



A survey on sentiment analysis methods, applications, and challenges

Mayur Wankhade^{1,2} · Annavarapu Chandra Sekhara Rao^{1,2}  · Chaitanya Kulkarni^{1,2}

Published online: 7 February 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

The rapid growth of Internet-based applications, such as social media platforms and blogs, has resulted in comments and reviews concerning day-to-day activities. Sentiment analysis is the process of gathering and analyzing people's opinions, thoughts, and impressions regarding various topics, products, subjects, and services. People's opinions can be beneficial to corporations, governments, and individuals for collecting information and making decisions based on opinion. However, the sentiment analysis and evaluation procedure face numerous challenges. These challenges create impediments to accurately interpreting sentiments and determining the appropriate sentiment polarity. Sentiment analysis identifies and extracts subjective information from the text using natural language processing and text mining. This article discusses a complete overview of the method for completing this task as well as the applications of sentiment analysis. Then, it evaluates, compares, and investigates the approaches used to gain a comprehensive understanding of their advantages and disadvantages. Finally, the challenges of sentiment analysis are examined in order to define future directions.

Keywords Sentiment analysis · Text analysis · Word embedding · Machine learning · Social media

1 Introduction

Sentiment analysis has gained widespread acceptance in recent years, not just among researchers but also among businesses, governments, and organizations (Sánchez-Rada and Iglesias 2019). The growing popularity of the Internet has lifted the web to the rank of the principal source of universal information. Lots of users use various online resources to express their views and opinions. To constantly monitor public opinion and

✉ Annavarapu Chandra Sekhara Rao
acsrao@iitism.ac.in

Mayur Wankhade
mayur.18dr0078@cse.iitism.ac.in

¹ Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad 826004, India

² Dayananda Sagar College of Engineering, Bangalore 560078, India

aid decision-making, we must employ user-generated data to analyze it automatically. As a result, sentiment analysis has increased its popularity across research communities in recent years. Sentiment analysis is also called as Opinion analysis or Opinion mining. We have seen a recent growth in the sentiment analysis task. The various research works in sentiment analysis (Lighthart et al. 2021) published an overview on Opinion mining in the earlier stage. In (Piryani et al. 2017) discusses the study topic from 2000 to 2015 and provides a framework for computationally processing unstructured data with the primary goal of extracting views and identifying their moods. Several recent surveys (Yousif et al. 2019; Birjali et al. 2021) authors has described the problem of sentiment analysis and suggested potential directions. Soleymani et al. (2017) and Yadav and Vishwakarma (2020) on sentiment classification have been published. Also the topic of detecting opinion spam and fraudulent reviews was investigated. Additionally, In the work of Yue et al. (2019) and Liu et al. (2012) conducted research on the effectiveness of internet reviews. The Authors (Jain et al. 2021b) discuss machine learning applications that incorporate online reviews in sentiment categorization, predictive decision-making, and the detection of false reviews. In the work of Balaji et al. (2021) conducted a thorough examination of the several applications of social media analysis utilizing sophisticated machine learning algorithms. Authors present a brief overview of machine learning algorithms used in social media analysis (Hangya and Farkas 2017).

The growth of social network sites has generated a slew of fields devoted to analyzing these networks and their contents in order to extract necessary information. Sentiment analysis is concerned with deriving the sentiments communicated by a piece of text from its content. Sentiment analysis is a subfield of NLP and that, given long and illustrious public opinion for decision making, there must be multiple early works addressing it. However, it still works going on sentiment analysis develop till the new millennium.

Several real-world applications require sentiment analysis for detailed investigation. For example, product analysis, discover which components or qualities of a product appeal to customers in terms of product quality. In work of Subhashini et al. (2021) presents the results of a comprehensive review of contemporary opinion mining literature. It also covers how to extract text features from opinions with noise or uncertainty, represent knowledge in opinions, and categorize them. Mowlaei et al. (2020) suggest a technique for adaptive aspect-based lexicons for sentiment classification. The authors described two strategies for constructing two dynamic lexicons to aid in the classification of sentiments depending on their aspects: a strategy based on statistics and genetic algorithms. A dynamic lexicon can be automatically updated and provides more precise grading for context-related concepts (Kumar and Uma 2021). To organize all the aspects of the reviews they selected a number of lexicons from several dictionaries. Sentiment analysis is previously being applied in various domains ranging from hotels to airlines and healthcare to the stock market (Zvarevashe and Olugbara 2018). Sentiment analysis has applied to hotel reviews to get a better understanding of customer's likes and dislikes. In comparison (Valencia et al. 2019), it is used to determine the trends of the stock market and cryptocurrencies based on the market sentiment. The Authors (Ahmad et al. 2019a) analyze tweets relating to various domains and analyze the sentiments of the tweets. The health care domain is seeing a surge in sentiment analysis applications in recent times, customer opinion analyses (Ruffer et al. 2020; Park et al. 2020; Cortis and Davis 2021; Arora et al. 2021), customer satisfaction analyses are few applications in the healthcare sector (Baashar et al. 2020; Miotto et al. 2018). The business sector has always utilized sentiment analysis for its improvement. Sentiment analysis

for various applications like reputation management, market research, and competitor analysis, product analysis, customer voice, etc.

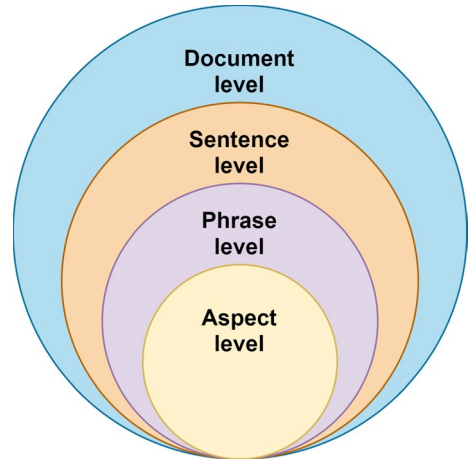
Various issues are associated with sentiment analysis and natural language processing, such as individuals informal writing style, sarcasm, irony, and language-specific challenges. There are many words in different languages whose meaning and orientation change depending on the context and domain in which they are employed. Therefore, there are not many tools and resources available for all the languages. Sarcasm and irony are two of the most critical challenges that have recently attracted the attention of researchers. There has been much development in detecting sarcasm and irony in text. There are many challenges in sentiment analysis. In this work we will analyze the various challenges, methodologies, applications, and algorithms that are employed in sentiment analysis. We present the task with comparative data analysis shown in tables, flow charts, and graphs that are simple to comprehend. Various description of abbreviation which use in our work shown in Table 1.

To our understanding, existing surveys frequently skip some of the sentiment analysis techniques in favor of machine learning, transformer learning, and lexicon-based approaches. Although this paper covered all the task as well, it varies from earlier research in that it covers the most frequently used techniques. Additionally, other surveys study sentiment analysis from a particular task, various challenges, or concentrate on a specific issue, such as product reviews. This study provides a comprehensive investigation of sentiment analysis by discussing this area from various perspectives since it encompasses numerous research components connected to sentiment analysis, such as problems, applications, tools, and approaches. This work is highly beneficial to scholars and beginners, as it allows them to access a wealth of knowledge about this area in a single paper. The survey important contributions can be summarized as follows

- Several literature have been analyzed in order to thoroughly define the sentiment analysis process and to identify well-known technologies for performing this work.
- Analyses of available methodologies in order to determine which one is most appropriate for a certain application.
- We classify and summarize frequently used sentiment analysis approaches to understand better accessible techniques such as machine learning, lexicon-based analysis, and hybrid analysis.
- Summarizing the benefits and challenges of sentiment analysis in order to keep up of current trending research.
- Each method comparison with their advantage and disadvantage, suggesting selecting the proper method sentiment analysis task.

The literature survey paper is organized as Sect. 2, Level of Sentiment Analysis, Sect. 3, contain the Data Collection, Feature Extraction, and Feature Selection Method, explaining all the steps from data extraction to various task of Sentiment Analysis, Sect. 4 contain General Methodology for Sentiment Analysis and its Summary, Sect. 5, Contain the Sentiment Analysis Application in Various Domain, Sect. 6, Contain the Challenges in Sentiment Analysis, In the final Sect. 7, We Conclude our research work.

Fig. 1 Level of sentiment analysis



2 Sentiment analysis levels

Sentiment analysis has been investigated on several levels: Document Level, Sentence Level, Phrase Level, and Aspect Level. Sentiment analysis in each level such as document, sentence and phrase, aspect level shown in Fig. 1.

2.1 Document level sentiment analysis

Document-level: Document level sentiment analysis is performed on a whole document, and single polarity is given to the whole document. This type of sentiment analysis is not used a lot. It can be used to classify chapters or pages of a book as positive, negative, or neutral. At this level, both supervised and unsupervised learning approaches can be utilized to classify the document (Bhatia et al. 2015). Cross-domain and cross-language sentiment analysis are the two most significant issues in document-level sentiment analysis. (Saunders 2021) Domain-specific sentiment analysis has been shown to achieve remarkable accuracy while staying highly domain-sensitive. In these tasks, the feature vector is a set of words that must be domain-specific and limited.

