Knowledge-Based Systems 226 (2021) 107134

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A comprehensive survey on sentiment analysis: Approaches, challenges and trends

Marouane Birjali *, Mohammed Kasri, Abderrahim Beni-Hssane

LAROSERI Laboratory, Computer Science Department, University of Chouaib Doukkali, Faculty of Sciences, El Jadida, Morocco

ARTICLE INFO

Article history:
Received 1 July 2020
Received in revised form 25 March 2021
Accepted 10 May 2021
Available online 14 May 2021

Keywords:
Sentiment analysis
Opinion mining
Machine learning
Lexicon-based
Sentiment classification
Deep learning

ABSTRACT

Sentiment analysis (SA), also called Opinion Mining (OM) is the task of extracting and analyzing people's opinions, sentiments, attitudes, perceptions, etc., toward different entities such as topics, products, and services. The fast evolution of Internet-based applications like websites, social networks, and blogs, leads people to generate enormous heaps of opinions and reviews about products, services, and day-to-day activities. Sentiment analysis poses as a powerful tool for businesses, governments, and researchers to extract and analyze public mood and views, gain business insight, and make better decisions. This paper presents a complete study of sentiment analysis approaches, challenges, and trends, to give researchers a global survey on sentiment analysis and its related fields. The paper presents the applications of sentiment analysis and describes the generic process of this task. Then, it reviews, compares, and investigates the used approaches to have an exhaustive view of their advantages and drawbacks. The challenges of sentiment analysis are discussed next to clarify future directions.

1. Introduction

Sentiment Analysis is a task of Natural Language Processing (NLP) that aims to extract sentiments and opinions from texts [1,2]. Besides, new sentiment analysis techniques start to incorporate the information from text and other modalities such as visual data [3,4]. This research topic is conjoined under the field of *Affective Computing* research alongside emotion recognition [3]. According to [5], *affective computing* and *sentiment analysis* are the keys to the development of Artificial Intelligence (AI). Moreover, they have a great potential when applied to various domains or systems. The task of sentiment analysis can be considered as a text classification problem [6–8] because the process includes several operations that end up with classifying whether a given text expresses a positive or negative sentiment. However, sentiment analysis may seem an easy process, but in fact, it requires taking into consideration many NLP subtasks like sarcasm and subjectivity detection [9,10]. Moreover, the text is not always organized as in the books or newspapers [11,12] and can contain many orthographic mistakes, idiomatic expressions, or abbreviations.

Nowadays, sentiment analysis has become well acknowledged, not only among researchers, but also companies, governments, and organizations [4,8,13]. The growing use of the Internet have made the web become the universal and the most important source of information. Millions of people express their opinions, and sentiments in forums, blogs, wikis, social networks, and other web resources [14–16]. Those opinions and sentiments are very relevant to our daily lives, and hence there is a need to analyze this user-generated data in order to automatically monitor the public opinion and assist decision-making [6,14]. For example, Twitter posts have been used to predict election results [17].

For this reason, the field of sentiment analysis gained more interest within the last one and a half decades among research communities. Since 2004, sentiment analysis has become the fastest growing and the most active research area, as there has been a massive increase in the number of papers focusing on sentiment analysis and opinion mining recently [18]. Fig. 1 shows the rising popularity of sentiment analysis according to Google Trends.

Numerous surveys and review articles on sentiment analysis have been presented. Liu and Zhang [19] presented earlier in 2012 a survey of opinion mining and sentiment analysis. In this survey, the authors defined the problem of opinion mining and discussed the issues that should be addressed in the future. Furthermore, they analyzed the problem of detecting opinion spam and fake reviews. The authors investigated also the utility of online reviews. Piryani et al. [20] presented a scientometric analysis of research work conducted between 2000 and 2015 on opinion mining and sentiment analysis. The authors exploited

* Corresponding author.
E-mail addresses: birjali.marouane@gmail.com (M. Birjali), mohammed.kasri1@gmail.com (M. Kasri), abenihssane@yahoo.fr (A. Beni-Hssane).

https://doi.org/10.1016/j.knosys.2021.107134
0950-7051/

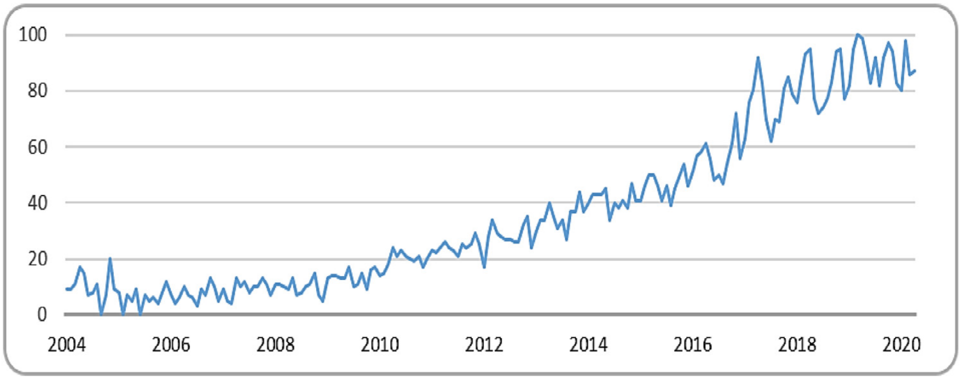

Fig. 1. Interest in "Sentiment Analysis" since 2004 according to Google Trends (trends.google.com/trends).

the research publication indexed in Web of Science database to identify year-wise publication pattern, rate of growth of publications, most productive countries, and more. This work was done computationally. Furthermore, a manual detailed analysis has been done also to identify the most popular approaches used in the publications, the levels of sentiment analysis have been addressed, and major application areas of sentiment analysis. Medhat et al. [21] proposed a comprehensive survey investigating and presenting sentiment analysis techniques and applications with brief details. The authors discussed also the related fields to sentiment analysis such as emotion detection and building resources. Ravi and Ravi [22] presented also a survey covering the published works during 2002–2015. In this survey, the authors explored the views presented by over one hundred papers. The focus was on necessary tasks, approaches, applications of sentiment analysis, and open issues of this field. In the same trend, Zhang et al. [23] presented a comprehensive survey on deep learning applied to sentiment analysis. Hemmatian and Sohrabi [24] provided a survey on classification techniques for opinion mining and sentiment analysis. In this survey, the authors compared and classified several techniques for aspect extraction and opinion mining to have a better understanding of their advantages and drawbacks. Chaturvedia et al. [2] proposed a survey that reviewed the hand-crafted and automatic models for subjectivity detection in the literature. The survey investigated these models by highlighting the key assumptions they have made in addition to the results they have obtained. The authors included also the comparison of advantages and limitations related to each subjectivity detection approach. Such papers are very useful to explore the open issues on this subtask of sentiment analysis. Rajalakshmi et al. [25], it briefly investigates various sentiment analysis methods in addition to some application domains and challenges. Moreover, many other studies [26–31] have been proposed in the literature.

To the best of our knowledge, the existing surveys often do not include the majority of sentiment analysis techniques and concentrate only on some supervised machine learning and lexicon-based techniques. Although this work has also discussed these approaches, it differs from the previous studies by covering the most used techniques. In addition to that, other surveys investigate sentiment analysis from specific points of view such as challenges or focus on specific domains such as movie reviews.

This paper presents a more comprehensive study of sentiment analysis as it discusses this field from different points of view, because it includes many research parts related to sentiment analysis including challenges applications tools, techniques, etc. This is very helpful for researchers and newcomers as they can find an enormous amount of information about this field in one paper. Our paper differs from other surveys also by providing detailed advantages and disadvantages of sentiment analysis techniques, which may help researchers to choose the appropriate approach to their problems. The significant contributions of this survey can be summarized as:

- A large number of literatures has been studied to describe the sentiment analysis process in detail and identify the well-known tools to perform this task.
- Categorizing the most used sentiment analysis approaches and summarizing them in brief details to have an overview of available techniques (machine learning, lexicon-based, hybrid and others techniques).
- Comparisons of available approaches to choose the appropriate one for a given application.
- Summarizing sentiment analysis applications and challenges to monitor the new trending researches.

The rest of this paper is organized as follows: Section 2 introduces the common application domains of sentiment analysis. Section 3 tackles the overall process of sentiment analysis in detail including the important steps such as data preprocessing and feature extraction. In Section 4, most sentiment analysis techniques are discussed and explained in brief detail. Section 5 presents sentiment analysis challenges and Section 6 discusses the conclusion and future trend.

2. Levels and applications of sentiment analysis

2.1. Levels of sentiment analysis

The task of sentiment analysis has been investigated at several levels. However, sentiments and opinions can be detected mainly at the document level, sentence level, or the aspect level [32–34]. Fig. 2 shows the levels of sentiment analysis. The first two levels are interesting and highly challenging. However, the third level is more difficult because it performs a fine-grained investigation [9]. A brief presentation of each level is as follows:

2.1.1. Aspect-level sentiment analysis

This level performs fined-grained analysis because it aims to find sentiments with respect to the specific aspects of entities. For example, consider the following sentence, "*The camera of iPhone 11 is awesome.*", the review is on "*camera*" which is an aspect of the entity "*iPhone 11*", and the review is positive. Therefore, the task at this level helps to detect exactly what people like or do not like [35]. It focuses on the aspects of entities (e.g., product features) instead of discovering the sentiment of paragraphs or sentences. According to [36], aspect extraction is the core task for sentiment analysis that can be implicit or explicit aspects. In this regard, the authors proposed a review of implicit aspect extraction techniques from different points of view.

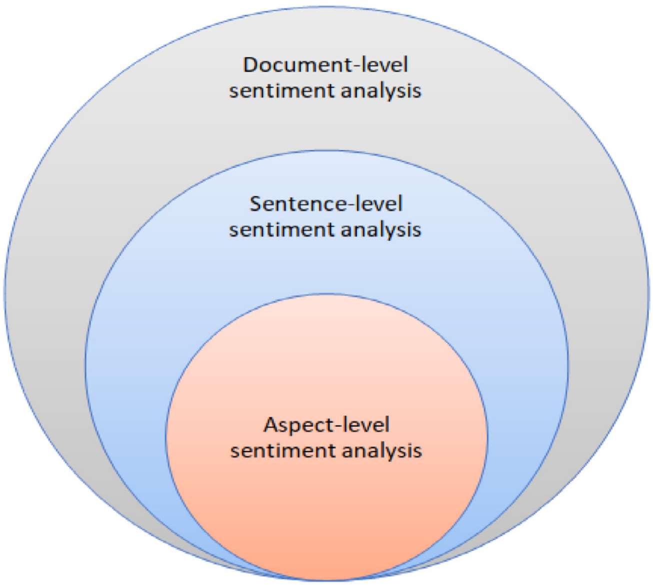

Fig. 2. Levels of sentiment analysis.

This level of detailed analysis is required by many real-life applications. For example, companies identify what components or aspects of the product are interesting to consumers in order to make product improvements. An approach for aspect-based sentiment analysis using Adaptive Aspect-Based lexicons was proposed by Mowlaei et al. [37]. The authors introduced two methods to generate two dynamic lexicons (Section 4.2 treats lexicon-based technique in more detail) for the improvement of aspect-based sentiment analysis; one using a statistical method, and the second using a genetic algorithm. Dynamic lexicon can be constantly updated without human supervision and they assign more accurate scores to context-related terms. They fused later the proposed lexicons with a set of well-known static lexicons in the literature to classify the aspects in reviews.

Instead of performing sentiment analysis based on one single level, two or more levels can be combined to reach a better performance. A joint approach of a sentence and aspect-level sentiment analysis of product comments on YouTube was proposed by Mai and Le [38]. The authors assumed that there is a strong mutual effect between sentence-level and aspect-level sentiment analysis because the sentiment polarity on the sentence-level depends and affects the sentiment polarity on the aspect-level and hence the joint approach can resolve the problem of the two levels together. After obtaining pre-processed comments a BERT-based model [39] was applied to extract the author's emotional reaction on both the sentence and aspect levels and then the analysis results are aggregated to generate statistical reports of the target product.

2.1.2. Sentence-level sentiment analysis

At this level, the focus is on the sentence. The main goal is to determine whether the sentence expresses positive, negative, or neutral opinion [40]. But to achieve this goal, the sentence needs to be classified as objective expressing factual information, or subjective expressing views and opinions. Several approaches tackled this level of analysis. Chen et al. [41] used a sentence type approach to improve the performance of sentence-level sentiment analysis. They applied first a neural network-based sequence model to classify sentences into three types based on the number of targets included in a sentence (sentence with non-target, one-target, or multi-target). For classification, they used a one-dimensional convolutional neural network (more information about this technique can be found in Section 4.1.5) where each type of sentence is fed to the model separately. Sentiment analysis at both the sentence and document level is important and useful, but it does not provide the necessary detail needed opinions on all aspects of the entity [21] as they do not find precisely what people like or dislike.

2.1.3. Document-level sentiment analysis

At this level, the process aims to classify whether a whole document expresses a negative or positive sentiment or opinion [42]. Each document is classified based on the overall sentiment of the opinion holder about a single entity (e.g., single product). The document-level classification works best when the document is written by one person and is not suitable for documents that evaluate or compare multiples entities. There have been many approaches proposed for document-level sentiment analysis. Zhao et al. [43] introduced a Domain-Independent Framework for Document-Level Sentiment Analysis (DFDS) with weighting rules based on Rhetorical Structure Theory (RST). The authors parsed the documents into rhetorical structure trees and then the sentiment scores of sentences are computed using two well-known lexicons. To identify the document sentiment polarity, they summed up the scores of sentences based on weighting rules. Sentiment analysis is very useful for many application domains, but sometimes the document may include some opposite sentiments which can impact the final decision.

2.2. Application domains

Sentiment analysis is very useful in a wide range of application domains starting from identifying customer opinion [44,45] to monitoring the mental health based on patient's social media posts [46]. In addition to this, the emergence of new technologies such as Big Data [47,48], Cloud Computing [49], and Blockchain [50] has widened the area of applications providing for sentiment analysis unlimited possibilities to be applied in almost every domain. As for instance, some of the common application domains of sentiment analysis are described in the following subsections.

2.2.1. Business intelligence

The usage of sentiment analysis in the domain of business intelligence has many advantages, for example, companies can exploit the results of sentiment analysis to make product improvements, study the customer's feedback, or adopt a new marketing strategy [51]. Analyzing customers' perceptions of products or services is the most common application of sentiment analysis in the area of business intelligence. However, these analyses are not applicable only to product manufacturers, but costumers can make use of them also to compare products and make a better decision. Bose et al. [45] tracked find food reviews on Amazon for ten years. They analyzed the reviews using NRC emotion lexicon that categorized customers' reviews into eight emotions (anger, fear, trust, anticipation, sadness, surprise, disgust, and joy) and two sentiments (positive and negative). Their results show that sentiment analysis can help to identify the customers' behaviors and overcome risks to meet the customers' satisfaction.

Sentiment analysis was applied also to market and Forex prediction. Rognone and al. [52] studied the impact of news sentiment on Bitcoin and traditional currency returns, volume, and volatility. A high-frequency intra-day data (15 min) for a sample period of seven years (2012–2018) was analyzed using Ravenpack News Analytics 4.0[1] to identify the sentiment of non-scheduled

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

news around Bitcoin and six traditional currencies. The authors found that traditional currencies react immediately and significantly to news wire messages coming from the economy. For bitcoin, the results were different from those on Forex which means that Bitcoin does not react similarly to news arrivals as traditional currencies.

Bitcoin and digital currencies (also called cryptocurrencies) refer to a new technology came to light recently called Blockchain, which is a decentralized digital ledger that can facilitate the transfer of peer-to-peer values (e.g. digital currency) rapidly and securely without the need of third-party such as banks and lawyers [50]. The transactions are verified concurrently by the blockchain network participants using peer-to-peer consensus protocols. The studies that apply sentiment analysis to the field of blockchain technology still scarce and the existing works generally use sentiment analysis to forecast digital currencies value as in the work of Kraaijeveld and De Smedt [53]. The authors used a cryptocurrency-specific lexicon-based approach to perform Tweeter sentiment analysis in order to predict the price returns of some well-known cryptocurrencies. Jing and Murugesan [54] proposed a theoretical framework to detect fake news automatically on social media using the principals and methods of blockchain technology. Although the effectiveness and the performance of this framework need to be validated, it promises that a combination of sentiment analysis and blockchain technology can be useful.

2.2.2. Recommendation system

A recommender system is an algorithm that aims to suggest relevant items (movies, music, or product to buy) to users [55]. An efficient recommender system can generate a huge amount of income for some industries. Thus, such systems [56–58] can benefit from the application of sentiment analysis to make a better recommendation. In the work of Li et al. [59], the authors proposed KBridge; an intelligent movie recommendation system using sentiment analysis of microblogs. The system identifies discussion groups in microblogs that are correlated with a given topic and use a proposed novel sentiment-aware association rule mining algorithm to investigate the correlation between groups. This investigation employs the sentiments expressed in microblogs to identify frequent program patterns and deduce the association rule of movie/TV program.

Shen et al. [60] introduced a new algorithm called Sentiment Based Matrix Factorization with reliability (SBMF+R) to leverage reviews for reliable recommendations. Their algorithm consists of three stages; First, they constructed a sentiment dictionary and used it to convert reviews into sentiment scores. In the second stage, they designed user reliability measures that combine user consistency and feedback on reviews. The third stage consists of incorporating the rating, reviews, and feedback into a probabilistic matrix factorization to improve the performance of recommender systems. In [61], the authors proposed an adaptive e-learning model based on social network analysis and they showed how Big Data and sentiment analysis can transform e-learning paradigm. The proposed sentiment analysis determines social indicators of learner that provide an appropriate learning rhythm.

2.2.3. Government intelligence

In addition to products and services, people also write comments on several subjects including, politics, religion, and social issues. Using sentiment analysis to identify opinions on government policies or other similar issues is very helpful for monitoring possible public reaction on implementation of certain policies as in the work of Georgiadou et al.[62] which used sentiment analysis of Twitter posts to investigate and aggregate public sentiment toward Brexit outcomes. Zavattaro et al. [63] analyzed the U.S local government tweets to determine if sentiment (tone) can positively influence citizen participation with government via social media. They used two algorithms to analyze tweets; the first algorithm was provided by a third party[2] and the second algorithm is a developed machine learning model by the authors. This work demonstrates that a positive tone tends to encourage citizen participation more than a neutral or negative tone. However, a positive tone alone is not enough to encourage this participation and undertake activities such as responding directly to citizens and sharing photos may have more encouraging influence.

Falck et al. [64] measured proximity between newspapers and political parties using the Sentiment Political Compass (SPC), a data-driven framework that classifies the attitude of newspapers toward political parties. The purpose of this paper is to study the impact of political tendencies of newspapers on voters' opinion forming. The authors crawled a dataset consisting of 180,000 newspaper articles from twenty-five newspapers during the German Federal Elections for 18 months and then used entity extraction and entity sentiment analysis to extract 740,000 political entities with their contextual sentiment. These data are exploited to analyze the relationship between newspapers and political parties.

Performing sentiment analysis to monitor the publics' mood should be in real-time for some situations [62,65]. However, real-time sentiment analysis requires the integration of other technologies such as Big Data [66]. Sentiment analysis and Big Data represent a perfect match and, it is not necessary to combine them just for real-time analysis [67]. Big Data refers to a collection of a large amount of complex data where state-of-the-art data processing technologies are unable to manage it efficiently [48]. EL Alaoui et al. [65] proposed an adaptable approach that analyzes users' social media posts to extract their opinions in real-time using Big Data tools. This approach consists of three stages; building sentiment words dictionaries for each entity, performing posts classification, and balancing the sentiment wights before executing a prediction algorithm. The authors gathered the 2016 US election-related tweets for evaluation. First, they evaluated the post classification performance, and then they compare their approach with other sentiment analysis tools.

2.2.4. Healthcare and medical domain

The application of sentiment analysis in the medical domain has gained so much interest recently. This application allows healthcare actors to obtain information about the diseases, adverse drug reactions, epidemics, and patients' mood [68], and analyze them to provide better healthcare services. However, it is difficult to apply sentiment analysis in such a domain because of some faced problems like terminology as found in the work of Jiménez-Zafra et al. [69]. Clark et al. [70] identified and analyzed tweets related to the patient experience as an additional informative tool for monitoring public health. They collected about 5.3 million breast cancer related tweets using Twitter's public streaming API for over a period of one year. After pre-processing, they analyzed tweets using a logistic regression classifier and a convolutional neural network model to sift tweets relevant to breast cancer patient experiences. The authors founded that positive experiences were shared regarding patient treatment, raising support, and spreading awareness. This work proves that social media can provide a positive outlet for patients to discuss their needs and concerns. Thus, analyzing patient's generated data on social media using sentiment analysis is very useful to deduce patient healthcare coverage and identify his treatment needs.

[2] www.repustate.com.

