kiritchenko, S., Mohammad, S.M.: Examining gender and race bias in two hundred sentiment analysis systems (2018). arXiv:1805.04508