2.2 Sentence level sentiment analysis

Sentence level: In this level of analysis, each sentence is analyzed and finding with a corresponding polarity. This is highly useful when a document has a wide range and mix of sentiments associated with it (Yang and Cardie 2014). This classification level is associated with subjective classification (Rao et al. 2018). Each sentence polarity will be determined independently using the same methodologies as the document level but with greater training data and processing resources. The polarity of each sentence may be aggregated to find the sentiment of the document or used individually. Occasionally, document-level sentiment analysis is insufficient for specific uses (Behdenna et al. 2018). In previous work on sentence-level analysis has been devoted to finding subjective sentences. However, more difficult tasks, such

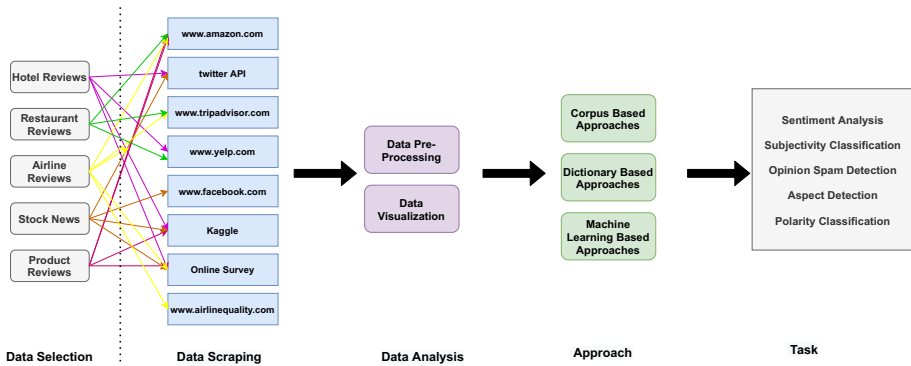


Fig. 2 General procedure of sentiment analysis

as working with conditional sentences or ambiguous statements (Ferrari and Esuli 2019). In these circumstances, sentence-level sentiment analysis is critical.

2.3 Phrase level sentiment analysis

Phrase level: Sentiment analysis also be performed where opinion words are mined at phrase level, and classification will be done. Each phrase may contain multiple aspects or single aspects. This may be useful product reviews of multiple lines; here, it is observed that a single aspect is expressed in a phrase (Thet et al. 2010). It has been a hot topic of researchers in recent times. While document-level analysis concentrated on categorizing the entire document as subjective, either positively or negatively, sentence-level analysis is more beneficial, as a document contains both positive and negative statements. Word is the most basic unit of language; its polarity is intimately related to the subjectivity of the sentence or document in which it appears. A sentence containing an adjective has a high probability of being a subjective sentence. (Fredriksen-Goldsen and Kim 2017) Additionally, the term chosen for expression represents the demographic characteristics of individuals, such as gender and age, and its desire, social standing, and personality, other psychological and social characteristics (Flek 2020). As a result, term serves as the foundation for text sentiment analysis.

2.4 Aspect level sentiment analysis

Aspect level: sentiment analysis is performed at the aspect level. Each sentence may contain multiple aspects; therefore, Aspect level sentiment analysis. Primary attention to all the aspects used in the sentence and assigns polarity to all the aspects after which an aggregate sentiment has calculated for the whole sentence (Schouten and Frasinca 2015; Lu et al. 2011).

Table 1 Various description of abbreviation

Abbreviation	Description of abbreviation
NLP	Natural Language Processing
NB	Naïve Bayes
SVM	Support Vector Machine
RF	Random Forest
LR	Logistic Regression
DT	Decision Tree
LSTM	Long Short-Term Memory
Bi-LSTM	Bi-directional Long Short-Term Memory
CNN	Convolutional Neural Network
RNTN	Recursive Neural Tensor Network
RNN	Recurrent Neural Network
ANN	Artificial Neural Network
RNTN	Recursive Neural Tensor Network
AI	Artificial Intelligence
KNN	K nearest Neighbours
LDA	Latent Dirichlet allocation
ME	Maximum Entropy
LSA	Latent Semantic Analysis
CBOW	Continuous Bag of words
SG	Skip Gram
TF-IDF	Term Frequency-Inverse Document Frequency
DNN	Deep Neural Network
BoW	Bag of Words
ABSA	Aspect Based Sentiment Analysis
PoS	Parts-of-Speech
NLTK	Natural Language tool kit
DCNN	Dynamic Convolutional Neural Network
BERT	Bidirectional Encoder Representations from Transformers
MSA	Multimodal Sentiment Analysis
GPT	Generative Pre-trained Transformer
SLC	Sentence Level Classification
AR-NN	Auto-regressive Neural Network
DMN	Dynamic Memory Networks
ELMo	Embedding Language Models

3 Data collection and feature selection

3.1 Data collections

Data can be collected from the internet via web scraping, social media, news channels, E-commerce websites, Forums, Weblog, some other websites shown in Fig. 2. Data Collection is the first stage in the Sentiment Analysis. Depending on task sentiment analysis

of findings, text data can be combined with other types of data like video, audio, location, etc. A few essential sources of data collection are:

Social media: Social data refers to information gathered via social media networks. It demonstrates how consumers interact with the product by accessing, posting, and exchanging over. Academic study on individual, group and behavior uses social media as a dynamic data source. It refers to Internet apps that are web or mobile-based that enable users to create, access, and trade user-generated content.

Forums: Users can use message boards to discuss various topics, exchange opinions and ideas, and solicit assistance via text messages. Forums are an intriguing source for sentiment analysis due to the dynamic nature of user-generated information. Additionally, researchers can undertake sentiment analysis on a specific domain by leveraging forums as a source (Korkontzelos et al. 2016).

Weblog: A short weblog consists of paragraphs conveying a viewpoint, facts, personal diary entries, or links. Together referred to as posts, that are chronologically sorted with the most recent entry appearing first, in the style of a research article (Kumar and Teeja 2012). Blogs are a valuable resource for performing sentiment analysis on a variety of entities (Annett and Kondrak 2008).

Electronic Commerce website: Electronic Commerce websites where users can give evaluations and express their opinions about a particular business or organization. In this instance, websites that do not specifically review sites have millions of reviews, such as e-commerce sites that feature product reviews¹ or professional review sites such as.² In the work of Jain et al. (2019) conducted a descriptive study of the various airline service classifications.

3.2 Feature selection

It is important to remember that developing a classification model requires first identifying relevant features in dataset (Ritter et al. 2012). Thus, a review can be decoded into words during model training and appended to the feature vector. For single word is considered, the technique is called a “Uni-gram”; when two words are considered, the technique is called a “Bi-gram”; and if three words are considered, the technique is referred to as a “Tri-gram.” combination of unigram and bigram helpful for analysis (Razon and Barnden 2015); the context feature which helpful for getting results most accurate.

Pragmatic features are those that emphasize the application of words rather than a methodological foundation. Pragmatics is the study of how context relates to perception in linguistics and related sciences. Pragmatics is the study of phenomena such as implicature, speech acts, relevance, and conversations.

Emoji are facial expressions used in sentiment analysis to convey emotions. Various emoticons are used to depict a wide variety of human emotions (Tian et al. 2017). Emoticons aid in conveying a person’s tone when composing a sentence and so aid in sentiment analysis. Substitute their meaning for the emoticons: Review contain a range of emotions, including happiness, sadness, and rage. Emoticons are classified into two categories: positive and negative sentiment emotions. Positive emotions are formed of positive emotions

¹ www.amazon.com.

² www.yellowpages.com.

such as love, happiness, and joy, while negative emoticons are composed of negative emotions such as sadness, depression, and wrath.

Punctuation marks, or exclamation marks, serve to highlight the force of a positive or negative remark. Similarly, the apostrophe and the question mark are other punctuation marks.

Words in slang, such as *lol* and *rofl*. These are frequently used to introduce a sense of humor into a remark. Given the nature of opinion tweets, it is plausible to assume that a slang expression in the text suggests sentiment analysis. Substitute their meaning for the slang term.

Punctuation marks, like exclamation marks, serve to highlight the force of a positive or negative remark. Similarly, the apostrophe and the question mark are other punctuation marks.

3.3 Feature extraction

Feature extraction is a key task in sentiment classification as it involves the extraction of valuable information from the text data, and it will directly impact the performance of the model. The approach tries to extract valuable information that encapsulates the text's most essential features. In the work of Venugopalan and Gupta (2015) incorporated other features as it is challenging to extract features from the text. In most cases, punctuations are removed from the text after lowering it in the pre-processing stage, but they used them to extract features and hashtags and emoticons commonly used techniques for feature extractions listed below.