As mentioned before, other technologies can be combined with sentiment analysis to facilitate its application in a variety of domains, because they can resolve some faced problems such as data scarcity or the need of computation resources. Cloud Computing technology, for example, can be exploited for conducting sentiment analysis with computationally expensive approaches, and many of its advantages. Cloud Computing is a model that allows network access on-demand to computing services (hardware and software) using shared and dynamically scalable resources that can be easily and rapidly provisioned and released with minimal interaction of a service provider [49]. Ayata et al. [71] designed a system for emotion recognition that can be incorporated into a health monitoring system. This system can be helpful for elderly people to benefit from improved healthcare service quality. The framework is based on a wearable computer that monitors and collects physiological signals and sends them to the cloud via a mobile network. The collected data are processed and analyzed to predict physiological or psychological conditions of subjects using machine learning algorithms.

3. Sentiment analysis pre-processing

Sentiment analysis is not just one single problem, but in fact, it is a "suitcase" research problem that requires tackling many NLP tasks [72] as illustrated in Fig. 3. In order to extract sentiments from a given text, various steps are needed and many NLPs problems must be resolved. Thus, sentiment analysis as a field of Information Retrieval always struggles with NLPs unresolved problems [73], such as *negation handling* and *sarcasm detection* [9,72,74,75], in addition to other problems like opinion summarization [76] and implicit/explicit features extraction [36].

The generic sentiment analysis process is illustrated in Fig. 4. After data has been collected and extracted from several resources in diverse formats, it is converted to text and processed using NLP methods. In particular, the processing step involves text pre-processing, features extraction, and features selection. The classification step can be performed using different sentiment analysis approaches (e.g., machine learning), which are discussed in the next section. Therefore, the output can be presented in various forms. Several tools are available to handle all these steps and perform sentiment analysis. Serrano-Guerrero et al. [73] reviewed some free access web services for sentiment analysis to investigate their classification performance. In addition to web services, many other toolkits can be used as the ones listed in Table 1 except MADAMIRA.

3.1. Data extraction

3.1.1. Data collection and extraction

The first step in sentiment analysis is to have textual data, various sources and many tools are available, in order to obtain it. In general, text data can be created or collected as a part of research [77–79], by third-party [80] or via web scraping and crawling [81]. Therefore, enhancing textual data with other types of data (e.g., telecoms data, geospatial data, and video data) to perform sentiment analysis can lead to interesting results. Some available data sources are listed below:

- **Social Media (SM):** is a high dynamic data-source for academic research to investigate individual and collective life and behavior. It is defined as web-based and mobile-based Internet applications that allow the creation, access, and exchange of user-generated content [26].
- **Review Websites:** where people can post reviews or opinions about a particular entity (e.g., people, businesses, products, and services). In this case, data can be obtained from websites that are not particularly dedicated for reviews, but they contain millions of them such as e-commerce websites (www.amazon.com – product reviews) or from professional review sites like www.yellowpages.com.
- **Weblog (i.e. Blog):** refers to a simple webpage consisting of brief paragraphs of opinion, information, personal diary entries, or links, called posts, arranged chronologically with the most recent first, in the style of an online journal [82]. Blogs can be a rich source to perform sentiment analysis about various entities [83,84].
- **Forums:** or message boards allow users to discuss all types of questions [85], share thoughts, ideas, or help by posting text messages. This user-generated content tends to be emotional, which makes forums an interesting source for sentiment analysis. Furthermore, using forums as a source allows researchers to do sentiment analysis with respect to specific-domain [86,87].
- **Interview Transcripts:** are a written record of completed oral interviews. This process can be done in real-time or from an audio or video recording. Sentiment analysis on interview transcripts has been used in many studies [78,88].

There are different options to get data for sentiment analysis. One of these options is using public datasets which can be very low cost, but sometimes it is difficult to find relevant data suitable for the research purpose. However, many tools are available to create new datasets that generate relevant data [26,89,90], but this method can be costly and time-consuming sometimes. Data for sentiment analysis from web resources can be obtained using:

- **APIs** (Application Programming Interface): that provide access to textual content using HTTP-based protocols (e.g., Twitter API and Facebook API). In this case, data is collected using search API which is based on search queries to get a text containing specific keywords or using stream API for capturing real-time text data filtered by one or multiple filters such as keyword and geographic location [26].
- **Free available dataset:** are datasets created or collected by academics institutions, researchers, students, or companies, and can be downloaded freely such as Sentiment140[3] [89] and Large Movie Review Dataset[4] [90].
- **Web scraping:** is the automatic process of extracting data from websites that can hold an enormous amount of valuable data such as product details. However, this data can be used for sentiments analysis or other purposes. Many web scraping software or services are available for free like ParseHub.[5]
- **Crowdsourcing:** a technique used to outsource data creation or annotation. It can be very helpful to construct an efficient large volume of data in a short time. A well-known service for crowdsourcing is Amazon Mechanical Turk.[6]

3.1.2. Input data

The development of Web 2.0 made data available in different formats. Sentiment analysis research fields can make use of this variety to perform better sentiment classification. Thence, the input of an SA system is a corpus of documents or media files in different formats [3,91–93] such as :

- **TXT:** a plain text file that contains unformatted text;
- **CSV:** Comma-separated values file is a plain text file that uses a comma to separate values;

3 http://help.sentiment140.com/for-students.

4 https://ai.stanford.edu/~amaas/data/sentiment/.

5 https://www.parsehub.com.

6 https://www.mturk.com.

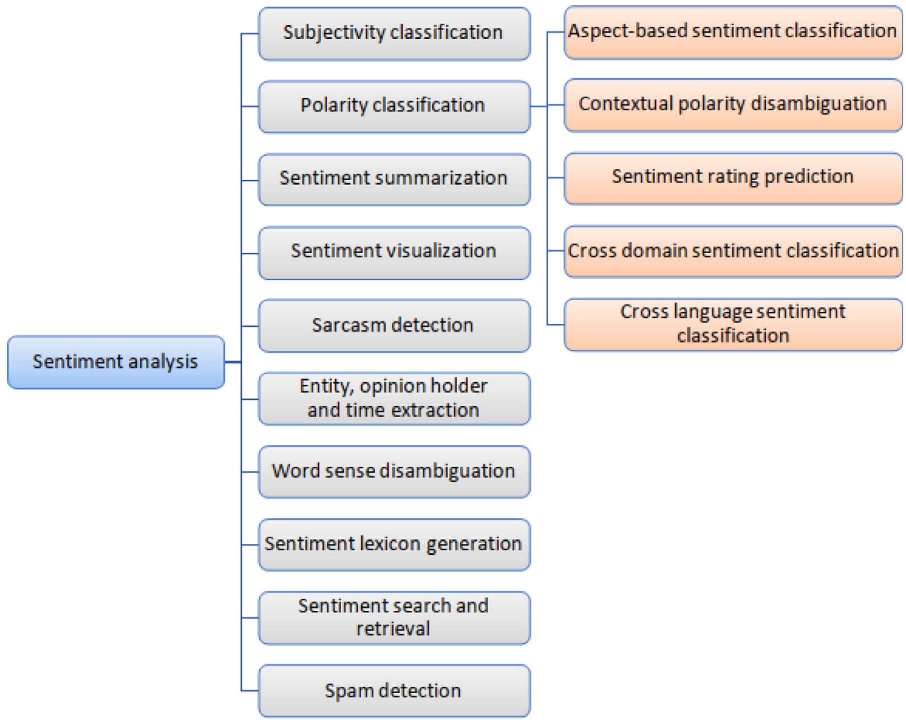

Fig. 3. Sentiment analysis common subtasks [72].

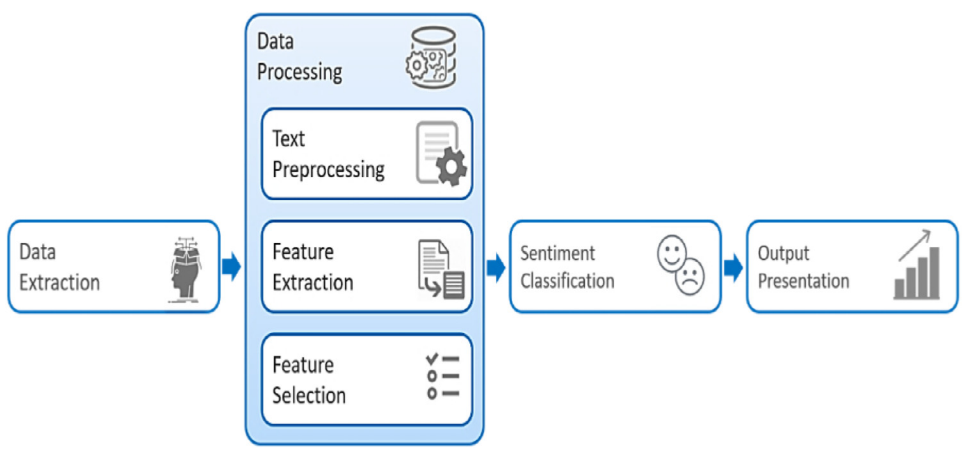

Fig. 4. The generic process of sentiment analysis.

- **XML:** Extensible Markup Language file is a hierarchical text format that describes the content using costume tags;
- **JASON:** a data format used for serializing and transmitting structured data over a network connection;
- **HTML:** a markup language used to describe the structure of a Web page;
- **Media file:** a file format that can contain an image, audio, or video.

3.2. Data pre-processing

The data acquired from different sources especially social media are usually unstructured. The raw form of these data may contain a lot of noise and all kinds of spelling and grammatical errors [40]. Therefore, it is necessary to clean and preprocess text before any analysis. The aim of the preprocessing step is not only to get better analysis but also to reduce the dimensionality of input data since many words are useless and should be removed, because they do not have any impact on the text polarity (e.g., articles, prepositions, punctuation, special characters). Some of the publicly available tools for different preprocessing and NLPs tasks are listed in Table 1. A performance comparison of several NLP toolkits in formal and social media text was conducted by Pinto et al. [94]. The whole process involves several common tasks:

- **Tokenization**: this step breaks text into smaller elements named tokens (e.g., document into sentences, sentence into words);
- **Stop words removal**: stop words are words (e.g., "the", "for", "under") that do not often contribute to analysis, and hence they are removed in advance.
- **Part-of-Speech (Post) tagging**: this step recognizes different structural elements of a text such as verbs, nouns, adjectives, and adverbs.
- **Lemmatization**: is the technique of converting a given word into a base form. This is similar to *stemming* technique,

except that lemmatization keeps word-related information such as PoS tags.

The pre-processing step can differ based on the input data format. However, some formats require additional processing and cleaning steps, for example expanding abbreviations and removing repeated characters like the "I" in "liiiiiiike". As mentioned before textual data can be very noisy, hence to perform better sentiment analysis two fundamental steps are needed, *features extraction* and *features selection*, which will be explained next.

3.3. Feature extraction

Feature extraction (FE) or feature engineering is a fundamental task in the sentiment analysis process because it can influence the performance of sentiment classification directly [99–101]. The purpose of this task is to extract valuable information (e.g., words that express sentiment) that describes important characteristics of the text. However, dealing with social media texts brings more challenges, and other features can be incorporated as in the work of Venugopalan and Gupta [102]. In this study, the authors take into consideration punctuations as features in addition to emoticons, hashtags, and capital text. This is not always the case, because generally punctuations are removed and text is converted to lower case. Some important features used in sentiment analysis are:

- **Terms presence and frequency**: is the simplest way of representing features and it is commonly used for Information Retrieval as well as for sentiment analysis. It considers single words or a list of *n* contiguous words which can be in the form of a unigram, bi-gram, or tri-gram, and their frequency counts as features. Term presence gives the words a binary value (zero if the word appears, or one if not). In term frequency, a term is given an integer value that represents its count in the document. TF–IDF weighting scheme can be applied to measure the importance of the term in the document.
- **Parts-of-Speech (PoS) tags**: are the labels or annotations that identify the word's function in a given language. In general, words can be categorized into several parts of speech categories (e.g., noun, verb, article, adjective, preposition, pronoun, adverb, conjunction, and interjection). For example the sentence "This camera is good." will be tagged as Stanford Log-linear Part-Of-Speech Tagger[7] [103]: *This* (determiner DT), *camera* (noun NN), *is* (verb VBZ), *good* (adjective JJ). Some sentiment analysis approaches rely on adjectives as they are important indicators of opinions [27, 104].
- **Opinion words and phrases**: opinion words are words that are commonly used to express positive or negative sentiments (e.g. *good* and *wonderful* for positive sentiment, *bad* and *terrible* for negative sentiment) [105]. In addition to many adjectives, there are nouns, verbs, some common phrases, and idioms that can also express opinions and sentiments without using opinion words.
- **Negations**: negation words (also called opinion shifters or valence shifters [105–107]) are words that may shift or change the opinion orientations and reverse the sentiment polarity. For example, *not*, *never*, *none*, *nobody*, *nowhere*, *neither*, and *cannot* are the most common negatives [106]. However, these words are often included in stop-word lists and removed from text analysis during the preprocessing step. Because of their impact, negation words should be handled with care, as not every appearance of negated words leads to negation.

Data used for sentiment analysis often exist in text format. Therefore, it is necessary to transform the input text into a fixed-length feature vector suitable for classification algorithms. This text representation is based on the bag-of-words model (BoW) and the vector space model (VSM). In general, text representation algorithms use a keyword set. Based on these predefined keywords, the feature extraction algorithm calculates the weights of the words in a text and then forms a digital vector which is the feature vector of the text [108]. Some typical text representation techniques are discussed in the following paragraphs.

3.3.1. Bag of words (BoW)

The BoW model is one of the simplest and the most common techniques to transfer text to numerical representation (vector) [100]. However, it has the disadvantage of losing syntactic information of the text, because it does not take on consideration ordering between words, sentence structure, or grammatical construction and only the occurrence of a word that matter [109]. For example, considering the following sentences:

S1: "The camera of this phone is awesome".

S2: "I want this phone; it is all about the camera. I love it".

First of all, BoW model creates a vocabulary (V in Table 2) of all unique words occurring in the document, then encode any sentence (S1 and S2 in this case) as a fixed-length vector with the length of the vocabulary of known words, where the value of each position in the vector represents a count or frequency of each word in the training set [100]. An extension of the BoW is Term Frequency–Inverse Document Frequency (TF–IDF), which is also simple and effective.

3.3.2. Distributed representation

Unlike the BoW model, in distributed representation (also called word embedding) of a qualitative concept (e.g., word, paragraph, document), the information about the concept is distributed all along the vector. Therefore, every position in the vector may be non-zero value for a given concept. Distributed representation is usually used with deep learning models which will be explained in Section 4.15. A brief introduction of some distributed representation methods is as follows:

- **Word2vec:** is one of the most common techniques for learning distributed representations of words developed by Mikolov et al. [110] using shallow neural networks. The word2vec architecture is a class of two models: Continuous Bag-of-Words (CBOW) and Skip-Gram (SG). The CBOW model predicts the current word from surrounding context words, whereas the Skip-Gram model predicts the surrounding context words from the current word.
- **Global Vectors (GloVe):** is an unsupervised learning algorithm developed by [111] for generating word embeddings by aggregating global word–word co-occurrence matrix from a corpus and the resultant representations show interesting linear substructures of the word vector space. The advantage of the GloVe model is that it can be trained quickly on more data as the implementation can be parallelized [112].

Recently, many approaches based on traditional word embedding methods or new ones have been proposed such as Doc2Vec [113], FastText [114]. However, the traditional word embedding methods learn word distributions that are independent of any specific task. For sentiment analysis, distributions can be reinforced with available resources such as sentiment lexicons [112,115].

[7] https://nlp.stanford.edu/software/tagger.shtml.

Table 1
Available toolkits for text preprocessing and NLPs tasks.

Toolkit	Language	Description
NLTK [95]	Python	Natural Language Toolkit is a suite of open-source Python modules that help to perform NLP tasks such as tokenization and PoS tagging. https://www.nltk.org/
CoreNLP [96]	Java	Stanford CoreNLP is a framework for basic and advanced NLP tasks as well as sentiment analysis. https://stanfordnlp.github.io/CoreNLP/
OpenNLP [97]	Java	Apache OpenNLP is an open-source Natural Language Processing Java library. It features an API for use cases like Named Entity Recognition, Sentence Detection, POS tagging, and Tokenization. https://opennlp.apache.org
MADAMIRA [98]	Java	MADAMIRA is a tool that performs Arabic NLP tasks like morphological analysis and tokenization. https://camel.abudhabi.nyu.edu/madamira/
TextBlob	Python	TextBlob is a Python library that provides a consistent API for diving into common natural language processing (NLP) tasks such as PoS tagging, sentiment analysis. https://textblob.readthedocs.io

Table 2
Example of BoW representation of a sentence.

V	The	camera	of	this	phone	is	awesome	I	want	it	all	about	Love
S1	1	1	1	1	1	1	1	0	0	0	0	0	0
S2	0	1	0	1	1	1	0	2	1	2	1	1	1

3.4. Feature selection

As mentioned before, a feature describes the characteristic of the data. A feature can be irrelevant, relevant, and redundant. In order to remove irrelevant and redundant features, various features selection (FS) methods are used. FS is a process of identifying and eliminating excessive and irrelevant features from the feature list to reduce the size of the feature dimension space and helps to improve the accuracy of sentiment classification [116].

Feature selection methods involve *lexicon-based methods* and *statistical methods* [117]. In lexicon-based approaches, features are generated by humans. The process is usually started by collecting terms that have a strong sentiment to build a small feature set. In the next step, this set is enriched with other terms through synonym detection or online resources. The advantage of these approaches is effectiveness because features are handled with care, however, handcrafted features selection is a long and difficult process. A popular example of this approach is SentiWordNet[8] [118] lexicon. On the other hand, statistical approaches are fully automatic and the most used for feature selection, but they often fail to separate features that carry sentiment from those that do not [119].

Statistical approaches are usually classified into four categories: *filter approach*, *wrapper approach*, *embedded approach*, and *hybrid approach* [120].

- **Filter approach**: This is the most common feature selection method [121]. It selects features based on the general characteristics of the training data without using any machine learning algorithm [122]. The feature is ranked based on some statistical measures and then the highest-ranking features are selected. Filter methods are computationally less expensive and suitable for datasets with a high number of features [120,123]. Some common filter approaches are *Information Gain* (IG) [99], *Chi-square* (CHI) [99], *Document Frequency* (DF) [124], and *Mutual information* (MI) [125].
- **Wrapper approach**: This approach depends on machine learning algorithms because it evaluates a subset of features based on the resulting performance of the applied machine learning algorithm. This dependency makes wrapper methods typically iterative and computationally intensive, but they can identify the best performing features set for that specific modeling algorithm [126,127]. A wrapper method is a combination of learning algorithms (e.g. Naïve Bayes [128] or SVM [129]) and a feature subset generation strategy (e.g. forward or backward selection [126]).
- **Embedded approach**: This approach incorporates the feature selection process during the modeling algorithm's execution. It uses classification algorithms that contain its own built-in ability to select features [130]. Therefore, it is computationally efficient compared to the wrapper approach. However, this approach is specific to the applied learning algorithm [120,126]. Common embedded approaches are based on various types of decision tree algorithms such as CART [131], C4.5, and ID3 [126,130] in addition to other algorithms like LASSO [132].
- **Hybrid approach**: This approach is a combination of filter and wrapper approaches, but in general hybrid methods combine different approaches to get the best possible feature subset. Hybrid approaches usually reach high performance and accuracy and take advantage of combined approaches. Many hybrid features selection approaches have been proposed for sentiment analysis [133–135].

4. Sentiment analysis techniques

Sentiment analysis is an active and flourishing research field and can be applied in many domains. For this reason, researchers propose, evaluate, and compare different approaches constantly. The aim is to increase the performance of sentiment analysis and to find solutions to this field challenges. Furthermore, applying sentiment analysis in new domains is a great incentive and makes this task more important. However, selecting the appropriate approach for sentiment analysis is very important and critical. The purpose of this section is to provide an overview of the most used approaches to perform sentiment analysis.

The existing approaches for sentiment analysis can be categorized based on various points of view (e.g. a view of the text,

[8] https://github.com/aesuli/SentiWordNet.