Terms frequency It is one of the simplest ways to express features that are more frequently used in various NLP applications, including Sentiment Analysis, for information retrieval. It considers a single word, i.e., uni-gram or group of two-three words, which can be in bi-gram and tri-gram, with their terms count representing features (Sharma et al. 2013). Term's presence gives the word a value of either 0 or 1. Term frequency is the integer value, which is its count in the given document. TF-IDF can be used as a weighted scheme for better results that will measure the importance of any token in the given document.

Parts of Speech tagging The process of tagging a word in a text (corpus) based on its definition and context is also known as grammatical tagging. Tokens are categorized as nouns, verbs, pronouns, adverbs, adjectives, and prepositions. For instance, "*This mobile is amazing*" may be tagged as follows: (Straka et al. 2016) This :determiner, mobile:noun, is:verb, amazing:adjective. For sentiment mining, an adjective is used more often as it represents the sentiment of the opinion. PoS taggers may be used for this task which is available in NLTK or Spacy. Researches most commonly use Stanford PoS-tagger (Weerasooriya et al. 2016).

Negations These are the words that can change or reverse the polarity of the opinion and shift the meaning of a sentence. Commonly used negation words include not, cannot, neither, never, nowhere, none, etc. Every word appearing in the sentence will not reverse the polarity; therefore, removing all negation words from stop-words may increase the computational cost and decrease the model's accuracy. Negation words must be handled with at most care (George et al. 2013). Negation words such as not, neither, nor, and so on are critical for sentiment analysis since they can revert the polarity of a given phrase. For instance, the line "This movie is good." is a positive sentence, but "The movie is not good." is a negative sentence. Regrettably, some systems eliminate negation words because they are

included in stop word lists or are implicitly omitted since they have a neutral sentiment value in a lexicon and do not affect the absolute polarity. However, reversing the polarity is not straight forward because negation words might occur in a sentence without affecting the text's emotion.

In some cases, neutral sentiment is also included, and neutral evaluations are frequently ignored in many sentiment analysis tasks due to their vagueness and lack of information. In the work of Valdivia et al. (2018) considered to empower neutrality by defining the boundary between positive and negative evaluations to improve the model performance. They use several sentiment analysis approaches to various corpora, extracting their sentiment and filtering out neutral evaluations by consensus, i.e., taking various models based on weighted aggregation. Finally, then they compared the performance of single and aggregated models in categorization. Other contribution introduced in the work of Wang et al. (2020) opinion analysis, namely multi-level fine-scaled sentiment detection with ambivalence handling. The ambivalence handler is detailed, as are the strength-level tuning settings for analyzing the strength and fine-scale of both positive and negative attitudes (Buder et al. 2021). It is capable of delving deeper into the text to uncover multi-level fine-scaled sentiments and distinct emotional types. In the work of Valdivia et al. (2017) suggest the usage of induced ordered weighted averaging operators based on the fuzzy majority for the aggregating polarity from many sentiment analysis methods. Their contribution is to establish neutrality for opinions guided by a fuzzy majority.

Bag of Words (BoW) BoW is one of the simplest approach for extracting text features. BoW will describe the occurrence of words in a document. Bag represents the vocabulary of words using which a vector is formed for each sentence. The main problem with this model is that it does not consider the syntactic meaning of the text. For instance, consider two sentences $s_1 =$ "the food was good", $s_2 =$ "the service was bad". The vocabulary is created for two sentences where $v = \{ 'the', 'food', 'was', 'service', 'bad', 'good' \}$ and the length of the vector is 6 and is represented as $v_1 = [1 1 1 0 0 1]$ and $v_2 = [1 0 1 1 1 0]$. BoW approach performance evaluated using (TF-IDF) which performs better in most cases.

3.3.1 Word embedding

Word embeddings represent words in a vector space by clustering words with similar meanings together. Each word is assigned to a vector, which is then learned in a manner similar to neural networks. It learns and chooses a vector from a predetermined vocabulary. The dimension of the words may be chosen by passing it as a hyperparameter. SG model and the continuous CBOW model are two of the most well-known algorithms for word embeddings. Both of these are shallow window approaches methods in which a short window of some size, such as four or six, is specified, and the current word is anticipated using context words in CBOW, while context words are forecasted using the current word in the SG model. Word embeddings are concerned with learning about words in the context of their local usage, which is specified by a window of nearby terms.

Word2vec word2vec is a 2-layer neural network that is used for vectorizing the tokens. It is one of the famous and widely used vectorizing techniques developed by Mikolov et al. (2013). Word2vec mainly has two models CBOW and SG. The CBOW model predicts the target word using context words, whereas the SG model predicts the target word using context words. With a larger dataset, the SG model performs better.

Global Vectors (GloVe) Global Vectors for word representation have developed (Pennington et al. 2014) by an unsupervised learning approach to generate word embeddings

from a corpus word-to-word co-occurrence matrix. GloVe is a popularly used method as it is straightforward and quick to train GloVe model because of its parallel implementation capacity (Al Amrani et al. 2018).

Fast Text It is an open-source and free library developed by FAIR (Facebook AI Research) mainly used for word classifications, vectorization, and creation of word embeddings. It uses a linear classifier to train the model, which is very fast in training the model (Bojanowski et al. 2017). It supports a CBOW and SG model. Semantic similarities may be found using this model.

ELMo ELMo is a deep contextualized text representation. ELMo contributes to overcoming the limitations of conventional word embedding approaches such as LSA, TF-IDF and n-grams models (Peng et al. 2019). ELMo generates embeddings to words based on the contexts in which they are used to record the word meaning and retrieve additional contextual information. Through pretraining, ELMo can more accurately represent polysemous words in a variety of contexts and is more informative about the text's higher-level semantics (Ling et al. 2020).

3.4 Feature selection approach

Feature Selection Approach is evaluated to identify a data characteristic. A characteristic can be insignificant, significant, or redundant. Various feature selection approaches are used to eliminate irrelevant and superfluous characteristics (Ahmad et al. 2019b; Lata et al. 2020). Feature Selection is a procedure that identifies and eliminates superfluous and irrelevant characteristics from the feature list and thus increases sentiment classification accuracy. In the work of (Hailong et al. 2014; Duric and Song 2012) sentiment analysis for feature selection include lexicon-based and statistical methods. Humans produce features in lexicon-based techniques. Typically, the procedure begins with the collection of phrases with a strong feeling to develop a limited feature set (Kolchyna et al. 2015). The set is augmented with additional terms via synonym detection or web resources (Ghazi et al. 2015; Rizos et al. 2019). The benefit of these approaches is their efficacy, as they carefully address aspects. Choosing handcrafted features is a lengthy and complicated process.

SentiWordNet is a sentiment lexicon built from the WordNet database, with each term accompanied by numerical values indicating positive and negative sentiment. A well-known example is a SentiWordNet lexicon. contrarily, statistical processes are entirely automated and are widely used for feature selection, although they typically fail to distinguish between sentimental and non-sentimental features (Poria et al. 2014; Varelas et al. 2005). The authors (Cambria et al. 2020) proposed SenticNet as a way to include logical reasoning into deep learning models for sentiment analysis.

Statistical techniques for feature selection are typically categorized into four categories: filter, embedding, wrapper, and hybrid.

Filter approach This is the most often used technique of selecting features. It selects features without utilizing any machine learning technique based on the general properties of the training data. The feature is ranked using several statistical metrics, and then the features with the highest rankings are chosen (Adomavicius and Kwon 2011). They are computationally inexpensive and well-suited for datasets with a high number of attributes. The words "Information Gain", "Chi-square", "Document Frequency", and "Mutual information" are all used to refer to fundamental filter algorithms.

Wrapper approach This approach is based on machine learning algorithms since it relies on the output of the machine learning algorithm. Approaches are often iterative and

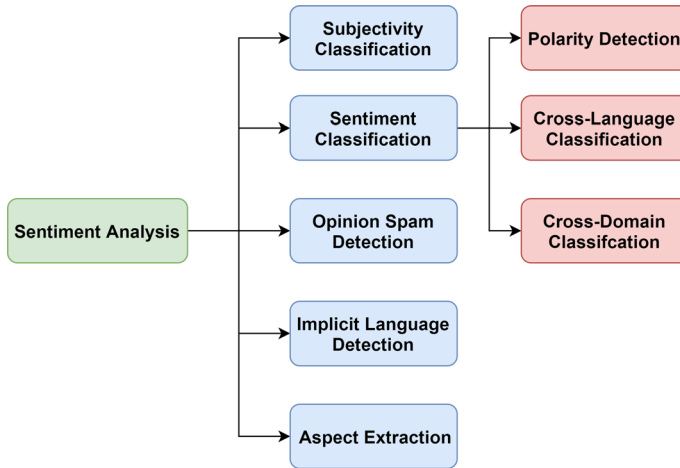


Fig. 3 Task of sentiment analysis

computationally demanding due to this dependency, but they can determine the optimal feature set for that particular modeling algorithm. Wrapper techniques include creating feature subsets (forward or backward selection) plus various learning algorithms (such as NB or SVM).