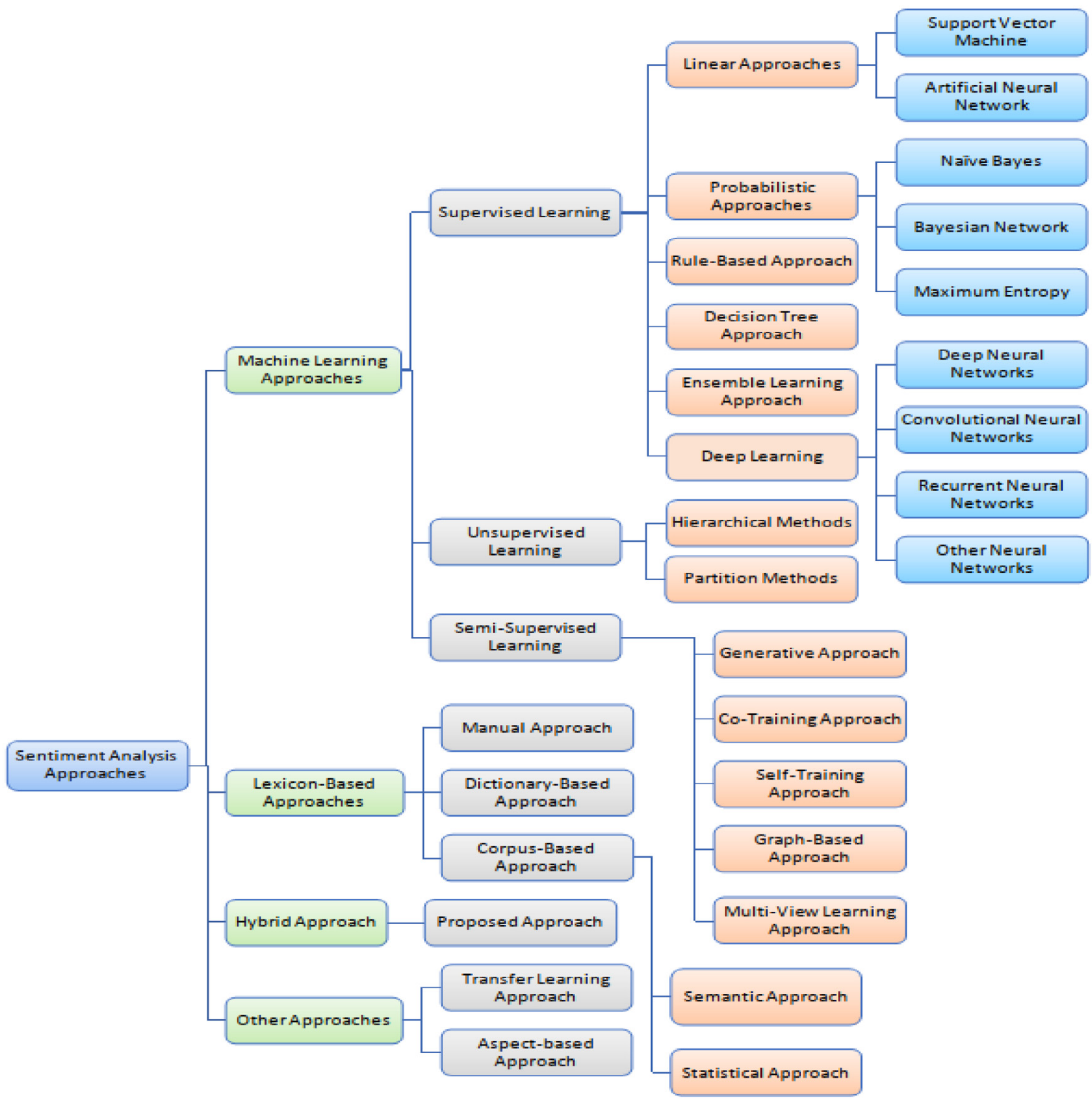

Fig. 5. Sentiment analysis approaches [21,73].

level of detail of text analysis) [136]. However, most literature usually divides sentiment analysis approaches into three categories: *Machine Learning approaches*, *Lexicon-Based approaches*, and *Hybrid approaches* [29,137,138]. Machine learning is the most widely-used approach. It relies on machine learning algorithms and linguistic features to perform sentiment classification. The lexicon-based approach uses sentiment lexicon which represents a list of words and phrases that are commonly used to express positive or negative sentiments [139]. On the other hand, hybrid approaches combine machine learning and lexicon-based approaches to improve sentiment analysis performance. Fig. 5 provides the outline of the sentiment analysis approaches.

4.1. Machine learning approaches

Machine learning approaches are used to classify sentiment polarity (e.g., negative, positive, and neutral) based on a train as well as test datasets. According to [140], these approaches can be divided into supervised learning [141], unsupervised learning [142], semi-supervised learning [143] and reinforcement learning [144]. The supervised approach is applied when the classification task has a specific set of classes, and when it is difficult to determine this set because of the absence of labeled data, the unsupervised approach can be the key in this case. On the other hand, the semi-supervised approach can be used for unlabeled datasets that include some labeled examples. The algorithms of reinforcement learning use trial and error mechanisms to help the agent interact with the surrounding environment to obtain maximum cumulative rewards.

Machine learning approaches can learn domain-specific patterns from the text which leads to better classification results, but the problem with these approaches is they often require large training datasets to achieve a good performance. However, a trained classifier on a specific dataset does not perform as well as for anther domain [145,146].

4.1.1. Supervised learning

Supervised approaches require labeled training documents, where the labels are generally the classes (e.g., positive, neutral, and negative). There are four types of supervised classification approaches which are *linear*, *probabilistic*, *rule-based*, and *decision tree* [138,140]. In the following subsections, a brief explanation and comparison of the most supervised classification approaches commonly used for sentiment analysis.

4.1.1.1. Linear approach. A Linear approach is a statistical approach that classifies sentiment using a linear or hyperplane decision boundaries [147]. Generally, the term hyperplane is used when there are more than two classes. This classification is performed by linear predictor $p = A.X + b$ which uses the given document features to predict which class it belongs. The vectors A and X are the vector of linear coefficients (weights) and the document frequency of the words, respectively. The predictions represent the dot product between A and X plus the bias b. Linear classifiers (also known as deterministic classifiers) are simple and often obtain state-of-the-art performances if the right features are used. Table 3 summarizes the advantages and disadvantages of SVM and ANN.

- **Support Vector Machine (SVM):** is a non-probabilistic classifier that can be used to separate data linearly or nonlinearly and can handle both discrete and continuous variables. It has a solid theoretical foundation and performs classification more accurately than most other algorithms in many applications [148,149]. According to [150] SVM is suitable for text classification, which makes it common in sentiment classification. The main goal of the SVM classifier is to find the optimal hyperplane to separate classes. An effective separation means that the hyperplane has the maximum margin to the closest training point from either class because a larger margin reduces the generalization error of the classifier. Many studies used SVM to perform sentiment analysis, Rana and Singh [151] analyzed movie reviews using linear SVM and Naïve Bayes. Their results show that the linear SVM method provided the best accuracy. Combining SVM with other algorithms achieved also positive results as in Al Amrani et al. [152] work. They proposed a hybrid approach based on SVM and Random Forest algorithm. Their work shows that the hybrid approach outperforms other algorithms individually.
- **Artificial Neural Network (ANN):** has emerged as an important method for classification and has gain more attraction recently [153]. It relies on the idea of extracting features from linear combinations of the data provided as an input, and models the output as a nonlinear function of these features [28,149]. The common architecture of a neural network involves three layers, namely input, output, and hidden layer where each layer consists of many organized neurons. The connection between two successive layers is established via links between the neurons of either layer. Each link has a corresponding weight value which is estimated by minimizing a global error function in a gradient descent training process [149]. In [154] Chen et al. proposed a neural network based approach that combines the advantages of machine learning and Information Retrieval techniques. They feed a back-propagation neural network with semantic orientation indexes. They found that the proposed approach increases the performance of sentiment classification and saves a considerable amount of training time. Several studies have explored ANN with more than one hidden layer which is discussed in a later separate subsection.

4.1.1.2. Probabilistic approach. Unlike the linear approach that output the most likely class of given input (either belongs to positive or negative class), a probabilistic classifier predicts a probability distribution over a set of classes, and they are usually based on Bayes' theorem [155]. The classifier uses mixture models to perform classification where each class is a part of the mixture. These kinds of classifiers are also called generative classifiers since each component of the mixture is a generative model. Probabilistic classifiers are easy to implement, computationally fast compared to other algorithms, and they do not require a lot of training data. However, the classification performance is sometimes worse if the data do not (at least nearly) meet the distribution assumptions [156]. Table 4 shows a summary of different advantages and disadvantages of Naïve Bayes, Bayesian Network and Maximum Entropy.

■ **Naïve Bayes (NB):** is a simple classifier and it is one of the most commonly used algorithms in the field of text classification. The model is based on Bayes Theorem and depends on BoW feature extraction. Therefore, the position of a word in the document is ignored and the presence of a particular word is independent to the presence of any other words. Naïve Bayes assigns a document d to the category c, that maximizes $P(c/d)$ by applying Bayes' rule:

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

where $p(c)$ is the prior probability of category c, $p(d|c)$ is the prior probability of document d being assigned to category c, and $p(d)$ is the prior probability of document d. Based on the assumption of independent feature conditions, Naïve Bayes calculates the posterior probability of a class, using word distribution in the document and the above equation could be rewritten as follows:

$$P(c|d) = \frac{p\,(c)\,p\,(w_1|c) * \ldots * (w_n|c)}{p(d)} \tag{2}$$

Various studies used Naïve Bayes as a classifier. In [157] Hasan et al. design a classifier using the Naïve Bayes algorithm to classify opinions expressed in English and Bangla language and they get a significant accuracy. They also checked some random reviews and tweets with their classifiers and achieved excellent outcomes in most of the cases.

- **Bayesian Network (BN):** it consists of a directed acyclic graph where each node represents a random variable and the edges between the nodes represent an influence relationship [158]. The model assumes that all nodes are independent as they are random while on the other hand assumes that these nodes are fully dependent because of the conditional dependencies between them. It is a complete architecture to describe relationships between a set of variables via the joint probability distribution and because of its extended structure, it is easy to add new variables. In the area of text classification, Bayesian networks used to find relationships among a large number of words. Bayesian network has been used directly by Wan and Gao [159] as sentiment classifier. They applied an ensemble sentiment classification system including Naive Bayes, SVM, Bayesian Network, C4.5, Decision Tree, and Random Forest algorithms. Their work shows that Bayesian Network outperformed all the six classifiers in the individual evaluation.
- **Maximum Entropy (ME):** also known as a conditional exponential classifier or Maxent classifier) it does not make assumptions about the relationships between features. It estimates the conditional distribution of the class label c given a document d to maximize the entropy of the system by using the following exponential form [160]:

$$P_{ME}\,(c|d) = \frac{1}{Z(d)} exp\left(\sum_i \lambda_{i,c} f_{i,c}(d,c)\right) \tag{3}$$

where $Z(d)$ is a normalization function, $f_{i,c}$ is a feature function for the feature f_i and the class c, and $\lambda_{i,c}$ is a parameter for the feature weight to make sure that the observed features match the expected features in the given set.

$$f_{i,c}\left(d, c'\right) = \begin{cases} 1, & n_i\,(d) > 0 \text{ and } c' = c \\ 0, & otherwise \end{cases} \tag{4}$$

Table 3
Advantages and disadvantages of SVM and ANN.

Classifier	Advantages	Disadvantages
SVM	•Effective and stable in high dimensional spaces.	•Poor performance if the number of features is much greater than the number of samples.
	•Obtains high accuracy and easy to train compared to other machine learning algorithms.	•The appropriate kernel function needs to be chosen.
	•Memory efficient due to its advantage of kernel mapping to high-dimensional feature spaces.	•Poor interpretability because there is no probabilistic explanation for the classification.
ANN	•The ability to deal with complex relations between variables and perform better generalization even against noisy data.	•Theoretically complex and difficult to implement.
	•Effective for high dimensionality problems.	•Requires high memory usage.
	•Fast execution time.	•Needs considerable training time compared to other algorithms and in some cases requires a large dataset.

Maxent used by Ficamos et al. [161] to perform sentiment analysis on Chinese social media. They evaluated their proposed features extraction method that is relies on POS tags using two classifiers Naive Bayes and Maxent. The Maxent classifier shows remarkable accuracy when used with unigrams and bigrams features.

4.1.1.3. Rule-based approach. The term rule-based classification can be used to refer to any classification scheme that makes use of IF-THEN rules for class prediction [162]. Therefore, the classifiers involved in this technique depend on a set of rules to perform sentiment classification. A rule can be expressed as LHS $\rightarrow$ RHS where the left-hand side (LHS) represents an antecedent of the rule or a set of conditions on the feature set expressed in DNF (Disjunctive Normal Form), and the right-hand side (RHS) represents a conclusion or consequence (class label) of the rule if the LHS is satisfied [138,162]. Rule-Based classifiers can classify new instances rapidly, and their performance is comparable to decision trees. Another advantage of the rule-based method is that can avoid over-fitting. However, their interpretation becomes difficult and extremely labor-intensive if there are too many rules. Moreover, it has poor performance against noisy data.

Tan et al. [163] used a rule-based approach with the aid of prior polarity lexicon to classify financial news articles. In order to determine the polarity of a sentence, they applied sentiment composition rules, and to calculate the overall sentiment of an article they used a mathematical formula called P/N ratio to average the sentiment values of all the sentences involved in the financial news article. Goa et al. [164] applied the rule-based approach to explore the emotions causes in a Chinese micro-blog. Their system triggers emotions based on an emotion model if a set of rules are met to extract the corresponding cause.

4.1.1.4. Decision tree approach. In this approach, training data space is decomposed hierarchically using a condition on the attribute value, in order to classify input data into a finite number of predefined classes. The condition on attribute values is the presence or absence of one or more words [21]. This based tree approach is a flowchart like structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent child node or class distributions [165]. Decision tree classifiers are easy to understand and interpret, moreover they can deal with noisy data. But, on the other hand, they are unstable and prone to over-fitting [166]. Decision tree approach performs very well on large datasets and hence it is not advisable for small datasets.

Ngoc et al. [167] implemented a new model based on the C4.5 algorithms of the decision tree which belongs to the data mining field, to perform document-level sentiment classification. The proposed model achieved 60.3% accuracy on the testing set. Several other decision tree algorithms such as CART, C5.0, C4.0 can be used to classify sentiment as in [168].

4.1.1.5. Ensemble learning approach. The main idea behind this approach is to combine several individual classifiers in order to obtain a classifier that outperforms every one of them [169]. This principle is used by humans when they want to make an important decision, they took on consideration several opinions. This technique takes advantage of the all used classifiers to make a better decision. In general, the final decision is obtained using a set of rules such as the Majority Vote method in Wan et Gao [159] work. An ensemble system tends to have a better generalization and good accuracy because of the classifiers' collaboration, but the main problem of such technique is it requires more computation and training time than a single algorithm. Therefore, it is advisable to choose algorithms carefully (e.g., fast algorithms such as decision trees). Ankit and Saleena [170] proposed an ensemble classification system using Naïve Bayes, Random Forest, Support Vector Machine, and Logistic Regression algorithms. Several studies [159,171,172] used ensemble-based to perform sentiment classification.

Another interesting sequential approach that falls inside the ensemble methods family is boosting technique. It consists of improving the prediction performance by training a sequence of weak classifiers. Each new classifier is trained only with samples that have been poorly classified by its predecessors [173]. The advantages of such a technique are: (1) the final classifier (a team of classifiers) learns to make accurate predictions on all types of data. (2) The bad performance of a single classifier does not matter since other members of the team will most likely solve the problem. Different boosting models have been proposed, including Adaptive Boosting (AdaBoost), Gradient Boosting Machine (GBM), and Boosted SVM. For sentiment analysis, this technique had been used by several studies [174–176]. Khalid et al. [174] proposed a voting classifier named GBSVM (Gradient Boosted SVM) which is constituted of gradient boosting and SVM. The evaluation of this model on different datasets proves that boosting techniques outperform state-of-the-art models.

Table 4
Advantages and disadvantages of NB, BN, and ME.

Classifier	Advantages	Disadvantages
NB	•Simple and easy to implement and interpret.	•Assumes that features are independent which cannot be the case most of the time.
	•Efficient in terms of computational resource requirements.	
	•Needs less training time and training data compared to other methods.	•Limited by data scarcity because for any possible value, a likelihood value should be estimated.
BN	•Requires less time to construct a model because it is easy to understand even in complex domains.	•Very computationally expensive; the reason why it is not widely used.
	•Handles missing data, and can achieve considerable accuracy with little training data.	•Not suitable for problems that contain many features.
	•Efficient against over-fitting.	
ME	•Suitable when the prior distributions are unknown.	•Tends to overfit.
	•Efficiency of acquiring information from textual data and handling a large amount of data.	

4.1.2. Unsupervised learning

Most of the existing approaches for sentiment analysis rely on supervised learning models trained from labeled corpora where each document has been labeled before training [177,178]. But sometimes, it is difficult to collect and create labeled datasets [179], especially for textual data which is unstructured most of the time. That is because their generation requires people labeling data which is too labor-intensive and time-consuming [180]. On the other hand, it is easier to collect unlabeled datasets and then, classify them using unsupervised learning approaches. These techniques make use of the documents' statistical properties such as word co-occurrence, NLP processes, and existing lexicons with emotional (or) polarized words [138]. However, in machine learning, unsupervised approaches in the field of sentiment analysis generally use clustering, which can classify data into different categories without specifying exactly which sentiment is represented by each category. In other words, the clustering approach divides data into groups (clusters), where the data of a cluster are very similar from a particular point of view than the data of different clusters. Ma et al. [181] explored the performance of some common clustering algorithms with respect to the task of sentiment analysis. According to [181,182] cluster analysis techniques can be categorized into *Hierarchical* and *Partition* clustering.

4.1.2.1. Hierarchical methods. Hierarchical methods create a hierarchical decomposition of a dataset which is represented by nested clusters (groups that have sub-groups) organized as a tree. Hierarchical techniques can be divided into two main strategies, namely *Agglomerative* and *Divisive* clustering [183]. The divisive clustering is called the top-down approach. This method starts from one individual cluster that groups all the data, and then assigns those data to sub-clusters through a recursive process based on the similarity between them. Tsagkalidou et al. [184] used this approach to propose a clustering framework that groups blog posts according to the closeness they present to certain emotions. The Agglomerative clustering (also called the bottom-up approach) considers that each data starts in its own cluster and then merges the clusters that contain similar data until one or a set of clusters remain. Archambault et al. [185] applied the agglomerative clustering technique to explore topics and sentiment in microblogging data.

4.1.2.2. Partition methods. Partition methods aim to partition data into a set of non-overlapping clusters where each element is assigned to only one cluster [186]. This partitioning is based on a similarity criterion which generally the Euclidean distance between elements. The data within a cluster have a very short distance to each other while having the largest distance to the data of other clusters. Table 5 shows commonly advantages and disadvantages of clustering approaches used in sentiment analysis.

The most popular partitioning algorithm is the k-means algorithm and its variants [182,187]. K-means algorithm starts with a pre-defined number of initial cluster centroids and assigns iteratively the data object in the dataset to cluster centroids based on the similarity between the data object and the cluster centroids. The process stops when a convergence criterion is met. The criterion can be a fixed iteration number, or the result does not change after a certain number of iterations. Sumbal et al. [188] applied K-means clustering to perform sentiment analysis on large scale data.

4.1.3. Semi-supervised learning

Semi-supervised learning (SSL) approaches are also used when there are difficulties in obtaining labeled data, but unlike unsupervised approaches, this technique uses a small set of initial labeled training data to guide the feature learning procedure. Thus, it fits in between supervised and unsupervised approaches. SSL approaches make full use of amounts of low-cost unlabeled data, save a lot of time and effort, and gain a classifier with strong generalization ability in addition to more labeled data [189]. Hussain and Cambria [143] proposed a novel semi-supervised learning model for Big Social Data analysis. It is based on the combined use of random projection scaling and SVM. The results demonstrate that such semi-supervised model can significantly improve the performance of some NLP tasks including sentiment analysis. Recent works on SSL-based sentiment analysis can be classified into five categories as generative, co-training, self-training, graph-based, and multi-view learning [177,190]. The advantages and disadvantages of each category are summarized in Table 6.

4.1.3.1. Generative approach. This approach assumes that data in different categories follow different distributions and the parameters of each distribution can be estimated if there is at least

Table 5
Advantages and disadvantages of clustering approaches used in sentiment analysis.

Approach	Advantages	Disadvantages
Hierarchical methods	•Easy to implement.	•Not suitable for large datasets because of the high-cost computation.
	•Perform well against noisy data.	•Very sensitive to outliers.
	•No need to define the number of clusters in advance.	•There is no possibility to move an object to another cluster once it was assigned to a specific cluster.
Partition methods	•Suitable for sentiment analysis.	•Sensitivity to noisy data.
	•Relatively scalable and simple.	•The problem of finding the initial cluster centroids and handling non-convex clusters.
	•Appropriate for large datasets because of the low computational requirements.	•Poor accuracy and stability.

one labeled data per category [177]. In other words, a generative model defines distributions on the inputs and use Bayes rule to predict the label (class) of a test input after training this model for each class. Mesnil et al. [191] proposed a very simple and powerful ensemble system for sentiment analysis that combines tree complementary and conceptually baseline models. One of them was based on a generative approach. The overall system achieves a new state of the art performance on IMDB movie review dataset4. That proves that ensemble learning can be used also with semi-supervised or unsupervised approaches.

4.1.3.2. Co-training approach. This algorithm was originally developed by Blum and Mitchell [192] and assumes that data can be represented using two independent views where each view has information about each data [193]. In co-training two separate classifiers will be trained to teach each other based on the shared information between them during the training process. Each classifier trained on a different feature set corresponding to the two views of the data. The training process is iterative, and at each iteration co-training updates the dataset by adding the most confident classified instances from each classifier to the labeled data. The process stops when all unlabeled data have been used or a specific number of iterations has been reached [31]. This algorithm has been used for sentiment analysis by several studies [192–194]. In the work of Xia et al. [190], Blum and Mitchell's algorithm was used to propose a dual-view co-training approach that addressed the negation problem and enhanced bootstrapping efficiency for semi-supervised sentiment classification.

4.1.3.3. Self-training approach. This approach is commonly used for semi-supervised learning. The procedure of self-training is divided into two steps. In the first step, the classifier is trained using a small amount of labeled data. In the second step, the trained classifier is used to classify unlabeled data in order to append the most confident samples to the original training set as new labeled data [195,196]. The last step will be repeated iteratively including the new labeled data. The resulting model is then evaluated using the test data. This approach has been used widely in the field of sentiment analysis [195,197,198]. He and Zhou [199] proposed a novel framework based on a self-training approach that learns from labeled features instead of a labeled instance. The experimental results show that their approach outperformed some existing methods.

4.1.3.4. Graph-based approach. In this approach, an architecture of a graph is used to represent the data. Vertices illustrate instances (e.g., sentences) in the graph while edges describe the similarity between instances. The strongly connected instances generally tend to belong to the same class [200]. Due to the wide use of this approach by many studies, its effectiveness was proved in many NLP tasks including sentiment analysis [200–202]. Jalilvand and Salim [203] used this approach to propose a sense level sentiment classification method using graph-based Word Sense Disambiguation (WSD) and a multiple meaning sentiment lexica. They compared their approach against a baseline method using two subjectivity lexicons to prove its effectiveness for sentiment classification.

4.1.3.5. Multi-view learning approach. This approach takes on consideration multiples points of view to treat the problem and the overall performance is obtained using the agreement between them [204]. Each classifier will be trained on a single view and then these classifiers used to label the unlabeled samples that will be added to the training set if they are classified with high reliability. This technique is generally applied to problems with multiple different feature sets. In [205] Lazarova and Koychev proposed an approach based on multi-view learning for movie review sentiment analysis in the Bulgarian language.

4.1.4. Reinforcement learning

Reinforcement learning (RL) is a machine learning method where an agent is rewarded in the next time step based on the evaluation of its previous action. The algorithms of RL use trial and error mechanisms to help the agent interact with the surrounding environment to obtain maximum cumulative rewards [206]. Reinforcement learning has been applied to solve different problems such as robot control, and it has been mostly used in games. However, applying this method to solve sentiment analysis problems is very scarce although its ability to handle complex tasks especially with the integration of Neural Networks. The main advantage of this method is the similarity to the learning process of human beings, which very desired in the field of sentiment analysis. Reinforcement learning uses the learned historical experiences to correct the errors committed during the training process and hence makes a better decision which makes it close to perfection. On the other hand, the conception of the reinforcement learning model can be laborious. Moreover, reinforcement learning requires a lot of data and it is computationally expensive.

Liu et al. [207] developed a reinforcement online learning method for real-time emotion state prediction using physiological signals. The authors exploited the concept of reward to modify the predictor on each iteration during the online training. They compared the effectiveness of their proposed method with least squares (LS) and support vector regression (SVR) algorithms. The

Table 6
Advantages and disadvantages of semi-supervised approaches for sentiment analysis.

Approach	Advantages	Disadvantages
Generative	•Achieves high accuracy if there is a small number of labeled instances.	•Inflexibility.
	•Effective, if the model is close to correct.	•Does not perform well for classification problems.
Co-training	•Good performance with limited number of labeled instances.	•Not suitable for datasets with one feature set.
	•Reduces mistake propagation. •Applies to most common classifiers.	•Sensitive to noisy data and outliers.
Self-training	•Simplicity	•Propagation of wrong results.
	•Suitable for dataset with significant amount of labeled data.	•Does not provide much information on convergence.
	•Applies to most common classifiers.	
Graph-based	•Good performance if the graph fits the task.	•Performance is sensitive to graph structure and edge weights.
	•Easy to interpret.	•Bad performance if the graph does not fit the task.
Multi-view learning	•Tackles the problem from different points of view.	•Assumes conditional independence between features.

experimental results show a significant time reduction and a prominent performance of the proposed method. Emotions can be used as a motivation to optimize the behavior of an agent as in the work of Broekens et al. [208]. The authors proposed a computational model of four emotions; joy, distress, hope, and fear, based on reinforcement learning primitives (e.g., reward). The model is instrumented as a mapping between RL primitives and emotional labels to study the relation between adaptive behavior and emotion. Agent-based simulation experiments show that emotions can be mapped to reinforcement learning providing a feedback signal for the agent to adapt its behavior which will optimize the communication between adaptive agents and humans.

4.1.5. Deep learning

Applying ANNs-based deep learning (DL) to sentiment analysis has become very popular recently. DL is an emergent area of machine learning that offers methods for learning feature representation in a *supervised* or *unsupervised* manner [209]. The term "deep learning" refers to neural networks with multiple layers of perceptron inspired by our brain [210]. Therefore, it is possible with this architecture to train more complex models on a much larger dataset, and hence, produce state-of-the-art results in many application domains, ranging from computer vision and speech recognition to NLP [23].

DL includes many neural network models such as CNN (Convolutional Neural Networks) [211], RNN (Recurrent Neural Networks) [212], and DBN (Deep Belief Networks) [213]. These models do not need to be provided with pre-defined features handpicked by an engineer, but they can learn sophisticated features from the dataset by themselves [214]. On the other hand, they are complicated and computationally very expensive. Several studies tackled deep learning approaches for sentiment analysis in detail [23,112,209,215,216]. However, the following subsections, give a brief description and summarization of the most common deep learning models used for sentiment analysis.

4.1.5.1. Deep neural networks (DNN). This model is an Artificial Neural Network (ANN) with multiple layers (hidden layers) between the input and output layers [217]. The input layer includes input data, the hidden layers include processing nodes called neurons, and the output layer includes one or several neurons used to yield the network outputs [215]. It uses sophisticated mathematical modeling and the learning power of ANN to find the right relationship whether to be linear or non-linear to map an input into an output. The flow process of ANNs and certainly DNNs can be categorized to feedforward and backward. Feedforward ANNs are straightforward networks and hence they are appropriate for sentiment classification. DNN architecture and its variants (e.g., CNN and RNN) have been used in many NLP tasks including sentiment analysis. Vassilev [218] designed a model named *BowTie* based on deep feedforward neural network, this model consists of one encoding layer, a cascade of hidden layers and an output layer. The evaluation of this model shows promising results compared to other methods.

4.1.5.2. Convolutional neural networks (CNN). This architecture is a special type of feedforward neural network originally employed in the area of computer vision [219], but recently it has achieved successful results in different areas such as recommender systems and NLPs. The layers of a CNN consist of an input layer, an output layer, and a hidden layer that involves multiple convolutional layers, pooling layers, normalization layers, and fully connected layers. Convolutional layers filter the inputs (e.g., word embedding in text sentiment classification) to extract features, while pooling layers reduce the resolution of features to make feature detection independent of noise and small changes. The normalization layer normalizes the output of a previous layer to improve the convergence during the training, and the fully connected layers used to perform the classification task. CNNs become very well-known recently in the field of sentiment analysis. One of the most popular CNN models for sentiment analysis was proposed by Kim [211]. The author evaluated a CNN model built on top of pre-trained word2vec to perform sentence-level sentiment classification. The model outperformed other methods and proves that pre-trained word-embedding can be good features for NLP tasks with deep learning.

4.1.5.3. Recurrent neural networks (RNN). This model uses a memory cell to process a sequence of inputs. The ability to capture and remember information about a long sequence makes RNNs widely used in NLP tasks such as sentiment analysis [220]. In

RNNs, the output is dependent on all the previous computation. For example, to predict the next word in a sentence, the model uses all the previous words states and the relation between them [221]. One of the main problems of standard RNN is vanishing gradient and to overcome this problem Hochreiter and Schmidhuber [222] introduced a special type of RNN called Long-Short Term Memory (LSTM) which becomes very popular in many fields. This architecture has been increasingly used by many researchers for sentiment classification. Li et al. [212] proposed a bidirectional LSTM model that can exploit the relationship between target words and sentiment polarity words in a sentence without relying on any sentiment lexicon. The experimental results show that this model outperformed other advanced methods.

4.1.5.4. Other neural networks. Several other types of deep neural networks used for sentiment analysis but not widely as the three models cited above. Among them, there is a special type of RNN model called Recursive Neural Network (RecNN) used by Li et al. [223] to introduce a novel Recursive Neural Deep Model. This model achieved high accuracy compared to Naïve Bayes, Maximum Entropy and SVM in binary sentiment classification of Chinese social data. Other unsupervised deep neural networks such as Autoencoders and its variants used also in the field of sentiment analysis but it is hard to use them directly for this task [224]. The general case is to use the encoder layer as a features extractor for classifiers as in the work of Zhou et al. [225]. Combining two or more deep learning models is a hybrid approach that is also commonly used as in the work of Rehman et al. [226]. The authors proposed a hybrid model using LSTM and a very deep CNN model named as Hybrid CNN-LSTM Model for sentiment classification. The following table illustrates the underlying advantages and disadvantages of DNN, CNN and RNN (see Table 7).

4.2. Lexicon-based approach

Lexicon-Based (also called knowledge-based) approach is one of the two main approaches used for sentiment analysis and requires a lexical resource named *opinion lexicon* (a predefined list of words) which associates word to their semantic orientation as negative or positive words using scores [139,227]. A score can be for example a simple polarity value such as $+1$, -1 or 0 for positive, negative, or neutral words respectively, or a value reflecting the sentiment strength or intensity. The final orientation of a document is obtained by calculating the semantic orientation values of the words that compose it. A document is tokenized into single words or micro phrases, and then sentiment values from the lexicon are assigned to each element. To conclude the overall sentiment of a given document, formula, or algorithm (e.g., sum and average) can be applied.

The lexicon-based approach is very practical at the sentence and feature level sentiment analysis. It does not require any training data and hence, it can be considered as an unsupervised approach. On the other hand, the main problem of this approach is domain dependency, because words can have multiple meanings and senses, thus, a positive word in a specific domain may not be in another. For example, given a word "small" and two sentences "The TV screen is too small"., and "This camera is very small"., the word "small" in the first sentence is negative, because generally, people prefer wide screens, while in the second sentence it is positive as if the camera is small, then it will be easy to carry. This problem can be avoided by the creation of a domain-specific sentiment lexicon or using a lexicon adaptation approach. Sanagar and Gupta [228] proposed a genre-level sentiment lexicon adaptation method. In contrary to other adaptation approaches that use labeled data, this new approach uses unlabeled data to learn the source and the target domain sentiment lexicons. The transfer learning approaches can be applied to learn new domain-specific lexicons as in the work of Sanagar and Gupta [229]. The authors proposed an unsupervised sentiment lexicon learning methodology that can be used for new domains of the same genre. After learning the polarity seed words from corpora of multiple source domains, the genre-level knowledge learned is then transferred to the target domains. Another problem of the lexicon-based approach is the dropped performance compared to machine learning approach if a large training dataset is provided. The three major techniques for creating and annotating sentiment lexicons [230] are listed below.

4.2.1. Manual approach

The manual approach requires human intervention to annotate the lexicon. The creation of sentiment lexicons contains two phases, namely, *generating of the sentiment bearing words list* and *the assignment of sentiment labels to these words.* This process is typically very laborious, costly, and time-consuming, but it can provide a consistent and reliable lexicon. To speed up this process, an automated approach can be implicated. In this case, a manual approach is used as a benchmarking procedure or to minimize the errors. Many lexicons have been created manually. Wilson et al. [231] created MPQA Subjectivity Lexicon, and Taboda et al. [232] created Semantic Orientation CALculator (SO-CAL) which are based on a manual lists of negators and intensifiers.

Researchers can also use *crowdsourcing* and *gamification*. Crowdsourcing is the practice to engage a group for a common goal on the Internet platforms. Mohammad and Turney [233] used Amazon Mechanical Turk to create word emotion and word polarity association lexicon. Gamification instead, is the application of game mechanics to non-game problems. Hong et al. [234] designed a game called *Tower of Babel* to engage players to assign a sentiment polarity to words for building a sentiment lexicon.

4.2.2. Dictionary-based approach

The assumption behind this approach is that synonymous words have the same sentiment polarities, while antonyms words have the opposite polarities. The sentiment lexicons in this approach, are created using the well-known dictionaries such as WordNet[9] [235] or thesauri [236]. First, a list of initial seed words with pre-known orientation is collected manually. The next step expands the list of words by searching for their synonyms and antonyms over other lexical resources. The newly found words are added iteratively to the previous list until no new words are found [140]. A manual examination can be done later to remove and correct errors. Baccianella and Esuli [118] created a well-known lexicon named SentiWordNet 3.0 using the automatic annotation of all synsets of WordNet 3. Park and Kim [237] proposed a method to build a thesaurus lexicon based on three online dictionaries. Some available lexicons to expand an initial seed are listed in Table 8. Sanagar and Gupta [238] presented a survey on polarity lexicon learning. The authors discussed the polarity lexicon in two aspects. They introduced in the first aspect, the polarity lexicon creation techniques starting with initial ones. In the second aspect, the authors gave valuable information about the available open-source polarity lexicon. At the end of the paper, open research problems and future directions of polarity lexicon creation were revealed.

The main problem of dictionary-based as all Lexicon-based approaches is the inability to find sentiment words with domain-specific orientation, and hence it is not suitable for context and domain-specific classification. Moreover, compiling dependency

9 https://wordnet.princeton.edu.

Table 7
Advantages and disadvantages of deep learning approaches for sentiment analysis.

Approach	Advantages	Disadvantages
DNN	•Easy to implement compared to other DL models	•Problem of over-fitting
	•Less training time required	•Considered as complex "black box"
CNN	•High accuracy	•Complex to design and maintain
	•Fast training	•Pooling layer may lead to losing the position or the order of a feature.
RNN	•The ability to capture sequential data which is very important for sentiment text classification.	•Require a longer time to train than other models.
	•High reliability	•Complex and computationally expensive.

rules is difficult and laborious, but on the other hand, this technique is not computationally expensive as long as there is no training of dataset, and represent a good strategy to easily and quickly build a lexicon with a large number of sentiment words and their orientation.

4.2.3. Corpus-based approach

Unlike dictionary-based, corpus-based approaches start with a list of seed sentiment words with pre-known orientation and exploit syntactic or co-occurrence patterns to find new sentiment words with their orientation in a large corpus. The identification of additional sentiment words uses linguistic constraints or conventions on connectives (e.g., AND, OR, BUT). For example, a pair of adjectives conjoined by a conjunction (e.g., "simple AND easy") usually have the same orientation. In addition to this idea which is called *sentiment consistency* although it is not always consistent in practice, a set of rules can be designed for these connectives. At the end of this process, several techniques such as clustering can be applied to construct sets of sentiment words (e.g. positive and negative words) [242,243].

This method was proposed for the first time in Hatzivassiloglou and McKeown [244]. In that work, authors constructed a set of frequently occurring adjectives with their orientation, and in order to expand the initial set they considered words that co-occurred alongside in the pattern *W1* and *W2* have the same orientation. They created a graph that contains words in vertices and their pairs in the edges and then used log-linear model to mark if two conjoined adjectives have the same or opposite orientation and cluster them in two sets of positive and negative words. The main advantage of the corpus-based approach is simplicity, however, it requires a large dataset to detect the polarity of words and hence the sentiment of the given text [245]. The corpus-based approach is usually divided into statistical and semantic approaches [246] as described in the following subsections.

4.2.3.1. Statistical approach. This approach obtains the sentiment orientation of a word depending on the statistics concept. The principle of this approach is that similar sentiment words usually have the same sentiment if they appeared together frequently in the same context. Therefore, the unknown polarity of a word is obtained based on the frequency of its co-occurrence with other words that appeared with them in the same context. The frequency of co-occurrence is calculated using Turney's method for computing mutual information [247]. Several studies have used this approach to generate sentiment lexicons and perform sentiment analysis. Han et al. [248] proposed a new domain-specific lexicon generation method for review sentiment analysis. They used mutual information to assign terms with their PoS tags in the lexicon. The authors got a good outcome using the proposed method.

4.2.3.2. Semantic approach. Unlike the previous approach, this technique (also called ontology-based approach) uses different rules to measure the similarity between words and assigns the same sentiment value directly to the semantically close words [249]. Generally, this method looks up in sentiment dictionaries for synonyms, antonyms, and words with a similar concept to extend a lexicon and to perform sentiment analysis as in the work of Zhang et al. [250]. The authors combined statistical and semantic approach to propose Weakness Finder, an expert system that find product weakness from Chinese reviews. They used the Chinese Hownet [251] lexicon to calculate the similarity of the words. The proposed expert system demonstrated a good performance on the experimental results.

4.3. Hybrid approach

The hybrid approach combines both lexicon and machine learning approaches. It combines the throughput of lexical analysis with the flexibility of machine learning approaches to cope with ambiguity and integrate the context of sentiment words [252]. The main reason behind the hybrid approach is to inherit high accuracy from machine learning and stability from lexicon-based approach.

The hybrid approach combines techniques from the two previous approaches in order to overcome their limitations and take advantage of their benefits. For that, the lexicon approach scores are utilized as input features to the sentiment classifier. Thus, sentiment lexica play an important role in the hybrid approach which is usually known to achieve a higher performance.

Only few models utilize the hybrid approach for sentiment analysis. Most of them used lexicon-based approaches to label word polarity to be further used in the sentiment analysis classifier. An early work of Devi et al. [253] used machine learning classifiers combined with dictionaries and HARN's algorithm which is proposed in lexicon-based approach to classified documents. First, they classified the reviews of each domain using two machine learning classifiers, namely Naïve Bayes and SVM and then they identified the polarity at document-level using HARN's algorithm. The hybrid approach achieved about 80%–85% more accurately than HARN's algorithm along. Deep Learning also can be combined with lexicons for the task of sentiment analysis. Shin et al. [254] integrated lexicon embeddings and an attention mechanism into Convolutional Neural Network. They constructed lexicon embeddings by taking word scores from multiple sources of lexicons. These embeddings are integrated into a CNN model using three methods; Naïve concatenation, Multichannel and Separate Convolution. They proved that lexicon integration can improve the accuracy, stability, and efficiency of the CNN model. A hybrid approach is proposed in [255] which combined both machine learning approaches and lexicon based

Table 8
List of common lexicons.

Lexicon Name	Description	Lexicon size	Output
WordNet [235]	A lexical database of English words which groups lexical units (e.g., verbs, nouns) according to their semantic and lexical relations.	117,000 synsets	A Synset that groups synonymous words expressing the same concept.
SentiWordNet [118,239]	A lexical resource for opinion mining based on WordNet dictionary. It contains synonym sets called synsets that groups sentence parts (e.g. nouns, verbs) with the same meaning and their polarity scores.	117,000 words	A score between 0.0 and 1.0 for positive, negative and objective polarity.
SenticNet[a] [240]	A semantic resource for SA based on conceptual primitives. It concludes the polarity of common-sense concepts at semantic level using dimensionality reduction.	200,000 concepts	Negative, Positive
MPQA [231]	A list of subjectivity clues and contextual polarity which represent a part of Opinion Finder developed and maintained by The University of Pittsburg [241].	20,611 words	Negative, Objective, Positive

[a] https://www.sentic.net.

in order to identify sentiments polarities of tweets. The authors presented their hybrid approach, HILATSA, a hybrid incremental learning approach for sentiment analysis of Arabic tweets. They used SVM, Logistic Regression and Recurrent Neural Network (RNN) classifiers for the classification and they built words lexicon, emoticon lexicon, idioms lexicon and some essential lexica. They also investigated the Levenshtein distance algorithm in sentiment analysis in order to deal with the different forms of words and spelling mistakes. The HILATSA approach has been tested by using six datasets. Five of them are used to build the lexicons in order to train, test and verify the classifier model, and the sixth one is used for evaluating and simulating the hybrid system.

4.4. Other approaches

4.4.1. Aspect-based approach

Aspect-based sentiment analysis is a fine-grained sentiment analysis task, which aims to predict the sentiment polarities of the given aspects or target terms in texts (e.g. product or service) [256]. Aspects can be attributes, characteristics or features of the target. Aspect-based sentiment classification consists of two stages: aspect-extraction and sentiment classification [257]. The first stage extracts aspects and groups synonyms of these aspects that may people use to refer to the same entity and the second stage aims to determine the sentiment of each aspect. Several approaches have been proposed for this level of sentiment analysis. Karagoz et al. [258] proposed a framework for aspect-based sentiment analysis on Turkish informal texts. For aspect extraction, they used an unsupervised approach presented in [259]. An explicit aspect-sentiment matching was used which consists of the steps of finding noun groups, finding sentiment word groups, matching noun groups to sentiment word groups, and extracting scores for aspects. The experimental results have shown that the proposed approach has more performance than the basic architectures of Long Short-Term Memory (LSTM) and Conditional Random Fields (CRF) models.

4.4.2. Transfer learning

Transfer learning is the method that uses the similarity of data, data distribution, model task, and so on to apply the knowledge already learned in one domain to the new domain [30]. This method has emerged as a new machine learning technique and it is very useful, especially for gaining time as there is no need to train an algorithm from scratch. For sentiment analysis, transfer learning is usually applied to transfer the obtained ability to perform sentiment classification from a domain to another. In the work of Meng et al. [260], a transfer learning method based on the multi-layer convolutional neural network (CNN) is proposed. The authors constructed a CNN model to extract features from a source domain dataset and share the weights in the convolutional and pooling layer between source and target domain samples. They completed the transfer of model from the source to the target domain and they fine-tuned the weights of the fully connected layer. Thus, there is no need to retrain the network for the target domain. This approach achieved relatively good results and proves that can solve the problem of the absence of the labeled data in the target domain. Similarly, Bartusiak et al. [261], used transfer learning to propose a new sentiment analysis method that evaluates the sentiment of documents in the Polish Language. In order to conduct transfer learning, they trained the proposed method on one dataset using Support Vector Machine and then validate it on another one.