Embedded approach This method combines the feature selection procedure into the execution of the modeling algorithm. It employs classification methods that have a built-in feature selection capability (Imani et al. 2013). As a result, it is more computationally efficient than the wrapper approach. However, this technique is algorithm-specific (Das et al. 2020). Embedded techniques are frequently based on a variety of decision tree algorithms, including CART (Kosamkar and Chaudhari 2013), C4.5, and ID3 (Quinlan 2014; Mezquita et al. 2020), and additional algorithms like LASSO (Hssina et al. 2014).

Hybrid approach This strategy combines filter and wrapper approaches; hybrid methods generally utilize multiple approaches to produce the optimum feature subset. Hybrid techniques typically achieve excellent performance and accuracy through the use of many approaches. Numerous hybrid feature selection algorithms for sentiment analysis have been developed (Chiew et al. 2019).

3.5 Task of sentiment analysis

Overview of the various task of sentiment analysis as shown in Fig. 3 and explain as follows.

Subjectivity classification This is frequently assumed to be the first stage in sentiment analysis. Subjectivity classification recognizes subjective hints, emotional phrases, and subjective ideas. Tokens like 'hard', 'amazing' and 'cheap' are identified (Kasmuri and Basiron 2017). These indications are used to distinguish objective or subjective text objects. In work of Kasmuri and Basiron (2017) involves determining whether or not there is a particular subject in the given text. Subjectivity classification aims to keep undesirable objective data items out of subsequent processing (Kamal 2013).

Sentiment classification Sentiment categorization is a well-known researched task in sentiment analysis. Polarity determination is one of the subtasks of sentiment classification,

and the term “Opinion analysis” is frequently used while referring to Sentiment Analysis. It is a little duty aimed on determining the sentiment of each piece of text. Polarity is traditionally either positive or negative (Wang et al. 2014). In the work of Xia et al. (2015), the opinion-level context is investigated, with intra-opinion and inter-opinion aspects being finely characterized. Neutral is also included in some cases. With a trained classifier, the cross-domain analysis predicts the sentiment of a target domain. Extracting the domain invariant features and where they are distributed is a commonly used approach (Peng et al. 2018). The cross-language analysis is done similarly by training the model on a dataset from a source language and then evaluating it on a dataset from a different language with limited data. The ambiguity of word polarity is one of the obstacles that sentiment analysis must overcome. In the work of Vechtomova (2017) and Singh et al. (2021b) demonstrated that retrieval-based models provide an alternative to Machine Learning based strategies for word polarity detection. Affective computing and sentiment analysis also have tremendous potential as a subsystem technology for other systems (Cambria et al. 2017). They can augment the capabilities of customer relationship management and recommendation systems by enabling the discovery of which features customers particularly enjoy or the omission of items that have received highly unfavourable feedback from the suggestions.

Opinion Spam Detection Spam Detection has become a significant challenge in sentiment analysis because of the rising interest in e-commerce and review platforms. Opinion spams, often known as fraudulent or phony reviews, are well-written comments supporting or criticizing a product for their benefit. Opinion spam detection seeks to identify three distinct characteristics of a phony review: the review’s content, the review’s metadata, and real-world product expertise (Crawford et al. 2015). Machine learning algorithms are frequently used to assess review material in order to detect dishonesty. The star or point ratings, IP address of the user, geolocation of user, and other information are few Metadata used in detecting spam opinions. In many circumstances, though, it is inaccessible for analysis. Real-world experience and knowledge are included in a third way. For example, if a product with bad ratings and reviews is being rated high for a period, that can be put under suspect and analyzed for opinion spam detection.

Implicit Language Detection Sarcasm, irony, and humor are generally referred to as Implicit Languages. These equivocal and ambiguous form is speech is an arduous task to detect, even by humans sometimes. However, this implicit language is an essential aspect of a sentence and can completely flip the meaning and polarity of the sentence. For instance, consider the phrase ‘Brilliant, I am fired’. The word Brilliant is very positive, but it describes irony or sarcasm when combined with later parts, i.e., “I am fired” it makes the phrase “I am fired” more negative. Investigating signs such as emoticons, laughter emotions, and extensive punctuation mark utilization are more classic approaches for detecting implicit language (Fang et al. 2020; Filatova 2012).

Aspect Extraction Aspect level sentiment analysis is mainly composed of three steps aspect extraction, polarity classification, and aggregation. The process of aspect-based sentiment analysis starts with the extraction of aspect, one of the key processes as this differentiates usual sentiment analysis. Aspects can be extracted using a predefined set of aspects which should be carefully predefined based on the domain on which it is used. Other approaches are more sophisticated approaches like Frequency-based methods, syntax-based methods, supervised and unsupervised machine learning approaches. It has been seen that in reviews (Kanapala et al. 2019), few words are used more frequently than others, and these most frequent terms are more likely to turn out as aspects; this straightforward method can turn out into quite a powerful approach by fact that a significant number of approaches. This approach has few shortcomings because all frequent nouns do

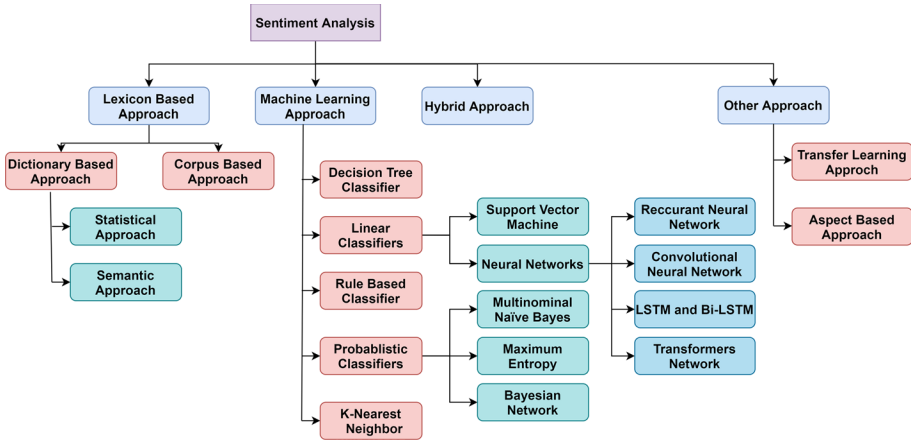


Fig. 4 Approach of sentiment analysis

not refer to aspects, terms like 'bucks,' 'dollars,' 'rupees,' etc. Also, aspects that are not mentioned frequently can be missed by this method. A set of rules can be supplemented with a frequency-based approach to overcome these problems, but these manually crafted rules tend to come from parameters that need to be tuned manually, which is a hectic and time-consuming task. Instead of focusing on the Frequency-based approach. Syntax-based approach can be used as this approach covers the flaws of the frequency-based approach of not detecting less frequent aspects (Bai et al. 2020). In this approach, For example, here, 'Awesome' refers to an adjective referring to the aspect "food" in 'Awesome food.' For this approach, many annotated data covering all syntactical relations should be collected for training the algorithm.

3.5.1 Need of sentiment analysis

Sentiment analysis is incredibly significant since it helps businesses understand their consumers sentiment towards their brand. By automatically classifying the emotions behind social media interactions, reviews, and more, organizations can make informed decisions. Sentiment Analysis refers to the methods and strategies that enable firms to examine data about how their customer base feels about a given service or product. To identify the Polarity: Indicates whether an emotion is good or negative.? Subject: What is the subject of discussion? Who is the holder of the opinion: A thing or person that conveys the sentiment.

Sentiment Analysis is a process that analyzes natural language utterances automatically, discovers essential claims or opinions, and classifies them according to their emotional attitude.

- In the business needs with sentiment analysis has increased consumer happiness through enhanced products, real-time problem detection, and market distinctiveness.
- Customer satisfaction analysis through sentiment analysis: The customer shares his experience with a product and communicates his opinion and attitude about it using natural language comments. This provides us with crucial insight into whether the consumer is satisfied and, if necessary, how we can improve the product.

- Identify and act in real-time problems: Through social media, a customer can immediately voice his discontent to the entire world.

4 Methodology

Three mainly used approaches for Sentiment Analysis include Lexicon Based Approach, Machine Learning Approach, and Hybrid Approach. In addition, researchers are continuously trying to figure out better ways to accomplish the task with better accuracy and lower computational cost. Overview various methods used in Sentiment Analysis as shown in Fig. 4. General Method about the Data collection, Feature selection and Sentiment analysis task are shown in Fig. 2 which understand the overall scenario of sentiment analysis task and overall method workflow.

4.1 Lexicon based approach

Lexicons are the collection of tokens where each token is assigned with a predefined score which indicates the neutral, positive and negative nature of the text (Kiritchenko et al. 2014). A score is assigned to tokens based on polarity such as $+1$, 0 , -1 for positive, neutral, negative or the score may be assigned based on the intensity of polarity and its values range from $[+1, -1]$ where $+1$ represents highly positive, and -1 represents highly negative. In Lexicon Based Approach, for a given review or text, the aggregation of scores of each token is performed, i.e., positive, negative, neutral scores are summed separately. In the final stage, overall polarity is assigned to the text based on the highest value of individual scores. Thus, the document is first divided into tokens of single words, where-after the polarity of each token is calculated and aggregated in the end.

The lexicon-based technique is extremely feasible for sentiment analysis at the sentence and feature level. Because no training data is required, it might be termed an unsupervised technique. On the other side, the primary disadvantage of this technique is domain dependence, as words can have several meanings and senses, and therefore a positive word in one domain may be negative in another. For instance, given the word “small” and the sentences “The TV screen is too small” and “This camera is extremely small”, the word “small” in the first sentence is negative, as people generally prefer large screens, whereas in the second sentence it is positive, as if the camera is small, it will be easy to carry. This issue can be overcome by developing a domain-specific sentiment lexicon or by adapting an existing vocabulary.

The advantage of the lexicon-based approach is that not require any training data and is considered an unsupervised approach by some experts (Yan-Yan et al. 2010). The main disadvantage with lexicon-based approach is that it is highly domain orientated and words pertaining to one domain cannot be used in another domain (Moreo et al. 2012). For instance, consider the word huge it may be positive or negative based on the domain in which it is being used. In “the queue for the movie was huge” the word may be considered positive whereas, in “there was a huge lag in network” the word can be considered negative. Therefore, the polarity should be assigned to words carefully, considering the domain. There are mainly two approaches used in Lexicon Based Approaches: Corpus Based and Statistical Approach explain below, Comparative Analysis of Lexicon Based Classification Method and its individual Advantage and Disadvantage are shown in Table 3.

4.1.1 Corpus based approach

The approach employs semantic and syntactic patterns to ascertain the sentence's emotion. This approach begins with a predefined set of sentiment terms and their orientation and then investigates syntactic or similar patterns to discover sentiment tokens and their orientation in a huge corpus. This is a situation-specific method that requires a significant amount of labeled data to train. However, it aids in resolving the issue of opinion words with context-dependent orientations.

In the work of Park and Kim (2016) used a corpus based method for sentiment analysis. They used linguistic constraints and connectives to find the sentiment of a new token. For instance, tokens on either side of correlative conjunctions like "AND" tend to have the same orientation while words like "OR", but point out opinion change or the tokens on opposite orientations. Although this idea is popularly known as Sentiment Consistency, in practice, this is not that consistent. So, they constructed a graph that contained tokens in vertices and their corresponding word in edges, after which a linear log model was used to identify if two conjoined adjectives were or same or opposite orientation and later clustered into a set of positive or negative words.

The corpus-based approach has the following types of approaches: Statistical Approach and Semantic Approach as explained below.

Statistical Approach The seed opinion words or co-occurrence patterns can be found using statistical approach. The rough idea behind this approach is that if it appears in positive texts more than negative texts, then it is more likely to be positive or vice versa. The key premise of this approach is that if comparable sentiment tokens are frequently observed in the same environment, they will likely have the same orientation. As a result, the orientation of the new token is determined by the frequency with which it appears alongside other tokens detected in a similar context. In Turney and Littman (2003) approach for calculating mutual information can be used to calculate the frequency of co-occurrences of tokens.

A statistical approach is mostly used in several sentiment analysis applications. One such application is detecting manipulated reviews by running a statistical test of randomness popularly known as training test. In work of Hu et al. (2012) expected that reviews written by customers would have random writing styles due to the random backgrounds of customers. They used a book review dataset from amazon.com to confirm their results but, it was found that close to 10.3 percent of products were subjected to online review manipulations.

LSA is another statistical technique for analyzing links between papers and tokens referenced in the documents in order to generate essential patterns connecting to the documents and phrases. In work of Cao et al. (2011) in used LSA to find semantic qualities from reviews to investigate the effect of various features. They engaged program user feedback dataset from the CNETdownload.com website. Their main objective was to find out why few reviews received helpful votes while few reviews helpful votes. They determined various factors which may affect the helpful voting pattern for reviews.

Semantic Approach In this approach, the similarity score is calculated between tokens that are used for Sentiment Analysis. Wordnet is commonly used for this task. Antonyms and synonyms can be easily found using this approach as similar words have a positive score or higher value. In Maks and Vossen (2012) proposed that semantic approach can be used in various applications to build a lexicon model that can be used to describe adjectives, verbs, and nouns to use in Sentiment Analysis. They described,

the in-depth description of subjectivity relations among the characters in a statement conveying distinct attitudes for each character. subjectivity tagged with the knowledge relating to both identity and orientation of attitude holder. In work of Bordes et al. (2014), Bhaskar et al. (2015), Rao and Ravichandran (2009) worked on the WordNet dataset in their work. They determined that the viewer's subjectivity and the actor's subjectivity might be distinguished in some instances (Herscovich and Donatelli 2021).

4.1.2 Dictionary based method

Dictionary based approach consists of a list of predefined set opinion words collected manually (Chetviorikin and Loukachevitch 2012; Kaity and Balakrishnan 2020). The primary assumption behind this approach is that synonyms have the same polarity as the base word, while antonyms have opposite polarity. Large corpora like thesaurus or wordnet are looked upon for antonyms and synonyms, after which it is appended to a group or seed list prepared earlier. In the first stage, initial set of words are collected manually with their orientation. Later the list is expanded by looking at the antonyms and synonyms in the available lexical resources (Singh et al. 2017; Ho et al. 2014). Then the words are iteratively added to the list, and list is expanded. Manual evaluation or correction may be done in the last stage to ensure the quality of it. Stefano and Andrea created SentiWordNet three-way in Baccianella et al. (2010) with the help of automatic annotations of WordNet 3's synsets. Another famous resource thesaurus was created based on online dictionaries. In the work of Park and Kim (2016) suggested a rule-based strategy for labelling sentiment sentences and words in contextual advertising using a dictionary-based approach. This approach is feasible only for small dictionary size. Another disadvantage of all lexicon-based approaches (Hajek et al. 2020), including the dictionary-based approach, is finding opinion words specific for each domain as the polarity may vary. General Procedure step in Lexicon Unsupervised learning category shown in Fig. 6. Summary Analysis of Lexicon Based Classification Method and its Advantage and Disadvantage shown in Table 3 and Summary Analysis of Clustering Method and its Advantage and Disadvantage shown in Table 2.

- *Lexicon Method Based tools* Summary Analysis of Lexicon Based method tools and available Dictionary as explain below
- *Pre-define Dictionary* Utilize a pre-defined list of positive and negative words to determine the polarity of texts based on the frequency with which each category is represented.³
- *SentiWordNet* SentiWordNet assigns numerical sentiment scores to WordNet synsets that are either positive or negative.⁴
- *Bing Liu's Sentiment Lexicon* A dictionary has 4783 helpful positive and negative words.⁵
- *SentiStrength* Unless modified by any additional classification rules, texts are categorized according to the highest positive or negative score for any constituent word.⁶

³ <http://www.wjh.harvard.edu/~inquirer/>

⁴ <http://sentiwordnet.isti.cnr.it>.

⁵ <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>.

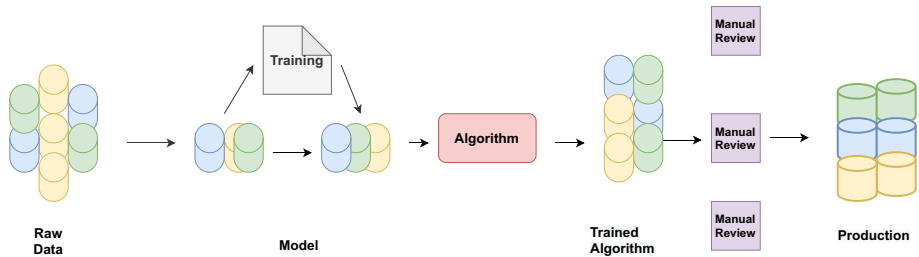
⁶ <http://sentistrength.wlv.ac.uk>.