4.4.3. Multimodal sentiment analysis

Multimodal sentiment analysis is a growing research field. In addition to text, this field aims to include other modalities such as audio and visual data in the process of sentiment analysis [262]. This is because the development of web 2.0 made people share sentiments within images, videos, and audios along with the text. Thus, sentiment analysis and affective computing have evolved to more complex forms of multimodal analysis instead of unimodal analysis [3]. A multimodal approach can be bimodal that uses different combinations of two modalities, or trimodal which includes three modalities. However, the fusion of multimodal content or information has several challenges such as fusion strategy, hyper-parameter tuning, interpretability, and speed [263]. Balazs and Velásquez [264] proposed a survey on information fusion applied to sentiment analysis and opinion mining. In this work, the authors defined opinion mining from different perspectives and explained many information fusion approaches. Besides, several studies that applied information fusion for opinion mining have been reviewed. In the same trend, Fan et al. [265] proposed an overview of information fusion processes and methods used to rank products based on online reviews. Majumder et al. [266] presented a novel feature fusion strategy to improve the multimodal fusion mechanism of three modalities. This strategy fuses the feature vectors of different modalities in a hierarchical fashion, instead of simply concatenating them. It fuses modalities two in two at the beginning, and then fuses all three modalities. The results show that this approach reduces the error rate with 5 to 10%. Other works focused on performing sentiment analysis with the incorporation of different modalities. Zhao et al. [267] proposed an image-text consistency driven multimodal sentiment analysis approach. This method explores the correlation between the image and the text. Moreover, it combines low-level visual and textual features to perform multimodal sentiment analysis. Several studies have been proposed using different modalities along with text such as physiological signal [268], video, and audio [269].

4.5. Evaluation metrics

The performance and the effectiveness of an approach or a proposed model are evaluated using different metrics. This last step of developing a model is very important because not all metrics are suitable for a given problem, and sometimes a new evaluation metric can be introduced to evaluate the newly proposed approach as in the work of Jiang et al. [270]. The choice of metrics can affect how the performance and effectiveness of a model are measured and compared. Among the techniques used for evaluating and summarizing the performance of a classification model, there is a Confusion Matrix (also called error matrix or truth table), Receiver Operator Characteristic (ROC) and Area Under the Curve (AUC). A basic confusion matrix is a 2 by 2 matrix as illustrated in Table 9 that summarizes the number of correct and incorrect samples predicted by a classifier, where:

- TP represents the number of positive samples that are predicted correctly as positive by the classifier.
- FP represents the number of negative samples that are predicted incorrectly as positive by the classifier.
- FN represents the number of positive samples that are predicted incorrectly as negative by the classifier.
- TN represents the number of negative samples that are predicted correctly as negative by the classifier.

These numbers are very useful for some measurement metrics which are summarized and briefly described in Table 10. A perfect classifier should have zero entry for FP and FN. Unfortunately, it is not the case in real life because any model can suffer from shortfalls which makes it not 100% accurate most of the time.

Other metrics have been used with or without the metrics mentioned above to evaluate several approaches. Among theme there is the ROC evaluation metric [152,168], AUC [168], Kappa [271] and Root Mean Square Error (RMSE) [205]. However, obtaining a model with good performance is not always easy and sometimes there is a need to solve some problems during or before the training process such as preventing overfitting or handling noisy data especially with machine learning algorithms. For example, K. Ravi and V. Ravi [75] used the L2 regularization technique to avoid overfitting for logistic regression classifier. Another technique to prevent overfitting and underfitting is Cross-Validation (CV). Although predicting sentiment more accurately is very important, it does not always offer complete information, because sentiments are subjective in nature. Akhtar et al. [272] proposed a stacked ensemble approach to predict the degree of sentiment intensity. They combined the outputs obtained from three deep learning models; LSTM, CNN and GRU using multiple layer perceptron. Predicting the level of sentiment does certainly help to understand the exact feeling of a given level of sentiment analysis.

5. Sentiment analysis challenges

5.1. Sarcasm detection

According to Macmillan English dictionary, sarcasm is defined as the activity of saying or writing the opposite of what someone means, or of speaking in a way intended to make someone else feel stupid or show them that he is angry [273,274]. The problem of sarcasm in sentiment analysis is for example when someone writes something positive but he actually means negative or vice versa which makes the task of sentiment analysis more complex. Sarcastic expressions are widely used in our daily lives. Thus, the interest in sarcasm detection is increasing to overcome the problem of getting deceitful sentiments by automatically identifying sarcastic expressions in a given text. The complexity and ambiguity of sarcasm make sarcasm detection a very challenging NLP task [275]. Many approaches have been proposed for sarcasm detection [275,276]. Jain et al. [277] used deep learning for real-time sarcasm detection in the mash-up of English with Indian native language (Hinglish). Their proposed model is a hybrid of bidirectional LSTM with a softmax attention layer and convolutional neural network. The softmax attention layer was used to learn the semantic context vector for English features from GloVe word representation and forward it to CNN. The CNN model supplemented also by HindiSenti (Hindi SentiWordNet) feature vector and auxiliary punctuation-based features combined. This model outperforms the baseline deep learning models with a superior classification accuracy of 92.71%.

5.2. Negation handling

Handling negation words such as not, neither, nor, etc. is very important for sentiment analysis because they can reverse the polarity of a given text. For example, given a sentence "This movie is good.", is classified as a positive sentence while "The movie is not good." should be classified as a negative sentence. Unfortunately, in some approaches, negation words are removed because they are included in Stop-Word lists or ignored implicitly because they have a neutral sentiment value in a lexicon which not impact the final polarity. However, it is not easy to handle this task by reversing the polarity because negation words can be found in a sentence without influencing the sentiment of the text. A syntactic path-based hybrid neural network for negation scope detection proposed by Lazib et al. [278]. This approach combined bidirectional LSTM and CNN where the CNN model used to capture relevant syntactic features between the token and the cue within the shortest syntactic path in both constituency and dependency parse trees, while The Bi-LSTM learns the context representation along the sentence in both forward and backward directions. Their model achieved a 90.82% F-score.

5.3. Spam detection

Spam detection plays an important role in the field of sentiment analysis. As online opinions influence the consumer purchase decisions, spam and fake reviews can damage the reputation of brands and artificially manipulate users' perceptions about products, services, companies, or other entities [279]. Developing a spam detection system that can identify fake reviews among many reviews is a very challenging task because there is no manifest difference between reviews. Among the systems that were proposed to perform the task of spam detection, a system developed by Saumya and Singh [280] which efficiently employs three features; sentiment of review and its comments, content-based factor, and rating deviation. This method makes use of the comment data to label the review as a spam or non-spam.

Table 9
Confusion matrix.

		Actual Class	
		Positive	Negative
Predicted class	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

Table 10
Description of the most common evaluation metrics for sentiment analysis.

Metric	Description	Calculation	Studies	Assessment
Accuracy	Accuracy is the most commonly used metric for classification problems, and it represents the ratio between the correctly predicted examples to the total number of examples. The complement of this metric is called *error* and can be calculated as 1-accuracy.	$\frac{TP+TN}{TP+TN+FP+FN}$	[41] [43] [65] [71] [168] [199]	Accuracy is a good choice for sentiment classification in machine learning when the classes in the dataset are nearly balanced.
Precision	The precision represents the proportion of samples that correctly predicted as positive to the total number of predicted positive samples. In other terms, precision measures the quality of being exact.	$\frac{TP}{TP+FP}$	[37] [164] [168] [248]	This metric is suitable for problems where prediction should be confident.
Recall (Sensitivity)	The recall is defined as the proportion of samples that correctly predicted as positive to all positive samples. Recall measures the misclassifying of the model.	$\frac{TP}{TP+FN}$	[37] [164] [168] [248]	Unlike precision, recall can be used with problems where the capture of a given class should be dominant (e.g. prediction depression). In this case, the prediction should not be very confident.
F1-score (F-measure)	F1-score is a number between 0 and 1 that helps to measure both precision and recall by calculating the harmonic mean of these two metrics.	$\frac{2*Precision.Recall}{Precision+Reacall}$	[43] [37] [59] [164] [168] [248]	Because it is not easy to compare two classifiers with high recall and low precision or vice versa, the F1-score can be the key to this problem. F1-score manages the tradeoff for a model that requires a confident prediction and dominant capture of a class.
Specificity	Specificity represents the opposite of recall metric.	$\frac{TN}{TN+FP}$	[168]	This metric can be used to conduct a confident precision.

The authors used those labeled data with a machine learning model to classify the other unlabeled data and two over-sampling techniques to make the class comparable because generally the number of spam reviews is much smaller that the number of real reviews. Their system achieved the F-score of 91%.

5.4. Anaphora and coreference resolution

Anaphora is a relation of coreference between linguistic terms [281]. In sentiment analysis and especially for aspect-based it is useful to identify what a pronoun refers to in a sentence because it helps to extract all the aspects of a given entity. Unfortunately, pronouns are usually ignored or removed in the preprocessing step. Sukthanker et al. [282] presented an exhaustive overview of the field of coreference resolution and the closely related field of anaphora resolution. An Enhanced Anaphora Resolution Algorithm was proposed by Deborah et al. [283]. This algorithm provides inter-sentential anaphora resolutions by uncovering compound nouns and resolving the PoS for each and every word. The algorithm achieved better performance compared to the traditional existing anaphora resolution methodology.

5.5. Word sense disambiguation (WSD)

A word can have different meanings and based on the context and the used domain the sense of this word can be different for each situation. Word sense disambiguation aims to determine which sense of a word has used in a sentence. For example, the word "curved" refers to a positive context if it is used with TV, and may refer to a negative sense if it is used with a mobile phone. Therefore, identifying a word sense from a sentence is highly challenging. A knowledge-based method which relies on the well-known lexicon WordNet was proposed by Wang et al. [284] to solve this challenging task. This method models the problem of WSD with semantic space and semantic path hidden behind a given sentence by using Latent Semantic Analysis (LSA) and PageRank respectively. The experimental results demonstrate the effectiveness of this method as it has achieved good performance. In a similar way, word polarity disambiguation (WPD) is another challenging problem. WPD aims to resolve polarity of the sentiment-ambiguous words in a specific context. Xia et al. [285] addressed this problem using Bayesian model and opinion-level features. They explored the level context by defining the intra and inter-opinion features. The Bayesian model was used to make the opinion-level features more effective and to resolve the polarity in a probabilistic manner.

5.6. Low-resource languages

In the field of sentiment analysis, most of the research works have focused on the English language [198], or other languages that have an acceptable amount of linguistic resources (e.g. sentiment lexicon and labeled text corpus). As mentioned before, supervised learning approaches are the most used for sentiment analysis. However, these approaches heavily rely on linguistic resources, which are costly to obtain for unpopular languages [202]. The types of languages that suffer from linguistic resource scarcity are called low-resource languages (or under-resourced languages). To over this problem several methods can be used: constructing linguistic resource from scratch; using unsupervised, semi-supervised and transfer learning approaches as mentioned in Section 4.1.2, Section 4.1.3, and Section 4.4.2, respectively. Zhou et al. [225] proposed an approach to exploit the rich English resources for Chinese sentiment classification. They trained two denoising autoencoder classifiers in English view and Chinese view, respectively. And then they combined the two results in two views to obtain the final sentiment classification results. This proves that cross-lingual sentiment classification approaches can be very useful for low-resource languages. Ren et al. [202] proposed a graph-based semi-supervised approach to solve the problem of document-level sentiment classification for under-resourced languages. With few labeled data, the authors investigated the usefulness of two graph-based algorithms, namely label propagation (LP) and modified adsorption (MAD). These methods help to increase labeled instances, thus more training data are available for scarce resource languages.

5.7. Sentiment analysis of code-mixed data

Code-Mixing (CM) is the use of vocabulary and syntax from multiples languages in the same sentence [286,287]. It is quite common in multilingual societies and poses a great challenge to NLP tasks such as sentiment analysis. The lack of a formal grammar of code-mixed sentences hampers the identification of compositional semantics, which are very important to conduct sentiment analysis using rule- and machine learning-based techniques. Besides, since the mixing is up to the person, there are non-determined mixing rules, which is one of the main hardships [287]. As a result, there is a need for new language models to handle sentiment analysis on code-mixed data. In the work of Chatterjee et al. [288] the authors explored the problem of language modeling for code-mixed Hinglish (Hindi–English language pair) text. Their study shows that switching points (where the person switches to another language) represent the essential problem for CM language model and the reason why traditional models would have a bad performance. However, although CM is a great challenge, few studies have addressed it as the work of Lal et al. [289]. The authors proposed a hybrid architecture for the task of sentiment analysis of English–Hindi code-mixed data. They divided this architecture into three components where each one of them seeking to handle different problems. The experiments on code-mixed social media dataset demonstrate that the proposed architecture can achieve an accuracy of 83.54%.

6. Conclusion and future works

This paper presented an overview of sentiment analysis and its related approaches. The main purpose of this paper is to examine and categorize the most used classification techniques to conduct the task of sentiment analysis. First, some of the most application domains were introduced, and then the process of classifying sentiment was briefly detailed including the important procedures such as pre-processing and feature selection. Various sentiment classification techniques were categorized and examined with their advantages and drawbacks. Supervised machine learning algorithms generally are the most used technique in this field due to their simplicity and high accuracy. Classification with Naïve Bayes and Support vector machine algorithms usually considered as baseline methods to compare the newly proposed approaches. However, other techniques (e.g., reinforcement learning) pose a powerful solution to some problems and challenges in the field such as the absence of labeled data or other related NLP tasks. The challenges presented later demonstrate that sentiment analysis remains an open research field.

The English language is the most tackled in this field, but recently other natural languages have gained more interest. Resources for these languages are still scarce. Therefore, addressing other natural languages than English by constructing useful resources such as building datasets and generating lexicons can be an interesting future work.

CRediT authorship contribution statement

Marouane Birjali: Conceptualization, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Mohammed Kasri:** Conceptualization, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Abderrahim Beni-Hssane:** Conceptualization, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] B. Liu, Sentiment Analysis, 2015, pp. 1–367, https://doi.org/10.1017/CBO9781139084789.
[2] I. Chaturvedi, E. Cambria, R.E. Welsch, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: Survey and challenges, Inf. Fusion. 44 (2018) 65–77, https://doi.org/10.1016/j.inffus.2017.12.006.
[3] S. Poria, E. Cambria, R. Bajpai, A. Hussain, A review of affective computing: From unimodal analysis to multimodal fusion, Inf. Fusion. 37 (2017) 98–125, https://doi.org/10.1016/j.inffus.2017.02.003.
[4] J.F. Sánchez-Rada, C.A. Iglesias, Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison, Inf. Fusion. 52 (2019) 344–356, https://doi.org/10.1016/j.inffus.2019.05.003.
[5] E. Cambria, Affective computing and sentiment analysis, IEEE Intell. Syst. 31 (2016) 102–107, https://doi.org/10.1109/MIS.2016.31.
[6] B. Schuller, A.E.D. Mousa, V. Vryniotis, Sentiment analysis and opinion mining: On optimal parameters and performances, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 5 (2015) 255–263, https://doi.org/10.1002/widm.1159.
[7] Y. Choi, H. Lee, Data properties and the performance of sentiment classification for electronic commerce applications, Inf. Syst. Front. 19 (2017) 993–1012, https://doi.org/10.1007/s10796-017-9741-7.
[8] S. Ju, S. Li, Active Learning on Sentiment Classification By Selecting Both Words and Documents, 2013, pp. 49–57, https://doi.org/10.1007/978-3-642-36337-5_6.
[9] E. Cambria, D. Das, S. Bandyopadhyay, A. Feraco, eds., A Practical Guide To Sentiment Analysis, 5, 2017, pp. 1–196, https://doi.org/10.1007/978-3-319-55394-8.
[10] A. Valdivia, M.V. Luzón, E. Cambria, F. Herrera, Consensus vote models for detecting and filtering neutrality in sentiment analysis, Inf. Fusion. 44 (2018) 126–135, https://doi.org/10.1016/j.inffus.2018.03.007.
[11] M. Erritali, A. Beni-Hssane, M. Birjali, Y. Madani, An approach of semantic similarity measure between documents based on big data, Int. J. Electr. Comput. Eng 6 (2016) 2454–2461, https://doi.org/10.11591/ijece.v6i5.pp2454-2461.
[12] M. Birjali, A. Beni-Hssane, M. Erritali, Measuring documents similarity in large corpus using MapReduce algorithm, in: 2016 5th Int. Conf. Multimed. Comput. Syst, IEEE, 2016, pp. 24–28, https://doi.org/10.1109/ICMCS.2016.7905587.
[13] L. Liu, X. Nie, H. Wang, Toward a fuzzy domain sentiment ontology tree for sentiment analysis, in: 2012 5th Int. Congr. Image Signal Process., IEEE, 2012, pp. 1620–1624, https://doi.org/10.1109/CISP.2012.6469930.
[14] F.J. Ramírez-Tinoco, G. Alor-Hernández, J.L. Sánchez-Cervantes, B.A. Olivares-Zepahua, L. Rodríguez-Mazahua, A Brief Review on the Use of Sentiment Analysis Approaches in Social Networks, 2018, pp. 263–273, https://doi.org/10.1007/978-3-319-69341-5_24.