Table 2 Advantages and disadvantages of sentiment analysis clustering approaches

Technique	Advantage	Disadvantage
K-means	<p>Due to the fact that these algorithms accomplish complexity in linear time, they are well-suited for massive datasets</p> <p>The algorithm is not required to know the class of a document in advance. Does not require training</p> <p>The algorithm is not required to know the class of a document in advance</p> <p>Does not require training</p> <p>This means that it is devoid of human intervention</p> <p>Memory requirements are minimal</p> <p>A low cost, efficient, and extremely convenient approach for sentiment analysis</p>	<p>If there are any ambiguities, it is not accurate enough</p> <p>The issue with these techniques is that they are sensitive to the starting center locations and make the assumption that the number of clusters is known</p> <p>These methods struggle with anomalies and noise, as well as a non-convex cluster</p>
Fuzzy c-means	<p>It always converges</p>	<p>It is susceptible to initial assumptions and may terminate at local minimum</p>
Agglomerative algorithm	<p>Do not depend on the number of clusters or the center gravity</p>	<p>Computation time is high</p> <p>Hierarchical algorithms have a high degree of time complexity. Although it has superior clustering properties, its high cost limits its use in large-scale data applications</p>
Divisive algorithm	<p>Do not rely on the number of clusters or the center</p> <p>Simple to implement and, in some situations, produces the best results</p>	<p>Algorithms can never erase prior actions</p> <p>No target reduction is achieved directly. At sometimes, it can be challenging to</p> <p>Determine the correct number of clusters using the dendrogram</p>

Table 3 Summary analysis of Lexicon based classification method and its advantage and disadvantage

Technique	Advantage	Disadvantage
Dictionary based	Trained data are not required Give good outcomes for domains with fewer bands Quick access to the vocabulary of word meanings	Opinion terms with a specific content orientation Incapable of finding opinion terms with a specified content oriented domain that are not included in the lexicon
Corpus based	The capacity to identify expressions of opinion with a particular content orientation. When domains are distinct, provides superior results	Due to the lexicon's broad scope, performance varies The difficulties of providing substantial texts while also being able to cover all of the text terms means that they cannot be used individually

**Fig. 5** General procedure for sentiment analysis in supervised machine learning category

- *Opinion Identification* Opinion Finder recognizes subjective statements automatically and highlights several characteristics of their subjectivity, such as the source (holder) of the subjectivity and terms included in phrases indicating positive or negative views.⁷
- *National Taiwan University Sentiment Dictionary* contain 2812 Positive and 8276 negatives words.⁸
- *WordNet-Affect* WordNet-Affect is a WordNet Domains extension that includes a subset of synsets that are appropriate for representing affective notions associated with emotional words.⁹
- *Affective Norms for English Words* Affective Norms for English Words is a collection of normative emotional ratings for a large number of English words. This collection of linguistic materials has been graded for enjoyment, arousal, and dominance in order to establish a baseline for future research on sentiment and attentiveness.¹⁰

⁷ <http://mpqa.cs.pitt.edu/opinionfinder>.

⁸ <http://academiasinicanlplab.github.io>.

⁹ <http://csea.php.ufl.edu/media/anevmessag.html>.

¹⁰ <http://csea.php.ufl.edu/media/anevmessag.html>.

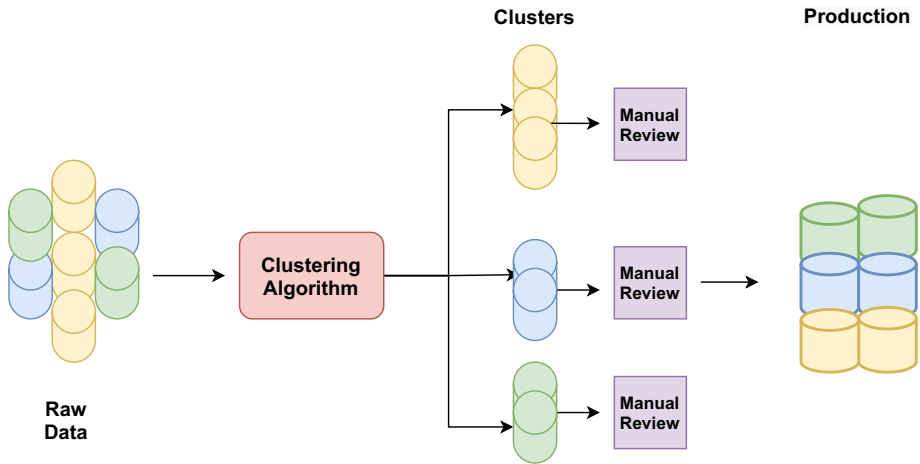


Fig. 6 General procedure for sentiment analysis in unsupervised machine learning category

- *LingPipe* can work on a wide range of activities, such as identifying topics, identifying named entities, parsing and indexing documents, database text mining, word segmentation, sentiment analysis, and language identification.¹¹
- *Apache OpenNLP* provides support for parsing sentence, tokenization, part-of-speech tagging, segmentation, chunking, named entity extraction, language recognition, and coreference resolution.¹²
- *Lexicon Sentiment Dictionary* A language used in politics.¹³

4.2 Machine learning approach

Machine Learning Algorithms can be used to categorize sentiments. Sentiment analysis is the process of identifying and quantifying the sentiment of text or audio using natural language processing, text analysis, computational linguistics, and other techniques. Data Collection from Social Media and processing step for sentiment analysis in Supervised Learning category shown in Fig. 5 and Unsupervised learning category shown in Fig. 6. There are two primary in Machine Learning approaches to sentiment analysis:

- Supervised machine learning
- Lexicon-based unsupervised learning

This task can be accomplished using both supervised and unsupervised learning methodologies. Unsupervised strategies for sentiment analysis by utilizing knowledge bases, ontologies, databases, and lexicons that include detailed knowledge that has been selected and

¹¹ <http://alias-i.com/lingpipe>.

¹² <http://incubator.apache.org/opennlp>.

¹³ <http://www.lexicoder.com>.

prepared specifically for sentiment analysis. Supervised learning methods are more commonly used due to their accurate results. These algorithms need to be trained on a training set before it is applied to the actual data. Features may be extracted from text data.

The machine learning technique utilizes syntactic and/or linguistic factors to address sentiment classification as a standard text classification issue utilizing syntactic and/or linguistic factors. The categorization model associates the underlying record's features with one of the class labels. The model is then used to predict a class label for a given instance of an unknown class. When an instance is assigned only one label, we have a difficult categorization challenge. When a probabilistic value of labels assigned to an instance, this is referred to as the soft classification issue. Machine learning enables systems to acquire new abilities without being explicitly programmed to do so. Sentiment analysis algorithms can be trained to read beyond simple definitions to comprehend contextual information, sarcasm, and misapplied words. Commonly used algorithms include:

4.2.1 Naive Bayes (NB)

NB technique is utilized for both categorization and training. NB is a Bayesian classification approach based on the theorem of Bayes. NB is a probabilistic classifier that uses Bayes theorem to predict the probability of a given set of features as part of any particular label. The conditional probability that event A occurs given the individual probabilities of A and B and conditional probability of occurrence of event B . Here it is assumed that features are not dependent. BoW model may be used for feature extraction. Generally, NB is applied when the training data size is small. NB classified as positive 10% more accurately than negative classification. This led to a decrease in average accuracy when it was taken. In the work of Kang et al. (2012) solved this problem using an improved version of the NB classifier. They tested this model on a restaurant review dataset. In work of Tripathy et al. (2015) used machine learning for the classification of reviews. They proposed a NB model along with a SVM model (Hajek et al. 2020; Bordes et al. 2014). They used a movie review dataset for training and testing the models. Two thousand reviews were trained after pre-processing and vectorization of the training dataset. Count Vectorizer and *TF-IDF* were used before training the machine learning model. NB model proposed in Tripathy et al. (2015) gave an accuracy of 89.05 percent in a K-fold Cross-validation. The performance was better when compared to other models using the probabilistic NB algorithm (Calders and Verwer 2010).

4.2.2 Support vector machine (SVM)

SVM approach, which uses hyper-planes, is used to analyze data and define decision limits in this technique. SVM are a type of non-probabilistic supervised learning technique that is frequently used for classification tasks. SVM primary objective is to determine the hyperplane that best separates the data into distinct classes. As a result, SVM seeks out the hyperplane with the highest feasible margin. In work of Li and Li (2013) used Support Vector Machines for sentiment polarity Classifier. Classifying reviews based on their quality is one of the many purposes for which SVM are utilized. Chen and Tseng (2011) used two multiple class SVM based approaches. First being One-vs- all SVM and Multi-class SVM to classify reviews. Second, a method was proposed to evaluate the quality of the product review dataset quality by considering it as a classification problem. In work of Dave et al. (2003) worked on MP3 reviews and digital camera reviews. Borg and Boldt

(2020) used Linear SVM and VADER to predict customer reviews sentiment. The review belonged to Huge Swedish Telecom Corporation. The dataset was huge and consisted of 168, 010 emails for training. They used Swedish sentiment Lexicon and VADER sentiment for initial labeling. Their (Li et al. 2019a) linear SVM model performed marvelously with an F1 score of 83.4 percent and a mean AUC of 0.896. Furthermore, their model highlighted a pattern that was predicted in Email conversations, using which sentiment of unseen Email was being predicted. In work of Xia et al. (2020) urged that subjectivity of opinion and credibility of expresser should be considered, unlike regular Binary Classification Problem. A framework (Wu et al. 2020) was proposed to summarize opinions on microblogs. They found and retrieved the topics mentioned in the opinions related to users inquiries and then categorized the opinions using SVM. Ali et al. (2020) also worked on Twitter tweet data for the experiment. They found it to be beneficial to aggregate the opinions for microblogs.

4.2.3 Logistic regression (LR)

A machine learning technique known as logistic regression works by multiplying an input value by a weight value. It is a classifier that learns which input properties are most helpful in identifying positive and negative classes. Logistic regression is a probabilistic regression analysis used for classification tasks. For binary classification applications, logistic regression is commonly deployed. When there are multiple explanatory variables, logistic regression calculates the ratio of odds. Logistic regression uses Maximum-likelihood to calculate best parameters. The independent variables may belong to any category i.e., Continuous, Discrete (ordinal and nominal). LR model (Hamdan et al. 2015) that the dependent variable is binary, and there is little or no multicollinearity between the predicting variables.

4.2.4 Decision tree (DT)

DT Classifier is a supervised learning technique where a tree is built using the training example to classify the polarity of the text. DT uses a condition to divide data into parts recursively. RF are used frequently than DT which combines multiple DT to avoid overfitting and improve accuracy. DT may be built using several algorithms like CART, ID3, C5.0, C4.5 (Revathy and Lawrance 2017; Hssina et al. 2014; Singh and Gupta 2014; Patel and Prajapati 2018). These are used to identify the best fitting attribute which needs to be placed in the root (Gower 1966; Revathy and Lawrance 2017; Patil et al. 2012). Yan-Yan et al. (2010) using a graph-based strategy, They proposed a propagation strategy for integrating sentence-level and sentence-level features. These two phrase characteristics are referred to as inter and intra document verification. They tried to argue that determining the sentiment classification of a review sentence entails more than simply examining the statement's components. They investigated the camera domain and compared their results to those obtained using SVM and NB Classifiers. In the work of Jain et al. (2021a) tagged data that can be used to distinguish between genuine and fraudulent reviews. Additionally, we used two distinct datasets to test various machine learning techniques for categorization (Yelp hotel review dataset, Yelp restaurant review dataset).

4.2.5 Maximum entropy (ME)

Conditional Exponential Classifiers: Conditional exponential classifiers encode labelled feature sets as vectors or arrays of integers. This vector is then used to compute feature weights, which can be used to select the most likely label for the feature set. Entropy is a measure of unpredictability. The Entropy is maximum for uniformly distributed data. The input data consists of texts and ratings from 1-5 polarity assigned to it. Most popularly used algorithms include SVM, NB, ME (Khairnar and Kinikar 2013; Kaufmann 2012) used ME Classifier to detect parallel sentences in any two-language pair, which have less training data. The other models used either required a massive amount of training dataset or used a language-specific technique (Bergsma et al. 2012), but their model showed improved results could be produced using any pair of languages. This will enable the establishment of parallel corpora for various languages.

4.2.6 K-nearest neighbours (KNN)

KNN algorithm is not extensively used in sentiment analysis but has shown to produce good results when trained carefully. It operates on the fact that the classification of a test sample will be similar to nearby neighbours. The K value may be selected on any hyper-parameter tuning algorithms like Grid search or Randomized search cross validation. The polarity may be hard voted based on K nearest neighbors values, or soft addition may be done to find overall polarity.

4.2.7 Semi-supervised learning

In this case, where the training dataset contains both labelled and unlabelled data, semi-supervised learning appears to be a viable option (Zhu and Goldberg 2009). It is motivated that while gathering unlabelled data is relatively easy in many real-world applications, such as collecting articles from various blogs, labelling is expensive or labour-consuming because labelling the training dataset is typically performed by humans. Ortigosa-Hernández et al. (2012) introduced in the work of a real-world situation in which the user attitude is defined by three distinct (but related) target variables: subjectivity, sentiment solarly, and will to influence. In the work of Janjua et al. (2021) framework for semi-supervised machine learning that combines pre-processing and classification algorithms for unlabelled datasets.

Summary of different Sentiment Analysis Techniques its Advantage and Disadvantage shown in Table 5.

4.3 Hybrid approach

Hybrid approach combines machine learning and lexicon-based approaches. Hybrid is a term that refers to the combination of machine learning and lexicon-based techniques to sentiment analysis. The hybrid technique combines the two and is extremely popular, with sentiment lexicons playing a significant role in the majority of systems. Sentiment analysis is a hybrid approach, including both statistical and knowledge-based methods for polarity recognition. In the work of Hassonah et al. (2020a) proposed a hybrid machine learning approach using SVM and two feature selection techniques using the multi-verse optimizer

Table 4 Summary analysis of machine learning classification algorithm and its advantage and disadvantage

Technique	Advantage	Disadvantage
NB	<p>Easy to Implement</p> <p>Less training data required</p> <p>Requires less training time and data than other approaches</p>	<p>Assumes that attributes are mutually independent</p> <p>Model can face Zero Frequency Problem</p> <p>Limited by class imbalance, as a probability value must be estimated for each conceivable value</p>
SVM	<p>Most popular algorithm for SA</p> <p>Gives good accuracy for large dataset</p>	<p>Fine tuning of model is very difficult and tedious</p> <p>Long training time for large datasets</p>
DT	<p>Easy to construct</p> <p>Less time for training</p> <p>Extensive dataset is not required for training</p>	<p>Model is more prone to overfit</p> <p>Highly Domain oriented model will be constructed</p>
KNN	<p>Non-linear decision boundaries can built</p> <p>Data can be added constantly with time without explicit training</p>	<p>Higher prediction complexity for large dataset and dimensions</p> <p>Equal importance is given to all features</p>
RF	<p>The model is robust and very stable</p> <p>Model is significantly less likely to overfit</p>	<p>The model is very complex</p> <p>Longer time for training the data</p>
ME	<p>Works on a probabilistic model. Therefore, lots of training data is not required</p>	<p>Model is domain-oriented and will not perform well for different dataset</p>
LR	<p>Simplest models to implement and training</p> <p>No assumptions made about distributions of classes in feature space</p>	<p>Linear boundaries are constructed</p> <p>Accuracy is low for complex dataset</p>
DNN	<p>In comparison to other DL models, this one is rather simple to implement</p>	<p>Over-fitting problem</p> <p>Train more slowly than other models</p> <p>Complicated and computationally costly</p>
RNN	<p>Capture of sequential data, which is critical for sentiment text categorization</p>	
LSTM	<p>More efficient than RNN</p> <p>Can map out long term dependencies</p>	<p>Very complex model</p> <p>Training time is high</p>
Bi-LSTM	<p>Can find out dependencies in both the directions</p> <p>Better results can be obtained as compare to other method</p>	<p>Computationally very costly</p> <p>Training is very slow</p>
CNN	<p>High Accuracy, Fast Training</p>	<p>Design and maintenance are time-consuming</p>
Transformer	<p>Uses self-attention models for finding dependencies</p> <p>Focuses only on important part of the sentence</p>	<p>Less popular and less used</p> <p>Require huge data</p>

Table 5 Summary of different machine learning sentiment analysis techniques its advantage and disadvantage

Technique	Advantage	Disadvantage
Supervised learning Method	With the ability to assess a wide range of topics, Effectiveness in identifying the issue of subjectivity	Designated training documents capability of revealing subjective it demands human work and linguistic knowledge Labeled data required
Unsupervised learning Method	It do not need much human effort Labeled data not required	Its capacity to do so is not yet fully established
Semi-supervised learning Method	Productive results in the presence of ambiguity	If the unlabelled samples are noisy, the classifier suffers difficulty
Domain-oriented Sentiment Analysis Lexicon-based Dictionary-based	Include a two-tiered classification system that includes both positive and negative opinions Classify sentiment on a certain scale Eliminating objective sentences from a document can help improve sentiment analysis effectiveness	Neglect the variety and depth of human feelings, as well as more emotional categories, aspects, and granularities Depends on Techniques

and Relief algorithms (Chang et al. 2020). Sentiment analysis task (Al Amrani et al. 2018) proposed using machine learning-based hybrid approach including RF and SVM. They have shown that the individual models of SVM and RF had an accuracy of 81.01 and 82.03 percent, respectively, whereas the hybrid model combining both the algorithms had an accuracy of close to 84% in the product review dataset provided by *amazon.com*. Few researchers have proposed a hybrid architecture involving both lexicon-based and automated learning techniques to enhance the results. This is still a hot topic for researchers, and lots of research needs to be done.

In work of Hassonah et al. (2020b) used Twitter data for training. As many as 6900 tweets were extracted for training using the Twitter API. The results showed that their model outperforms most of the models while reducing the total number of features up to 96%. They also pointed out the capacities of Hybrid models and concluded that Hybrid models could outperform all the models with proper architecture and precise selection of hyperparameters (Chang et al. 2020). The Hybrid model outperformed both the model in all other metrics and comparisons. They concluded that although their Hybrid model performs better than individual models, there are still many research opportunities available to improve the performance of the hybrid model by tweaking and training the model. There are various Method Summary Analysis of Supervised Machine learning Classification Algorithm and its Advantage and Disadvantage shown in Table 4.