[15] S. Mukherjee, Sentiment analysis of reviews, in: Encycl. Soc. Netw. Anal. Min, Springer New York, New York, NY, 2017, pp. 1–10, https://doi.org/10.1007/978-1-4614-7163-9_110169-1.
[16] M. Rushdi Saleh, M.T. Martín-Valdivia, A. Montejo-Ráez, L.A. Ureña López, Experiments with SVM to classify opinions in different domains, Expert Syst. Appl. 38 (2011) 14799–14804, https://doi.org/10.1016/j.eswa.2011.05.070.
[17] B. O'Connor, R. Balasubramanyan, B.R. Routledge, N.A. Smith, From tweets to polls: Linking text sentiment to public opinion time series, in: W.W. Cohen, S. Gosling (Eds.), Proc. Fourth Int. Conf. Weblogs Soc. Media, The AAAI Press, 2010, pp. 1–8.
[18] M.V. Mäntylä, D. Graziotin, M. Kuutila, The evolution of sentiment analysis—A review of research topics, venues, and top cited papers, Comput. Sci. Rev. 27 (2018) 16–32, https://doi.org/10.1016/j.cosrev.2017.10.002.
[19] B. Liu, L. Zhang, A survey of opinion mining and sentiment analysis, in: Min. Text Data, Springer, US, Boston, MA, 2012, pp. 415–463, https://doi.org/10.1007/978-1-4614-3223-4_13.
[20] R. Piryani, D. Madhavi, V.K. Singh, Analytical mapping of opinion mining and sentiment analysis research during 2000–2015, Inf. Process. Manag. 53 (2017) 122–150, https://doi.org/10.1016/j.ipm.2016.07.001.
[21] W. Medhat, A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: A survey, Ain Shams Eng. J. 5 (2014) 1093–1113, https://doi.org/10.1016/j.asej.2014.04.011.
[22] K. Ravi, V. Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications, Knowledge-Based Syst. 89 (2015) 14–46, https://doi.org/10.1016/j.knosys.2015.06.015.
[23] L. Zhang, S. Wang, B. Liu, Deep learning for sentiment analysis: A survey, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 8 (2018) 1–25, https://doi.org/10.1002/widm.1253.
[24] F. Hemmatian, M.K. Sohrabi, A survey on classification techniques for opinion mining and sentiment analysis, Artif. Intell. Rev. 52 (2019) 1495–1545, https://doi.org/10.1007/s10462-017-9599-6.
[25] S. Rajalakshmi, S. Asha, N. Pazhaniraja, A comprehensive survey on sentiment analysis, in: 2017 Fourth Int. Conf. Signal Process. Commun. Netw, IEEE, 2017, pp. 1–5, https://doi.org/10.1109/ICSCN.2017.8085673.
[26] B. Batrinca, P.C. Treleaven, Social media analytics: a survey of techniques, tools and platforms, AI Soc. 30 (2015) 89–116, https://doi.org/10.1007/s00146-014-0549-4.
[27] V.A. Kharde, P.S. Sonawane, Sentiment Analysis of Twitter Data: A Survey of Techniques, 2016, pp. 5–15, https://doi.org/10.5120/ijca2016908625.
[28] P. Yang, Y. Chen, A survey on sentiment analysis by using machine learning methods, in: 2017 IEEE 2nd Inf. Technol. Networking, Electron. Autom. Control Conf, IEEE, 2017, pp. 117–121, https://doi.org/10.1109/ITNEC.2017.8284920.
[29] H. Thakkar, D. Patel, Approaches for Sentiment Analysis on Twitter: A State-of-Art Study, 2015, pp. 1–8, http://arxiv.org/abs/1512.01043.
[30] R. Liu, Y. Shi, C. Ji, M. Jia, A survey of sentiment analysis based on transfer learning, IEEE Access. 7 (2019) 85401–85412, https://doi.org/10.1109/ACCESS.2019.2925059.
[31] N.F.F. Da Silva, L.F.S. Coletta, E.R. Hruschka, A survey and comparative study of tweet sentiment analysis via semi-supervised learning, ACM Comput. Surv. 49 (2016) 1–26, https://doi.org/10.1145/2932708.
[32] H.H. Do, P. Prasad, A. Maag, A. Alsadoon, Deep learning for aspect-based sentiment analysis: A comparative review, Expert Syst. Appl. 118 (2019) 272–299, https://doi.org/10.1016/j.eswa.2018.10.003.
[33] C.C. Aggarwal, Machine learning for text, Mach. Learn. Text. (2018) 1–493, https://doi.org/10.1007/978-3-319-73531-3.
[34] S. Behdenna, F. Barigou, G. Belalem, Sentiment analysis at document level, in: SmartCom 2016, 2016, pp. 159–168, https://doi.org/10.1007/978-981-10-3433-6_20.
[35] N. Indurkhya, F.J. Damerau, Handbook of Natural Language Processing, 2010, pp. 1–704.
[36] M. Tubishat, N. Idris, M.A.M. Abushariah, Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges, Inf. Process. Manag. 54 (2018) 545–563, https://doi.org/10.1016/j.ipm.2018.03.008.
[37] M.E. Mowlaei, M. Saniee Abadeh, H. Keshavarz, Aspect-based sentiment analysis using adaptive aspect-based lexicons, Expert Syst. Appl. 148 (2020) 1–13, https://doi.org/10.1016/j.eswa.2020.113234.
[38] L. Mai, B. Le, Joint sentence and aspect-level sentiment analysis of product comments, Ann. Oper. Res. (2020) https://doi.org/10.1007/s10479-020-03534-7.
[39] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding, 2018, pp. 1–16, http://arxiv.org/abs/1810.04805.
[40] B. Liu, Sentiment analysis and opinion mining, Synth. Lect. Hum. Lang. Technol. 5 (2012) 1–167, https://doi.org/10.2200/S00416ED1V01Y201204HLT016.
[41] T. Chen, R. Xu, Y. He, X. Wang, Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN, Expert Syst. Appl. 72 (2017) 221–230, https://doi.org/10.1016/j.eswa.2016.10.065.
[42] O. Alqaryouti, N. Siyam, A.A. Monem, K. Shaalan, Aspect-based sentiment analysis using smart government review data, Appl. Comput. Informatics. (2019) 1–20, https://doi.org/10.1016/j.aci.2019.11.003.
[43] Z. Zhao, G. Rao, Z. Feng, DFDS: A Domain-Independent Framework for Document-Level Sentiment Analysis Based on RST, 2017, pp. 297–310, https://doi.org/10.1007/978-3-319-63579-8_23.
[44] S. Kumar, M. Yadava, P.P. Roy, Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction, Inf. Fusion. 52 (2019) 41–52, https://doi.org/10.1016/j.inffus.2018.11.001.
[45] R. Bose, R.K. Dey, S. Roy, D. Sarddar, Sentiment Analysis on Online Product Reviews, 2020, pp. 559–569, https://doi.org/10.1007/978-981-13-7166-0_56.
[46] A. Rajput, Natural Language Processing Sentiment Analysis and Clinical Analytics, 2019, pp. 1–18, http://arxiv.org/abs/1902.00679.
[47] M. Birjali, A. Beni-Hssane, M. Erritali, Learning with big data technology: The future of education, in: Adv. Intell. Syst. Comput., 2018, pp. 209–217, https://doi.org/10.1007/978-3-319-60834-1_22.
[48] I. Yaqoob, I.A.T. Hashem, A. Gani, S. Mokhtar, E. Ahmed, N.B. Anuar, A.V. Vasilakos, Big data: From beginning to future, Int. J. Inf. Manage. 36 (2016) 1231–1247, https://doi.org/10.1016/j.ijinfomgt.2016.07.009.
[49] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, Cloud computing — The business perspective, Decis. Support Syst. 51 (2011) 176–189, https://doi.org/10.1016/j.dss.2010.12.006.
[50] J. Frizzo-Barker, P.A. Chow-White, P.R. Adams, J. Mentanko, D. Ha, S. Green, Blockchain as a disruptive technology for business: A systematic review, Int. J. Inf. Manage. 51 (2020) 1–14, https://doi.org/10.1016/j.ijinfomgt.2019.10.014.
[51] J. Bernabé-Moreno, A. Tejeda-Lorente, J. Herce-Zelaya, C. Porcel, E. Herrera-Viedma, A context-aware embeddings supported method to extract a fuzzy sentiment polarity dictionary, Knowledge-Based Syst. 190 (2020) 1–13, https://doi.org/10.1016/j.knosys.2019.105236.
[52] L. Rognone, S. Hyde, S.S. Zhang, News sentiment in the cryptocurrency market: An empirical comparison with forex, Int. Rev. Financ. Anal. 69 (2020) 1–17, https://doi.org/10.1016/j.irfa.2020.101462.
[53] O. Kraaijeveld, J. De Smedt, The predictive power of public Twitter sentiment for forecasting cryptocurrency prices, J. Int. Financ. Mark. Institutions Money. 65 (2020) 1–22, https://doi.org/10.1016/j.intfin.2020.101188.
[54] T.W. Jing, R.K. Murugesan, A Theoretical Framework To Build Trust and Prevent Fake News in Social Media using Blockchain, 2019, pp. 955–962, https://doi.org/10.1007/978-3-319-99007-1_88.
[55] Z.Y. Khan, Z. Niu, S. Sandiwarno, R. Prince, Deep learning techniques for rating prediction: a survey of the state-of-the-art, Artif. Intell. Rev. (2020) 1–41, https://doi.org/10.1007/s10462-020-09892-9.
[56] J. Serrano-Guerrero, J.A. Olivas, F.P. Romero, A T1OWA and aspect-based model for customizing recommendations on eCommerce, Appl. Soft Comput. 97 (2020) 1–10, https://doi.org/10.1016/j.asoc.2020.106768.
[57] B. Ray, A. Garain, R. Sarkar, An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews, Appl. Soft Comput. 98 (2021) 1–18, https://doi.org/10.1016/j.asoc.2020.106935.
[58] X. Fu, T. Ouyang, Z. Yang, S. Liu, A product ranking method combining the features–opinion pairs mining and interval-valued Pythagorean fuzzy sets, Appl. Soft Comput. 97 (2020) 1–17, https://doi.org/10.1016/j.asoc.2020.106803.
[59] H. Li, J. Cui, B. Shen, J. Ma, An intelligent movie recommendation system through group-level sentiment analysis in microblogs, Neurocomputing. 210 (2016) 164–173, https://doi.org/10.1016/j.neucom.2015.09.134.
[60] R.-P. Shen, H.-R. Zhang, H. Yu, F. Min, Sentiment based matrix factorization with reliability for recommendation, Expert Syst. Appl. 135 (2019) 249–258, https://doi.org/10.1016/j.eswa.2019.06.001.
[61] M. Birjali, A. Beni-Hssane, M. Erritali, A novel adaptive e-learning model based on Big Data by using competence-based knowledge and social learner activities, Appl. Soft Comput. J. 69 (2018) 1–34, https://doi.org/10.1016/j.asoc.2018.04.030.
[62] E. Georgiadou, S. Angelopoulos, H. Drake, Big data analytics and international negotiations: Sentiment analysis of brexit negotiating outcomes, Int. J. Inf. Manage. 51 (2020) 1–9, https://doi.org/10.1016/j.ijinfomgt.2019.102048.
[63] S.M. Zavattaro, P.E. French, S.D. Mohanty, A sentiment analysis of U.S local government tweets: The connection between tone and citizen involvement, Gov. Inf. Q. 32 (2015) 333–341, https://doi.org/10.1016/j.giq.2015.03.003.
[64] F. Falck, J. Marstaller, N. Stoehr, S. Maucher, J. Ren, A. Thalhammer, A. Rettinger, R. Studer, Measuring Proximity Between Newspapers and Political Parties: The Sentiment Political Compass, Policy & Internet, 2019, pp. 1–33, https://doi.org/10.1002/poi3.222.

[65] I. El Alaoui, Y. Gahi, R. Messoussi, Y. Chaabi, A. Todoskoff, A. Kobi, A novel adaptable approach for sentiment analysis on big social data, J. Big Data. 5 (2018) 1–18, https://doi.org/10.1186/s40537-018-0120-0.
[66] M. Birjali, A. Beni-Hssane, M. Erritali, Evaluation of high-level query languages based on mapreduce in big data, J. Big Data. (2018) 1–21, https://doi.org/10.1186/s40537-018-0146-3.
[67] M. Birjali, A. Beni-Hssane, M. Erritali, Analyzing social media through big data using infosphere biginsights and apache flume, in: Procedia Comput. Sci, 2017, pp. 1–6, https://doi.org/10.1016/j.procs.2017.08.299.
[68] F.J. Ramírez-Tinoco, G. Alor-Hernández, J.L. Sánchez-Cervantes, M. del P. Salas-Zárate, R. Valencia-García, Use of Sentiment Analysis Techniques in Healthcare Domain, 2019, pp. 189–212, https://doi.org/10.1007/978-3-030-06149-4_8.
[69] S.M. Jiménez-Zafra, M.T. Martín-Valdivia, M.D. Molina-González, L.A. Ureña López, How do we talk about doctors and drugs? Sentiment analysis in forums expressing opinions for medical domain, Artif. Intell. Med. 93 (2019) 50–57, https://doi.org/10.1016/j.artmed.2018.03.007.
[70] E.M. Clark, T. James, C.A. Jones, A. Alapati, P. Ukandu, C.M. Danforth, P.S. Dodds, A Sentiment Analysis of Breast Cancer Treatment Experiences and Healthcare Perceptions Across Twitter, 2018, pp. 1–17, http://arxiv.org/abs/1805.09959.
[71] D. Ayata, Y. Yaslan, M.E. Kamasak, Emotion recognition from multimodal physiological signals for emotion aware healthcare systems, J. Med. Biol. Eng. 40 (2020) 149–157, https://doi.org/10.1007/s40846-019-00505-7.
[72] E. Cambria, S. Poria, A. Gelbukh, M. Thelwall, Sentiment analysis is a big suitcase, IEEE Intell. Syst. 32 (2017) 74–80, https://doi.org/10.1109/MIS.2017.4531228.
[73] J. Serrano-Guerrero, J.A. Olivas, F.P. Romero, E. Herrera-Viedma, Sentiment analysis: A review and comparative analysis of web services, Inf. Sci. (Ny). 311 (2015) 18–38, https://doi.org/10.1016/j.ins.2015.03.040.
[74] C. Diamantini, A. Mircoli, D. Potena, A negation handling technique for sentiment analysis, in: 2016 Int. Conf. Collab. Technol. Syst, IEEE, 2016, pp. 188–195, https://doi.org/10.1109/CTS.2016.0048.
[75] K. Ravi, V. Ravi, A novel automatic satire and irony detection using ensembled feature selection and data mining, Knowledge-Based Syst. 120 (2017) 1–40, https://doi.org/10.1016/j.knosys.2016.12.018.
[76] M.E. Moussa, E.H. Mohamed, M.H. Haggag, A survey on opinion summarization techniques for social media, Futur. Comput. Informatics J. 3 (2018) 82–109, https://doi.org/10.1016/j.fcij.2017.12.002.
[77] B. O'Dea, S. Wan, P.J. Batterham, A.L. Calear, C. Paris, H. Christensen, Detecting suicidality on twitter, Internet Interv. 2 (2015) 183–188, https://doi.org/10.1016/j.invent.2015.03.005.
[78] M. Parmar, B. Maturi, J.M. Dutt, H. Phate, Sentiment analysis on interview transcripts: An application of NLP for quantitative analysis, in: 2018 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2018, 2018, pp. 1063–1068, https://doi.org/10.1109/ICACCI.2018.8554498.
[79] J.A. Caetano, H.S. Lima, M.F. Santos, H.T. Marques-Neto, Using sentiment analysis to define twitter political users' classes and their homophily during the 2016 American presidential election,, J. Internet Serv. Appl. 9 (2016) 1–15, https://doi.org/10.1186/s13174-018-0089-0.
[80] Z. Xiaomei, Y. Jing, Z. Jianpei, H. Hongyu, Microblog sentiment analysis with weak dependency connections, Knowledge-Based Syst. 142 (2018) 170–180, https://doi.org/10.1016/j.knosys.2017.11.035.
[81] J. Mei, R. Frank, Sentiment crawling, in: Proc. 2015 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. 2015 - ASONAM '15, ACM Press, New York, New York, USA, 2015, pp. 1024–1027, https://doi.org/10.1145/2808797.2809373.
[82] P. Anderson, What is web 2.0? Ideas technologies and implications for education, 2007, pp. 1–64, http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf.
[83] Y.H. Gu, S.J. Yoo, Z. Jiang, Y.J. Lee, Z. Piao, H. Yin, S. Jeon, Sentiment analysis and visualization of Chinese tourism blogs and reviews, in: 2018 Int. Conf. Electron. Information, Commun, IEEE, 2018, pp. 1–4, https://doi.org/10.23919/ELINFOCOM.2018.8330589.
[84] R. Folgieri, M. Bait, J.P.M. Carrion, A Cognitive Linguistic and Sentiment Analysis of Blogs: Monterosso 2011 Flooding, 2016, pp. 499–522, https://doi.org/10.1007/978-3-319-27528-4_34.
[85] A. Lommatzsch, F. Bütow, D. Ploch, S. Albayrak, Towards the Automatic Sentiment Analysis of German News and Forum Documents, 2017, pp. 18–33, https://doi.org/10.1007/978-3-319-60447-3_2.
[86] I. Korkontzelos, A. Nikfarjam, M. Shardlow, A. Sarker, S. Ananiadou, G.H. Gonzalez, Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts, J. Biomed. Inform. 62 (2016) 148–158, https://doi.org/10.1016/j.jbi.2016.06.007.
[87] J. Neidhardt, N. Rümmele, H. Werthner, Predicting happiness: user interactions and sentiment analysis in an online travel forum, Inf. Technol. Tour. 17 (2017) 101–119, https://doi.org/10.1007/s40558-017-0079-2.
[88] P. Choudhury, D. Wang, N.A. Carlson, T. Khanna, Machine learning approaches to facial and text analysis: Discovering CEO oral communication styles, Strateg. Manag. J. 40 (2019) 1705–1732, https://doi.org/10.1002/smj.3067.
[89] A. Go, R. Bhayani, L. Huang, Twitter Sentiment classification using distant supervision, Processing 150 (2009) 1–6.
[90] A.L. Maas, R.E. Daly, P.T. Pham, D. Huang, A.Y. Ng, C. Potts, Learning Word Vectors for Sentiment Analysis, Association for Computational Linguistics, Portland, Oregon, USA, 2011, pp. 142–150, http://www.aclweb.org/anthology/P11-1015.
[91] R. Feldman, Techniques and applications for sentiment analysis, Commun. ACM. 56 (2013) 1–8, https://doi.org/10.1145/2436256.2436274.
[92] Y. Wang, B. Li, Sentiment analysis for social media images, in: 2015 IEEE Int. Conf. Data Min. Work, IEEE, 2015, pp. 1584–1591, https://doi.org/10.1109/ICDMW.2015.142.
[93] S. Amiriparian, N. Cummins, S. Ottl, M. Gerczuk, B. Schuller, Sentiment analysis using image-based deep spectrum features, in: 2017 Seventh Int. Conf. Affect. Comput. Intell. Interact. Work. Demos, IEEE, 2017, pp. 26–29, https://doi.org/10.1109/ACIIW.2017.8272618.
[94] A. Pinto, H.G. Oliveira, A.O. Alves, Comparing the performance of different NLP toolkits in formal and social media text, in: H.G. Oliveira M. Mernik (Ed.), 5th Symp. Lang. Appl. Technol, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2016, pp. 3:1–3:16, https://doi.org/10.4230/OASIcs.SLATE.2016.3.
[95] S. Bird, E. Klein, E. Loper, Natural Language Processing with Python, 2009, pp. 1–504.
[96] C.D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S.J. Bethard, D. McClosky, The Stanford CoreNLP natural language processing toolkit, in: Assoc. Comput. Linguist. Syst. Demonstr, 2014, pp. 55–60, http://www.aclweb.org/anthology/P/P14/P14-5010.
[97] T. Kwartler, The OpenNLP project, in: Text Min. Pract. with R, John Wiley & Sons, Ltd, Chichester, UK, 2017, pp. 237–269, https://doi.org/10.1002/9781119282105.ch8.
[98] A. Pasha, M. Al-Badrashiny, M. Diab, A. El Kholy, R. Eskander, N. Habash, M. Pooleery, O. Rambow, R. Roth, MADAMIRA: A fast, comprehensive tool for morphological analysis and disambiguation of Arabic, in: Proc. Ninth Int. Conf. Lang. Resour. Eval, European Language Resources Association (ELRA), Reykjavik, Iceland, 2014, pp. 1094–1101, http://www.lrec-conf.org/proceedings/lrec2014/pdf/593_Paper.pdf.
[99] H. Liang, X. Sun, Y. Sun, Y. Gao, Text feature extraction based on deep learning: a review, EURASIP J. Wireless Commun. Networking 2017 (2017) 1–12, https://doi.org/10.1186/s13638-017-0993-1.
[100] M. Kasri, M. Birjali, A. Beni-Hssane, A comparison of features extraction methods for Arabic sentiment analysis, in: Proc. 4th Int. Conf. Big Data Internet Things, ACM, New York, NY, USA, 2019, pp. 1–6, https://doi.org/10.1145/3372938.3372998.
[101] M. Avinash, E. Sivasankar, A Study of Feature Extraction Techniques for Sentiment Analysis, 2019, pp. 475–486, https://doi.org/10.1007/978-981-13-1501-5_41.
[102] M. Venugopalan, D. Gupta, Exploring sentiment analysis on twitter data, in: 2015 Eighth Int. Conf. Contemp. Comput, IEEE, 2015, pp. 241–247, https://doi.org/10.1109/IC3.2015.7346686.
[103] K. Toutanova, D. Klein, C.D. Manning, Y. Singer, Feature-rich part-of-speech tagging with a cyclic dependency network, in: Proc. 2003 Hum. Lang. Technol. Conf. North {A}merican Chapter Assoc. Comput. Linguist, 2003, pp. 252–259, https://www.aclweb.org/anthology/N03-1033.
[104] V. Hatzivassiloglou, J.M. Wiebe, Effects of Adjective Orientation and Gradability on Sentence Subjectivity, in: COLING 2000 1 18th Int. Conf. Comput. Linguist., 2000 pp. 1–7. https://www.aclweb.org/anthology/C00-1044.
[105] C.C. Aggarwal, C. Zhai, eds., Mining Text Data, 2012, pp. 1–526, https://doi.org/10.1007/978-1-4614-3223-4.
[106] L. Polanyi, A. Zaenen, Contextual Valence Shifters, in: Comput. Attitude Affect Text Theory Appl., Springer-Verlag, Berlin/Heidelberg, n.d., pp. 1–10. https://doi.org/10.1007/1-4020-4102-0_1.
[107] V. Loia, S. Senatore, A fuzzy-oriented sentic analysis to capture the human emotion in web-based content, Knowledge-Based Syst. 58 (2014) 75–85, https://doi.org/10.1016/j.knosys.2013.09.024.
[108] R. Dzisevič, D. Šešok, Text classification using different feature extraction approaches, in: 2019 Open Conf. Electr. Electron. Inf. Sci, 2019, pp. 1–4, https://doi.org/10.1109/eStream.2019.8732167.
[109] S. Sadhukhan, S. Banerjee, P. Das, A.K. Sangaiah, Producing better disaster management plan in post-disaster situation using social media mining, in: Comput. Intell. Multimed. Big Data Cloud with Eng. Appl, Elsevier, 2018, pp. 171–183, https://doi.org/10.1016/B978-0-12-813314-9.00009-8.
[110] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient Estimation of Word Representations in Vector Space, 2013, pp. 1–12, http://arxiv.org/abs/1301.3781.
[111] J. Pennington, R. Socher, C.D. Manning, GloVe: Global vectors for word representation, in: Empir. Methods Nat. Lang. Process, 2014, pp. 1532–1543, http://www.aclweb.org/anthology/D14-1162.
[112] A. Yadav, D.K. Vishwakarma, Sentiment analysis using deep learning architectures: a review, Artif. Intell. Rev. (2019) 1–51, https://doi.org/10.1007/s10462-019-09794-5.

[113] Q. Le, T. Mikolov, Distributed representations of sentences and documents, in: Proc. 31st Int. Conf. Int. Conf. Mach. Learn. - vol. 32, JMLR.org, 2014, pp. II–1188–II–1196.
[114] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriching Word Vectors with Subword Information, 2016, pp. 1–12, http://arxiv.org/abs/1607.04606.
[115] M. Kasri, M. Birjali, A. Beni-Hssane, Word2Sent: A new learning sentiment-embedding model with low dimension for sentence level sentiment classification, Concurr. Comput. Pract. Exp. (2020) 1–12, https://doi.org/10.1002/cpe.6149.
[116] S.R. Ahmad, A.A. Bakar, M.R. Yaakub, A review of feature selection techniques in sentiment analysis, Intell. Data Anal. 23 (2019) 159–189, https://doi.org/10.3233/IDA-173763.
[117] R. Kumar, J. Kaur, Random Forest-Based Sarcastic Tweet Classification using Multiple Feature Collection, 2020, pp. 131–160, https://doi.org/10.1007/978-981-13-8759-3_5.
[118] S. Baccianella, A. Esuli, F. Sebastiani, SentiWordNet, Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining, in: Proc. Int. Conf. Lang. Resour. Eval. {LREC} 2010, 17-23 2010, European Language Resources Association, Valletta, Malta, 2010, pp. 1–5, http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf.
[119] A. Duric, F. Song, Feature selection for sentiment analysis based on content and syntax models, Decis. Support Syst. 53 (2012) 704–711, https://doi.org/10.1016/j.dss.2012.05.023.
[120] N. Hoque, D.K. Bhattacharyya, J.K. Kalita, MIFS-ND: A mutual information-based feature selection method, Expert Syst. Appl. 41 (2014) 6371–6385, https://doi.org/10.1016/j.eswa.2014.04.019.
[121] G. Kou, P. Yang, Y. Peng, F. Xiao, Y. Chen, F.E. Alsaadi, Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods, Appl. Soft Comput. 86 (2020) 1–14, https://doi.org/10.1016/j.asoc.2019.105836.
[122] M. Birjali, A. Beni-Hssane, M. Erritali, Machine learning and semantic sentiment analysis based algorithms for suicide sentiment prediction in social networks, Procedia Comput. Sci. 113 (2017) 65–72, https://doi.org/10.1016/j.procs.2017.08.290.
[123] N. Sánchez-Maroño, A. Alonso-Betanzos, M. Tombilla-Sanromán, Filter Methods for Feature Selection – A Comparative Study, in: Intell. Data Eng. Autom. Learn. - IDEAL 2007, Springer Berlin Heidelberg, Berlin, Heidelberg, n.d., pp. 178–187. https://doi.org/10.1007/978-3-540-77226-2_19.
[124] Y. Yang, J.O. Pedersen, A comparative study on feature selection in text categorization, in: Proc. Fourteenth Int. Conf. Mach. Learn, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997, pp. 412–420.
[125] L. Paninski, Estimation of entropy and mutual information, Neural Comput. 15 (2003) 1191–1253, https://doi.org/10.1162/089976603321780272.
[126] R.J. Urbanowicz, M. Meeker, W. La Cava, R.S. Olson, J.H. Moore, Relief-based feature selection: Introduction and review, J. Biomed. Inform. 85 (2018) 189–203, https://doi.org/10.1016/j.jbi.2018.07.014.
[127] N. Sánchez-Maroño, A. Alonso-Betanzos, R.M. Calvo-Estévez, A Wrapper Method for Feature Selection in Multiple Classes Datasets, 2009, pp. 456–463, https://doi.org/10.1007/978-3-642-02478-8_57.
[128] J.C. Cortizo, I. Giraldez, Multi Criteria Wrapper Improvements To Naive Bayes Learning, 2006, pp. 419–427, https://doi.org/10.1007/11875581_51.
[129] S. Maldonado, R. Weber, F. Famili, Feature selection for high-dimensional class-imbalanced data sets using support vector machines, Inf. Sci. (Ny). 286 (2014) 228–246, https://doi.org/10.1016/j.ins.2014.07.015.
[130] H. Liu, M. Zhou, Q. Liu, An embedded feature selection method for imbalanced data classification, IEEE/CAA J. Autom. Sin. 6 (2019) 703–715, https://doi.org/10.1109/JAS.2019.1911447.
[131] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1157–1182.
[132] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B. 58 (1996) 267–288, http://www.jstor.org/stable/2346178.
[133] G. Ansari, T. Ahmad, M.N. Doja, Hybrid filter–wrapper feature selection method for sentiment classification, Arab. J. Sci. Eng. 44 (2019) 9191–9208, https://doi.org/10.1007/s13369-019-04064-6.
[134] S.R. Ahmad, A.A. Bakar, M.R. Yaakub, Ant colony optimization for text feature selection in sentiment analysis, Intell. Data Anal. 23 (2019) 133–158, https://doi.org/10.3233/IDA-173740.
[135] Huan Liu, Lei Yu, Toward integrating feature selection algorithms for classification and clustering, IEEE Trans. Knowl. Data Eng. 17 (2005) 491–502, https://doi.org/10.1109/TKDE.2005.66.
[136] A. Collomb, L. Brunie, C. Costea, A study and comparison of sentiment analysis methods for reputation evaluation, in: Cogn. Informatics Soft Comput, 2013, pp. 1–10, https://liris.cnrs.fr/Documents/Liris-6508.pdf.
[137] Z. Madhoushi, A.R. Hamdan, S. Zainudin, Sentiment analysis techniques in recent works, in: 2015 Sci. Inf. Conf, IEEE, 2015, pp. 288–291, https://doi.org/10.1109/SAI.2015.7237157.
[138] H. Sankar, V. Subramaniyaswamy, Investigating sentiment analysis using machine learning approach, in: 2017 Int. Conf. Intell. Sustain. Syst, IEEE, 2017, pp. 87–92, https://doi.org/10.1109/ISS1.2017.8389293.
[139] A. Jurek, M.D. Mulvenna, Y. Bi, Improved lexicon-based sentiment analysis for social media analytics, Secur. Inform. 4 (2015) 1–13, https://doi.org/10.1186/s13388-015-0024-x.
[140] N.N. Yusof, A. Mohamed, S. Abdul-Rahman, Reviewing Classification Approaches in Sentiment Analysis, 2015, pp. 43–53, https://doi.org/10.1007/978-981-287-936-3_5.
[141] L. Oneto, F. Bisio, E. Cambria, D. Anguita, Statistical learning theory and ELM for big social data analysis, IEEE Comput. Intell. Mag. 11 (2016) 45–55, https://doi.org/10.1109/MCI.2016.2572540.
[142] Y. Li, Q. Pan, T. Yang, S. Wang, J. Tang, E. Cambria, Learning word representations for sentiment analysis, Cognit. Comput. 9 (2017) 843–851, https://doi.org/10.1007/s12559-017-9492-2.
[143] A. Hussain, E. Cambria, Semi-supervised learning for big social data analysis, Neurocomputing. 275 (2018) 1662–1673, https://doi.org/10.1016/j.neucom.2017.10.010.
[144] W. Rong, Y. Nie, Y. Ouyang, B. Peng, Z. Xiong, Auto-encoder based bagging architecture for sentiment analysis, J. Vis. Lang. Comput. 25 (2014) 840–849, https://doi.org/10.1016/j.jvlc.2014.09.005.
[145] B. Agarwal, N. Mittal, Prominent Feature Extraction for Sentiment Analysis, 2016, pp. 1–118, https://doi.org/10.1007/978-3-319-25343-5.
[146] G. Yoo, J. Nam, A hybrid approach to sentiment analysis enhanced by sentiment lexicons and polarity shifting devices, in: K. Shirai (Ed.), 13th Work. Asian Lang. Resour, Miyazaki, Japan, 2018, pp. 21–28, https://hal.archives-ouvertes.fr/hal-01795217.
[147] A. Joulin, E. Grave, P. Bojanowski, T. Mikolov, Bag of Tricks for Efficient Text Classification, 2016, pp. 1–5, http://arxiv.org/abs/1607.01759.
[148] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297, https://doi.org/10.1007/BF00994018.
[149] R. Moraes, J.F. Valiati, W.P. Gavião Neto, Document-level sentiment classification: An empirical comparison between SVM and ANN, Expert Syst. Appl. 40 (2013) 621–633, https://doi.org/10.1016/j.eswa.2012.07.059.
[150] T. Joachims, Text categorization with support vector machines: Learning with many relevant features, 1998, pp. 137–142, https://doi.org/10.1007/BFb0026683.
[151] S. Rana, A. Singh, Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques, in: 2016 2nd Int. Conf. Next Gener. Comput. Technol, IEEE, 2016, pp. 106–111, https://doi.org/10.1109/NGCT.2016.7877399.
[152] Y. Al Amrani, M. Lazaar, K.E. El Kadiri, Random forest and support vector machine based hybrid approach to sentiment analysis, Procedia Comput. Sci. 127 (2018) 511–520, https://doi.org/10.1016/j.procs.2018.01.150.
[153] G. Vinodhini, R.M. Chandrasekaran, A comparative performance evaluation of neural network based approach for sentiment classification of online reviews, J. King Saud Univ. - Comput. Inf. Sci. 28 (2016) 2–12, https://doi.org/10.1016/j.jksuci.2014.03.024.
[154] L.-S. Chen, C.-H. Liu, H.-J. Chiu, A neural network based approach for sentiment classification in the blogosphere, J. Informetr. 5 (2011) 313–322, https://doi.org/10.1016/j.joi.2011.01.003.
[155] D. Fisch, E. Kalkowski, B. Sick, Knowledge fusion for probabilistic generative classifiers with data mining applications, IEEE Trans. Knowl. Data Eng. 26 (2014) 652–666, https://doi.org/10.1109/TKDE.2013.20.
[156] E. Shanthi, D. Sangeetha, Analyzing Data Through Data Fusion using Classification Techniques, 2015, pp. 165–173, https://doi.org/10.1007/978-81-322-2208-8_16.
[157] K.M.A. Hasan, M.S. Sabuj, Z. Afrin, Opinion mining using Naïve Bayes, in: 2015 IEEE Int. WIE Conf. Electr. Comput. Eng, IEEE, 2015, pp. 511–514, https://doi.org/10.1109/WIECON-ECE.2015.7443981.
[158] L. Gutiérrez, J. Bekios-Calfa, B. Keith, A Review on Bayesian Networks for Sentiment Analysis, 2019, pp. 111–120, https://doi.org/10.1007/978-3-030-01171-0_10.
[159] Y. Wan, Q. Gao, An ensemble sentiment classification system of Twitter data for airline services analysis, in: 2015 IEEE Int. Conf. Data Min. Work, IEEE, 2015, pp. 1318–1325, https://doi.org/10.1109/ICDMW.2015.7.
[160] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? Sentiment classification using machine learning techniques, in: Proc. 2002 Conf. Empir. Methods Nat. Lang. Process. ({EMNLP} 2002), Association for Computational Linguistics, 2002, pp. 79–86, https://doi.org/10.3115/1118693.1118704.
[161] P. Ficamos, Y. Liu, W. Chen, A Naive Bayes and maximum entropy approach to sentiment analysis: Capturing domain-specific data in Weibo, in: 2017 IEEE Int. Conf. Big Data Smart Comput, IEEE, 2017, pp. 336–339, https://doi.org/10.1109/BIGCOMP.2017.7881689.
[162] A.K.H Tung, Rule-based classification, in: Encycl. Database Syst, Springer, US, Boston, MA, 2009, pp. 2459–2462, https://doi.org/10.1007/978-0-387-39940-9_559.
[163] L.I. Tan, W.S. Phang, K.O. Chin, P. Anthony, Rule-based sentiment analysis for financial news, in: 2015 IEEE Int. Conf. Syst. Man, Cybern., IEEE, 2015, pp. 1601–1606, https://doi.org/10.1109/SMC.2015.283.
[164] K. Gao, H. Xu, J. Wang, A rule-based approach to emotion cause detection for Chinese micro-blogs, Expert Syst. Appl. 42 (2015) 4517–4528, https://doi.org/10.1016/j.eswa.2015.01.064.

[165] J. Han, M. Kamber, J. Pei, 1 - Introduction, in: J. Han, M. Kamber, J. Pei (Eds.), Data Min, Third Ed., Third Edit, Morgan Kaufmann, Boston, 2012, pp. 1–38, https://doi.org/10.1016/B978-0-12-381479-1.00001-0.

[166] R. Nisbet, G. Miner, K. Yale, Classification, in: Handb. Stat. Anal. Data Min. Appl, Elsevier, 2018, pp. 169–186, https://doi.org/10.1016/B978-0-12-416632-5.00009-8.

[167] P.V. Ngoc, C.V.T. Ngoc, T.V.T. Ngoc, D.N. Duy, A C4.5 algorithm for english emotional classification, Evol. Syst. 10 (2019) 425–451, https://doi.org/10.1007/s12530-017-9180-1.

[168] E. Bastı, C. Kuzey, D. Delen, Analyzing initial public offerings' short-term performance using decision trees and SVMs, Decis. Support Syst. 73 (2015) 15–27, https://doi.org/10.1016/j.dss.2015.02.011.

[169] L. Rokach, Ensemble-based classifiers, in: Artif. Intell. Rev, 33, 2010, pp. 1–39, https://doi.org/10.1007/s10462-009-9124-7.

[170] Ankit N. Saleena, An ensemble classification system for Twitter sentiment analysis, Procedia Comput. Sci. 132 (2018) 937–946, https://doi.org/10.1016/j.procs.2018.05.109.

[171] J.J. Bird, A. Ekárt, C.D. Buckingham, D.R. Faria, High Resolution Sentiment Analysis By Ensemble Classification, 2019, pp. 593–606, https://doi.org/10.1007/978-3-030-22871-2_40.

[172] N. Sultana, M.M. Islam, Meta Classifier-Based Ensemble Learning for Sentiment Classification, 2020, pp. 73–84, https://doi.org/10.1007/978-981-13-7564-4_7.

[173] O. Araque, I. Corcuera-Platas, J.F. Sánchez-Rada, C.A. Iglesias, Enhancing deep learning sentiment analysis with ensemble techniques in social applications, Expert Syst. Appl. 77 (2017) 236–246, https://doi.org/10.1016/j.eswa.2017.02.002.

[174] M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, G.S. Choi, GB-SVM: Sentiment classification from unstructured reviews using ensemble classifier, Appl. Sci. 10 (2020) 1–20, https://doi.org/10.3390/app10082788.

[175] M.M. Jawad Soumik, S. Salvi Md Farhavi, F. Eva, T. Sinha, M.S. Alam, Employing machine learning techniques on sentiment analysis of google play store bangla reviews, in: 2019 22nd Int. Conf. Comput. Inf. Technol, IEEE, 2019, pp. 1–5, https://doi.org/10.1109/ICCIT48885.2019.9038348.

[176] F. Alzamzami, M. Hoda, A. El Saddik, Light gradient boosting machine for general sentiment classification on short texts: A comparative evaluation, IEEE Access. 8 (2020) 101840–101858, https://doi.org/10.1109/ACCESS.2020.2997330.

[177] Y. Han, Y. Liu, Z. Jin, Sentiment analysis via semi-supervised learning: a model based on dynamic threshold and multi-classifiers, Neural Comput. Appl. 32 (2020) 5117–5129, https://doi.org/10.1007/s00521-018-3958-3.

[178] C. Lin, Y. He, R. Everson, S. Ruger, Weakly supervised joint sentiment-topic detection from text, IEEE Trans. Knowl. Data Eng. 24 (2012) 1134–1145, https://doi.org/10.1109/TKDE.2011.48.

[179] K.K. Mohbey, B. Bakariya, V. Kalal, A Study and Comparison of Sentiment Analysis Techniques using Demonetization, 2019, pp. 1–14, https://doi.org/10.4018/978-1-5225-4999-4.ch001.

[180] M. Fernández-Gavilanes, T. Álvarez López, J. Juncal-Martínez, E. Costa-Montenegro, F. Javier González-Castaño, Unsupervised method for sentiment analysis in online texts, Expert Syst. Appl. 58 (2016) 57–75, https://doi.org/10.1016/j.eswa.2016.03.031.

[181] B. Ma, H. Yuan, Y. Wu, Exploring performance of clustering methods on document sentiment analysis, J. Inf. Sci. 43 (2017) 54–74, https://doi.org/10.1177/0165551515617374.

[182] S.H. Al-Harbi, V.J. Rayward-Smith, Adapting k-means for supervised clustering, Appl. Intell. 24 (2006) 219–226, https://doi.org/10.1007/s10489-006-8513-8.

[183] H. Suresh, S. Gladston Raj, A Fuzzy Based Hybrid Hierarchical Clustering Model for Twitter Sentiment Analysis, 2017, pp. 384–397, https://doi.org/10.1007/978-981-10-6430-2_30.

[184] K. Tsagkalidou, V. Koutsonikola, A. Vakali, K. Kafetsios, Emotional Aware Clustering on Micro-Blogging Sources, 2011, pp. 387–396, https://doi.org/10.1007/978-3-642-24600-5_42.

[185] D. Archambault, D. Greene, P. Cunningham, TwitterCrowds: Techniques for Exploring Topic and Sentiment in Microblogging Data, 2013, pp. 1–19, http://arxiv.org/abs/1306.3839.

[186] X. Cui, T.E. Potok, P. Palathingal, Document clustering using particle swarm optimization, in: Proc. 2005 IEEE Swarm Intell. Symp. 2005. SIS 2005, IEEE, 2005, pp. 185–191, https://doi.org/10.1109/SIS.2005.1501621.

[187] G. Li, F. Liu, Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions, Appl. Intell. 40 (2014) 441–452, https://doi.org/10.1007/s10489-013-0463-3.

[188] S. Riaz, M. Fatima, M. Kamran, M.W. Nisar, Opinion mining on large scale data using sentiment analysis and k-means clustering, Cluster Comput. 22 (2019) 7149–7164, https://doi.org/10.1007/s10586-017-1077-z.

[189] S. Zhu, B. Xu, D. Zheng, T. Zhao, Chinese Microblog Sentiment Analysis Based on Semi-Supervised Learning, 2013, pp. 325–331, https://doi.org/10.1007/978-1-4614-6880-6_28.

[190] R. Xia, C. Wang, X.-Y. Dai, T. Li, Co-training for semi-supervised sentiment classification based on dual-view bags-of-words representation, in: Proc. 53rd Annu. Meet. Assoc. Comput. Linguist. 7th Int. Jt. Conf. Nat. Lang. Process. vol. 1 Long Pap, Association for Computational Linguistics, Stroudsburg, PA, USA, 2015, pp. 1054–1063, https://doi.org/10.3115/v1/P15-1102.

[191] G. Mesnil, T. Mikolov, M. Ranzato, Y. Bengio, Ensemble of Generative and Discriminative Techniques for Sentiment Analysis of Movie Reviews, 2014, pp. 1–5, http://arxiv.org/abs/1412.5335.

[192] A. Blum, T. Mitchell, Combining labeled and unlabeled data with co-training, in: Proc. Elev. Annu. Conf. Comput. Learn. Theory - COLT' 98, ACM Press, New York, New York, USA, 1998, pp. 92–100, https://doi.org/10.1145/279943.279962.

[193] P. Biyani, C. Caragea, P. Mitra, C. Zhou, J. Yen, G.E. Greer, K. Portier, Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community, in: Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min. - ASONAM '13, ACM Press, New York, New York, USA, 2013, pp. 413–417, https://doi.org/10.1145/2492517.2492606.

[194] V. Iosifidis, E. Ntoutsi, Large scale sentiment learning with limited labels, in: Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min, ACM, New York, NY, USA, 2017, pp. 1823–1832, https://doi.org/10.1145/3097983.3098159.

[195] W. Gao, S. Li, Y. Xue, M. Wang, G. Zhou, Semi-Supervised Sentiment Classification with Self-Training on Feature Subspaces, 2014, pp. 231–239, https://doi.org/10.1007/978-3-319-14331-6_23.

[196] V. Van Asch, W. Daelemans, Predicting the Effectiveness of Self-Training: Application To Sentiment Classification, 2016, pp. 1–9, http://arxiv.org/abs/1601.03288.

[197] S. Hong, J. Lee, J.-H. Lee, Competitive self-training technique for sentiment analysis in mass social media, in: 2014 Jt. 7th Int. Conf. Soft Comput. Intell. Syst. 15th Int. Symp. Adv. Intell. Syst, IEEE, 2014, pp. 9–12, https://doi.org/10.1109/SCIS-ISIS.2014.7044857.

[198] M.S. Hajmohammadi, R. Ibrahim, A. Selamat, H. Fujita, Combination of active learning and self-training for cross-lingual sentiment classification with density analysis of unlabelled samples, Inf. Sci. (Ny). 317 (2015) 67–77, https://doi.org/10.1016/j.ins.2015.04.003.

[199] Y. He, D. Zhou, Self-training from labeled features for sentiment analysis, Inf. Process. Manag. 47 (2011) 606–616, https://doi.org/10.1016/j.ipm.2010.11.003.

[200] M.S. Hajmohammadi, R. Ibrahim, A. Selamat, Graph-Based Semi-Supervised Learning for Cross-Lingual Sentiment Classification, 2015, pp. 97–106, https://doi.org/10.1007/978-3-319-15702-3_10.

[201] T.-J. Lu, Semi-supervised microblog sentiment analysis using social relation and text similarity, in: 2015 Int. Conf. Big Data Smart Comput, IEEE, 2015, pp. 194–201, https://doi.org/10.1109/35021BIGCOMP.2015.7072831.

[202] Y. Ren, N. Kaji, N. Yoshinaga, M. Kitsuregawa, Sentiment classification in under-resourced languages using graph-based semi-supervised learning methods, IEICE Trans. Inf. Syst. E97.D (2014) 790–797, https://doi.org/10.1587/transinf.E97.D.790.

[203] A. Jalilvand, N. Salim, Sentiment Classification using Graph Based Word Sense Disambigution, 2012, pp. 351–358, https://doi.org/10.1007/978-3-642-35326-0_35.

[204] Y. Su, S. Li, S. Ju, G. Zhou, X. Li, Multi-view learning for semi-supervised sentiment classification, in: 2012 Int. Conf. Asian Lang. Process, IEEE, 2012, pp. 13–16, https://doi.org/10.1109/IALP.2012.53.

[205] G. Lazarova, I. Koychev, Semi-Supervised Multi-View Sentiment Analysis, 2015, pp. 181–190, https://doi.org/10.1007/978-3-319-24069-5_17.

[206] Y. Li, Y. Fang, Z. Akhtar, Accelerating deep reinforcement learning model for game strategy, Neurocomputing. (2020) 1–12, https://doi.org/10.1016/j.neucom.2019.06.110.

[207] W. Liu, L. Zhang, D. Tao, J. Cheng, Reinforcement online learning for emotion prediction by using physiological signals, Pattern Recognit. Lett. 107 (2018) 123–130, https://doi.org/10.1016/j.patrec.2017.06.004.

[208] J. Broekens, E. Jacobs, C.M. Jonker, A reinforcement learning model of joy, distress, hope and fear, Conn. Sci. 27 (2015) 215–233, https://doi.org/10.1080/09540091.2015.1031081.

[209] L.M. Rojas-Barahona, Deep learning for sentiment analysis, Lang. Linguist. Compass. 10 (2016) 701–719, https://doi.org/10.1111/lnc3.12228.

[210] P. Vateekul, T. Koomsubha, A study of sentiment analysis using deep learning techniques on Thai Twitter data, in: 2016 13th Int. Jt. Conf. Comput. Sci. Softw. Eng., IEEE, 2016, pp. 1–6, https://doi.org/10.1109/JCSSE.2016.7748849.

[211] Y. Kim, Convolutional neural networks for sentence classification, in: Proc. 2014 Conf. Empir. Methods Nat. Lang. Process, Association for Computational Linguistics, Stroudsburg, PA, USA, 2014, pp. 1746–1751, https://doi.org/10.3115/v1/D14-1181.

[212] W. Li, F. Qi, M. Tang, Z. Yu, Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification, Neurocomputing. 387 (2020) 63–77, https://doi.org/10.1016/j.neucom.2020.01.006.

[213] S. Zhou, Q. Chen, X. Wang, Fuzzy deep belief networks for semi-supervised sentiment classification, Neurocomputing. 131 (2014) 312–322, https://doi.org/10.1016/j.neucom.2013.10.011.

[214] H. Shirani-mehr, Applications of Deep Learning To Sentiment Analysis of Movie Reviews, 2015, pp. 1–8, https://cs224d.stanford.edu/reports/Shirani-MehrH.pdf.

[215] N.C. Dang, M.N. Moreno-García, F. De la Prieta, Sentiment analysis based on deep learning: A comparative study, Electronics. 9 (2020) 1–29, https://doi.org/10.3390/electronics9030483.

[216] S. Sohangir, D. Wang, A. Pomeranets, T.M. Khoshgoftaar, Big data: Deep learning for financial sentiment analysis, J. Big Data. 5 (2018) 1–28, https://doi.org/10.1186/s40537-017-0111-6.

[217] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks. 61 (2015) 85–117, https://doi.org/10.1016/j.neunet.2014.09.003.

[218] A. Vassilev, Bowtie – A deep learning feedforward neural network for sentiment analysis, 2019, pp. 1–13, https://doi.org/10.6028/NIST.CSWP.04222019.

[219] X. Ouyang, P. Zhou, C.H. Li, L. Liu, Sentiment analysis using convolutional neural network, in: 2015 IEEE Int. Conf. Comput. Inf. Technol. Ubiquitous Comput. Commun. Dependable, Auton. Secur. Comput. Pervasive Intell. Comput, IEEE, 2015, pp. 2359–2364, https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.349.

[220] A. Aziz Sharfuddin, M. Nafis Tihami, M. Saiful Islam, A deep recurrent neural network with bilstm model for sentiment classification, in: 2018 Int. Conf. Bangla Speech Lang. Process, IEEE, 2018, pp. 1–4, https://doi.org/10.1109/ICBSLP.2018.8554396.

[221] C. Chen, R. Zhuo, J. Ren, Gated recurrent neural network with sentimental relations for sentiment classification, Inf. Sci. (Ny). 502 (2019) 268–278, https://doi.org/10.1016/j.ins.2019.06.050.

[222] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.

[223] C. Li, B. Xu, G. Wu, S. He, G. Tian, H. Hao, Recursive deep learning for sentiment analysis over social data, in: 2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol, IEEE, 2014, pp. 180–185, https://doi.org/10.1109/WI-IAT.2014.96.

[224] H. Sagha, N. Cummins, B. Schuller, Stacked denoising autoencoders for sentiment analysis: a review, Wiley Interdiscip. Rev. Data Min. Knowl. 7 (2017) 1–9, https://doi.org/10.1002/widm.1212.

[225] H. Zhou, L. Chen, D. Huang, Cross-Lingual Sentiment Classification Based on Denoising Autoencoder, 2014, pp. 181–192, https://doi.org/10.1007/978-3-662-45924-9_17.

[226] A.U. Rehman, A.K. Malik, B. Raza, W. Ali, A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis, Multimedia Tools Appl. 78 (2019) 26597–26613, https://doi.org/10.1007/s11042-019-07788-7.

[227] M. Hu, B. Liu, Mining and summarizing customer reviews, in: Proc. 2004 ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD '04, ACM Press, New York, New York, USA, 2004, pp. 168–177, https://doi.org/10.1145/1014052.1014073.

[228] S. Sanagar, D. Gupta, Automated genre-based multi-domain sentiment lexicon adaptation using unlabeled data, J. Intell. Fuzzy Syst. 38 (2020) 6223–6234, https://doi.org/10.3233/JIFS-179704.

[229] S. Sanagar, D. Gupta, Unsupervised genre-based multidomain sentiment lexicon learning using corpus-generated polarity seed words, IEEE Access. 8 (2020) 118050–118071, https://doi.org/10.1109/ACCESS.2020.3005242.

[230] M.Z. Asghar, A. Sattar, A. Khan, A. Ali, F. Masud Kundi, S. Ahmad, Creating sentiment lexicon for sentiment analysis in Urdu: The case of a resource-poor language, Expert Syst. 36 (2019) 1–19, https://doi.org/10.1111/exsy.12397.

[231] T. Wilson, J. Wiebe, P. Hoffmann, Recognizing contextual polarity in phrase-level sentiment analysis, in: Proc. Conf. Hum. Lang. Technol. Empir. Methods Nat. Lang. Process. - HLT '05, Association for Computational Linguistics, Morristown, NJ, USA, 2005, pp. 347–354, https://doi.org/10.3115/1220575.1220619.

[232] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-based methods for sentiment analysis, Comput. Linguist. 37 (2011) 267–307, https://doi.org/10.1162/COLI_a_00049.

[233] S.M. Mohammad, P.D. Turney, Crowdsourcing a Word-Emotion Association Lexicon, 2013, pp. 1–25, http://arxiv.org/abs/1308.6297.

[234] Y. Hong, H. Kwak, Y. Baek, S. Moon, Tower of babel, in: Proc. 22nd Int. Conf. World Wide Web - WWW '13 Companion, ACM Press, New York, New York, USA, 2013, pp. 549–556, https://doi.org/10.1145/2487788.2487993.

[235] G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross, K.J. Miller, Introduction to WordNet: An on-line lexical database *, Int. J. Lexicogr. 3 (1990) 235–244, https://doi.org/10.1093/ijl/3.4.235.

[236] S. Mohammad, C. Dunne, B. Dorr, Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus, in: Proc. 2009 Conf. Empir. Methods Nat. Lang. Process. vol. 2-2, Association for Computational Linguistics, USA, 2009, pp. 599–608.

[237] S. Park, Y. Kim, Building thesaurus lexicon using dictionary-based approach for sentiment classification, in: 2016 IEEE 14th Int. Conf. Softw. Eng. Res. Manag. Appl, IEEE, 2016, pp. 39–44, https://doi.org/10.1109/SERA.2016.7516126.

[238] S. Sanagar, D. Gupta, Roadmap for Polarity Lexicon Learning and Resources: A Survey, 2016, pp. 647–663, https://doi.org/10.1007/978-3-319-47952-1_52.

[239] A. Esuli, F. Sebastiani, SENTIWORDNET: A Publicly available lexical resource for opinion mining, in: Proc. Fifth Int. Conf. Lang. Resour. Eval, European Language Resources Association (ELRA), Genoa, Italy, 2006, pp. 1–6, http://www.lrec-conf.org/proceedings/lrec2006/pdf/384_pdf.pdf.

[240] E. Cambria, Y. Li, F. Xing, S. Poria, K. Kwok, SenticNet 6: Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis, in: Proc. 29th ACM Int. Conf. Inf. Knowl. Manag., 2020, pp. 105–114.

[241] D. Sarkar, R. Bali, T. Sharma, Practical Machine Learning with Python, 2018, pp. 1–545, https://doi.org/10.1007/978-1-4842-3207-1.

[242] B. Liu, Opinion mining and sentiment analysis, in: Web Data Min, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 459–526, https://doi.org/10.1007/978-3-642-19460-3_11.

[243] T. Luo, S. Chen, G. Xu, J. Zhou, Sentiment analysis, in: Trust. Collect. View Predict, Springer New York, New York, NY, 2013, pp. 53–68, https://doi.org/10.1007/978-1-4614-7202-5_4.

[244] V. Hatzivassiloglou, K.R. McKeown, Predicting the semantic orientation of adjectives, in: 35th Annu. Meet. Assoc. Comput. Linguist. 8th Conf. {E}Uropean Chapter Assoc. Comput. Linguist, Association for Computational Linguistics, Madrid, Spain, 1997, pp. 174–181, https://doi.org/10.3115/976909.979640.

[245] B. Agarwal, N. Mittal, Semantic Orientation-Based Approach for Sentiment Analysis, 2016, pp. 77–88, https://doi.org/10.1007/978-3-319-25343-5_6.

[246] V. Vyas, V. Uma, Approaches To Sentiment Analysis on Product Reviews, 2019, pp. 15–30, https://doi.org/10.4018/978-1-5225-4999-4.ch002.

[247] P.D. Turney, M.L. Littman, Measuring Praise and Criticism: Inference of Semantic Orientation from Association, 2003, pp. 1–37.

[248] H. Han, J. Zhang, J. Yang, Y. Shen, Y. Zhang, Generate domain-specific sentiment lexicon for review sentiment analysis, Multimedia Tools Appl. 77 (2018) 21265–21280, https://doi.org/10.1007/s11042-017-5529-5.

[249] O. Araque, G. Zhu, C.A. Iglesias, A semantic similarity-based perspective of affect lexicons for sentiment analysis, Knowledge-Based Syst. 165 (2019) 346–359, https://doi.org/10.1016/j.knosys.2018.12.005.

[250] W. Zhang, H. Xu, W. Wan, Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis, Expert Syst. Appl. 39 (2012) 10283–10291, https://doi.org/10.1016/j.eswa.2012.02.166.

[251] Z. Dong, Q. Dong, Hownet and the Computation of Meaning, 2006, pp. 1–316, https://doi.org/10.1142/5935.

[252] I. C, N. Joshi, Enhanced Twitter sentiment analysis using hybrid approach and by accounting local contextual semantic, J. Intell. Syst. 29 (2019) 1611–1625, https://doi.org/10.1515/jisys-2019-0106.

[253] D.V.N. Devi, T. Venkata Rajini Kanth, K. Mounika, N. Sowjanya Swathi, Assay: Hybrid Approach for Sentiment Analysis, 2019, pp. 309–318, https://doi.org/10.1007/978-981-13-1742-2_30.

[254] B. Shin, T. Lee, J.D. Choi, Lexicon Integrated CNN Models with Attention for Sentiment Analysis, 2016, pp. 1–10, http://arxiv.org/abs/1610.06272.

[255] K. Elshakankery, M.F. Ahmed, HILATSA: A hybrid incremental learning approach for arabic tweets sentiment analysis, Egypt. Informatics J. (2019) 163–171, https://doi.org/10.1016/j.eij.2019.03.002.

[256] X. Tan, Y. Cai, J. Xu, H.-F. Leung, W. Chen, Q. Li, Improving aspect-based sentiment analysis via aligning aspect embedding, Neurocomputing. 383 (2020) 336–347, https://doi.org/10.1016/j.neucom.2019.12.035.

[257] Q. Xu, L. Zhu, T. Dai, C. Yan, Aspect-based sentiment classification with multi-attention network, Neurocomputing. 388 (2020) 135–143, https://doi.org/10.1016/j.neucom.2020.01.024.

[258] P. Karagoz, B. Kama, M. Ozturk, I.H. Toroslu, D. Canturk, A framework for aspect based sentiment analysis on turkish informal texts, J. Intell. Inf. Syst. 53 (2019) 431–451, https://doi.org/10.1007/s10844-019-00565-w.

[259] B. Kama, M. Ozturk, P. Karagoz, I.H. Toroslu, O. Ozay, A Web Search Enhanced Feature Extraction Method for Aspect-Based Sentiment Analysis for Turkish Informal Texts, 2016, pp. 225–238, https://doi.org/10.1007/978-3-319-43946-4_15.

[260] J. Meng, Y. Long, Y. Yu, D. Zhao, S. Liu, Cross-domain text sentiment analysis based on CNN_FT method, Information. 10 (2019) 1–13, https://doi.org/10.3390/info10050162.

[261] R. Bartusiak, L. Augustyniak, T. Kajdanowicz, P. Kazienko, Sentiment analysis for polish using transfer learning approach, in: 2015 Second Eur. Netw. Intell. Conf, IEEE, 2015, pp. 53–59, https://doi.org/10.1109/ENIC.2015.16.

[262] E. Cambria, B. Schuller, Y. Xia, C. Havasi, New avenues in opinion mining and sentiment analysis, IEEE Intell. Syst. 28 (2013) 15–21, https://doi.org/10.1109/MIS.2013.30.
[263] I. Chaturvedi, R. Satapathy, S. Cavallari, E. Cambria, Fuzzy commonsense reasoning for multimodal sentiment analysis, Pattern Recognit. Lett. 125 (2019) 264–270, https://doi.org/10.1016/j.patrec.2019.04.024.
[264] J.A. Balazs, J.D. Velásquez, Opinion mining and information fusion: A survey, Inf. Fusion. 27 (2016) 95–110, https://doi.org/10.1016/j.inffus.2015.06.002.
[265] Z.-P. Fan, G.-M. Li, Y. Liu, Processes and methods of information fusion for ranking products based on online reviews: An overview, Inf. Fusion. 60 (2020) 87–97, https://doi.org/10.1016/j.inffus.2020.02.007.
[266] N. Majumder, D. Hazarika, A. Gelbukh, E. Cambria, S. Poria, Multimodal sentiment analysis using hierarchical fusion with context modeling, Knowledge-Based Syst. 161 (2018) 124–133, https://doi.org/10.1016/j.knosys.2018.07.041.
[267] Z. Zhao, H. Zhu, Z. Xue, Z. Liu, J. Tian, M.C.H. Chua, M. Liu, An image-text consistency driven multimodal sentiment analysis approach for social media, Inf. Process. Manag. 56 (2019) 1–9, https://doi.org/10.1016/j.ipm.2019.102097.
[268] D. Ayata, Y. Yaslan, M.E. Kamasak, Emotion recognition from multimodal physiological signals for emotion aware healthcare systems, J. Med. Biol. Eng. 40 (2020) 149–157, https://doi.org/10.1007/s40846-019-00505-7.
[269] S. Poria, E. Cambria, N. Howard, G.-B. Huang, A. Hussain, Fusing audio visual and textual clues for sentiment analysis from multimodal content, Neurocomputing. 174 (2016) 50–59, https://doi.org/10.1016/j.neucom.2015.01.095.
[270] D. Jiang, X. Luo, J. Xuan, Z. Xu, Sentiment computing for the news event based on the social media big data, IEEE Access. 5 (2017) 2373–2382, https://doi.org/10.1109/ACCESS.2016.2607218.
[271] S. Ahmed, A. Danti, Effective Sentimental Analysis and Opinion Mining of Web Reviews using Rule Based Classifiers, 2016, pp. 171–179, https://doi.org/10.1007/978-81-322-2734-2_18.
[272] M.S. Akhtar, A. Ekbal, E. Cambria, How intense are you? Predicting intensities of emotions and sentiments using stacked ensemble [Application Notes], IEEE Comput. Intell. Mag. 15 (2020) 64–75, https://doi.org/10.1109/MCI.2019.2954667.
[273] M. Rundell, Macmillan English Dictionary for Advanced Learners, 2007, pp. 1–1748, https://books.google.co.ma/books?id=v6SISQAACAAJ.
[274] M. Birjali, A. Beni-Hssane, M. Erritali, A method proposed for estimating depressed feeling tendencies of social media users utilizing their data, in: Adv. Intell. Syst. Comput., 2017, pp. 1–8, https://doi.org/10.1007/978-3-319-52941-7_41.
[275] L. Ren, B. Xu, H. Lin, X. Liu, L. Yang, Sarcasm detection with sentiment semantics enhanced multi-level memory network, Neurocomputing. (2020) 1–7, https://doi.org/10.1016/j.neucom.2020.03.081.
[276] Y. Ren, D. Ji, H. Ren, Context-augmented convolutional neural networks for twitter sarcasm detection, Neurocomputing. 308 (2018) 1–7, https://doi.org/10.1016/j.neucom.2018.03.047.
[277] D. Jain, A. Kumar, G. Garg, Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN, Appl. Soft Comput. 91 (2020) 1–11, https://doi.org/10.1016/j.asoc.2020.106198.
[278] L. Lazib, B. Qin, Y. Zhao, W. Zhang, T. Liu, A syntactic path-based hybrid neural network for negation scope detection, Front. Comput. Sci. 14 (2020) 84–94, https://doi.org/10.1007/s11704-018-7368-6.
[279] E.F. Cardoso, R.M. Silva, T.A. Almeida, Towards automatic filtering of fake reviews, Neurocomputing 309 (2018) 106–116, https://doi.org/10.1016/j.neucom.2018.04.074.
[280] S. Saumya, J.P. Singh, Detection of spam reviews: a sentiment analysis approach, CSI Trans. ICT. 6 (2018) 137–148, https://doi.org/10.1007/s40012-018-0193-0.
[281] I. Toledo-Gómez, E. Valtierra-Romero, A. Guzmán-Arenas, A. Cuevas-Rasgado, L. Méndez-Segundo, AnaPro, tool for identification and resolution of direct anaphora in Spanish, J. Appl. Res. Technol. 12 (2014) 14–40, https://doi.org/10.1016/S1665-6423(14)71602-5.
[282] R. Sukthanker, S. Poria, E. Cambria, R. Thirunavukarasu, Anaphora and coreference resolution: A review, Inf. Fusion. 59 (2020) 139–162, https://doi.org/10.1016/j.inffus.2020.01.010.
[283] L.J. Deborah, V. Karthika, R. Baskaran, A. Kannan, Enhanced Anaphora Resolution Algorithm Facilitating Ontology Construction, 2011, pp. 526–535, https://doi.org/10.1007/978-3-642-22555-0_54.
[284] Y. Wang, M. Wang, H. Fujita, Word sense disambiguation: A comprehensive knowledge exploitation framework, Knowledge-Based Syst. 190 (2020) 1–13, https://doi.org/10.1016/j.knosys.2019.105030.
[285] Y. Xia, E. Cambria, A. Hussain, H. Zhao, Word polarity disambiguation using Bayesian model and opinion-level features, Cognit. Comput. 7 (2015) 369–380, https://doi.org/10.1007/s12559-014-9298-4.
[286] C. Myers-Scotton, Common and uncommon ground: Social and structural factors in codeswitching, Lang. Soc. 22 (1993) 475–503, https://doi.org/10.1017/S0047404500017449.
[287] S. Poria, D. Hazarika, N. Majumder, R. Mihalcea, Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research, IEEE Trans. Affect. Comput. (2020) 1, s.
[288] A. Chatterjere, V. Guptha, P. Chopra, A. Das, Minority positive sampling for switching points - an anecdote for the code-mixing language modeling, in: Proc. 12th Lang. Resour. Eval. Conf, European Language Resources Association, Marseille, France, 2020, pp. 6228–6236.
[289] Y.K. Lal, V. Kumar, M. Dhar, M. Shrivastava, P. Koehn, De-mixing sentiment from code-mixed text, in: Proc. 57th Annu. Meet. Assoc. Comput. Linguist. Student Res. Work, Association for Computational Linguistics, Stroudsburg, PA, USA, 2019, pp. 371–377, https://doi.org/10.18653/v1/P19-2052.