4.4 Neural network

Neural Network- In work of Van de Camp and Van den Bosch (2012) presented the use of Neural networks and SVM in supportive relationships. They used biographical texts to confirm their results. They were successful in marking relations between two individuals as neutral, positive, or negative. They concluded that a SVM and a single layer Neural Network had shown improved results. In work of Moraes et al. (2013a) presented a comparative empirical analysis between SVM and ANN in document-level sentiment analysis. The motive of this comparison is that SVM was widely being used as an algorithm for opinion mining as it had shown its capacity of getting accuracy. ANN, even though with good potential, did not have much attention. In Moraes et al. (2013b) they discussed all the aspects related to both ANN and SVM, including their requirements, their accuracies, and other contexts in which each model can perform the best. They have also implemented a consistent evaluation framework using well-known participants in supervised methods for selecting features and weights in orthodox BoW models. According to Medhat et al. (2014), Ravi and Ravi (2015) ANN had the edge over SVM in most of the cases, except very few cases where there was a data imbalance. They used product reviews of Books, GPS, and camera from *amazon.com* and Movie reviews to come to this conclusion. They also pointed out potential problems and drawbacks of each model. An important disadvantage of each model is the computational cost taken in training by ANN and the computational cost of SVM model in run time.

The traditional RNN (Liu et al. 2016) were used for various NLP tasks as they used the previous time step information to predict the current time step, which ensures the usage of previous information and acts as memory as it remembers some information about a sequence. The most significant achievement or advantage of RNN was that it used previous information, thus remembering the previous information, which acted as memory. The main disadvantage of a traditional RNN is that it suffers from vanishing and exploding gradient descent, which means it cannot remember long-term

relationships in the sequence. In the case of Bi-LSTM (Plank et al. 2016) uses the previous time step information along with next time step information to predict the current time step, as pass the sequence in both the ways forward as well as backward. Deep learning has identified new avenues for emulating the peculiarly human potential, for example-based learning. While this method of bottom-up learning is successful for picture classification and object recognition, it is ineffective for NLP (Cambria et al. 2020). They blend top-down and bottom-up learning in their work using an array of symbolic and subsymbolic AI tools and apply them to the intriguing challenge of text polarity detection.

RNN (Donkers et al. 2017) have proven to improve results when trained on sufficient data and computations. Variants of RNN (Pham and Le-Hong 2017) like LSTM (Bandara et al. 2020), GRU (Cheng et al. 2020), Bi-LSTM (Abid et al. 2019; Cho and Lee 2019) have been used extensively in Sentiment analysis and related NLP task (Abid et al. 2019; Khan et al. 2016). Attention models are being introduced recently, which gives models an edge over another model. Recent transfer learning techniques using BERT (Devlin et al. 2018) and GPT (Ethayarajh 2019) are gaining the attention of researchers as the model is already trained on a massive corpus for days on high-end GPU and Super computers. Weights can be fine-tuned using the training dataset to get accurate results. Deep learning-based techniques are becoming highly popular due to their outstanding performance in recent times. In the work of Yadav and Vishwakarma (2020) and Wadawadagi and Pagi (2020) gives a detailed assessment of common deep learning techniques that are widely employed in sentiment analysis. To detect the intensity of sentiments and emotions, a stacked-ensemble model based on deep learning was developed (Akhtar et al. 2020). To better capture both long-term dependencies and local features, they employ GloVe word embedding, bidirectional GRU, bidirectional LSTM, attention mechanism, and CNN. The authors (Basiri et al. 2021) suggested a model for sentiment analysis based on attention (CNN-RNN). In the work of Alhumoud and Al Wazrah (2021) conduct a systematic review of the literature to identify, categorize, and evaluate state-of-the-art works utilizing RNNs for Arabic sentiment analysis.

In 2017, researchers at the Google Brain team, Google Research and University of Toronto came up with the concept of Transformers in their paper (Vaswani et al. 2017) "Attention is all you need," which revolutionized the NLP applications. This model is a stack of encoder-decoder models consisting of self-attention, multi-headed attention layers, and normalizing and feed-forward layers. The input is word embeddings along with the position vector, which specifies the position of the vector and the inputs can be given parallelly, unlike other models which take serial or sequential inputs. In the encoder part, Self attention is calculated for each token with the help of key-value and query vector. This is done multiple times and stacked over each other, forming a multi-headed attention layer passed to the feed-forward layer (Kitaev and Klein 2018). There are six encoder and six decoder layers present in the model. The input to the decoder from an encoder is the two vectors K and V . The decoder layer has three layers: the self attention layer, then a normalization layer which is the same as the encoder layer second one is the encoder-decoder attention. The output self-attention and the input from encoder are used to produce an output vector followed a feed forward network along with a linear and a SoftMax layer (Juraska and Walker 2021). There are few skip connections or residuals present in both encoders and decoders for better results.

There are various methods for sentiment analysis using machine learning and deep learning used by the author are shown in Table 6. we are using several terms in Table 6

Table 6 Sentiment analysis by using deep learning and machine learning method

References	Method	Lang.	Dataset	Task	Result
Qiu et al. (2010)	Rule Based	English	automotvieforums.com	SA	Recall = 0.375
Bai (2011)	Markov Model, SVM, NB	English	IMDb	SC	SVM = 76.68, NB = 75.05
Fan and Chang (2011)	Chi-square, SVM	English	ebay.com, wikipedia.com, opinions.com	SA	Chi (matching) Pos = 74.8 Neg = 57.6 Uni-gram Pos = 59.1 Neg = 67.4
Rui et al. (2013)	NB, SVM, rule-based	English	convinceme.net	SC	Precision:Recall NB = 0.98:0.65 ANN = 0.86:0.85 SVM = 0.83:0.85
Moraes et al. (2013b)	NB, ANN, SVM	English	amazon.com	SC	
Socher et al. (2013)	RNTN and RNN	English	Sentiment treebank	SC	Beer = 77.04, Camera = 81.02
Lakkaraju et al. (2014)	Joint Multi aspect sentiment model + RNTN	English	Beer Review and camera review	SC	
Vateekul and Koomsubha (2016)	LSTM and Dynamic CNN	English	twitter	SLC	LSTM-75.30, DCNN-75.35
Poria et al. (2016)	Temporal CNN,RNN	English	MOUD Youtube dataset	Multi-modal SC AND EC	CNN,RNN = 96.55
Zhao et al. (2017)	CNN LSTM	English	amazon reviews on digital camera, phone and laptop	SLC	CNN-87.7, LSTM-87.9
Dragoni and Petrucci (2017)	RNN	English	Dranziera dataset	SC	Dranziera dataset = 0.81
Chen et al. (2017)	Deep fusion CNN	English	VSO : Flickr Images MVSO-EN: the dataset includes concepts related to emotions expressed	Multi-modal SC	VSO = 84.7, MVSO-EN = 73.7
Wu and Chi (2017)	Quadratic LSTM-RNN	English	Stanford sentiment treebank(SST) and SICK dataset	SLC	SST-86.6(binary) SICK-87.28
Uysal and Murphey (2017)	CNN,LSTM, CNN+LSTM	English	IMDb, Sentiment 140, Nine public sentiment reviews,amazon multi domain dataset	Document level SC	S140-71.5 Nine-77.1 IMDB-89.1 Amazon-85.4

Table 6 (continued)

References	Method	Lang.	Dataset	Task	Result
Yuan et al. (2018)	Bi-LSTM with attention mechanism	English	amazon multi-domain and sanders twitter sentiment dataset	Multi-domain SC	amazon = 87.69, sanders = 86.32
Zhang et al. (2018)	Dynamic Memory Networks	English	semeval2014, semeval 2016	SC	DMN = 0.84
Rao et al. (2018)	LSTM	English	Yelp 2014, 2015, IMDB	Document level SC	Yelp2014-63.9 Yelp2015-63.8, IMDB-44.3
Naseem et al. (2020)	Deep Intelligent Contextual Embedding	English	Airline Dataset	SA	airline dataset = 0.95
Basiri et al. (2021)	Attention based Bidirectional CNN-RNN Deep Model (ABCDM)		Sentiment140, Airline, Kindle dataset, movie review	SA	(ABCDM) Kindle dataset = 0.93, Airline = 0.92, movie review = 0.90, Sentiment140 = 0.81

as SA indicates Sentiment Analysis, SC indicates Sentiment Classification. Categorization of an individual method as Supervised Learning Method, Unsupervised Learning Method, Semi-supervised Learning Method, Domain-oriented Sentiment Analysis Lexicon-based and Dictionary-based with its advantage and disadvantage are shown in Table 5.

4.5 Other approaches

4.5.1 Aspect based sentiment analysis (ABSA)

ASBA is valuable and rapidly growing part of sentiment analysis that has gained prominence in recent years. Three critical phases compose aspect-level sentiment analysis: aspect detection, polarity or sentiment categorization, and aggregation. Aspect detection is a critical stage in Aspect-based Sentiment analysis, as it is followed by sentiment calculation. Aspects are mined either by using pre-defined implicit aspects or can be mined explicitly (Rana and Cheah 2016). Machine learning techniques, along with NLP techniques, are used to mine aspects out of a sentence. Aspect level sentiment analysis is most popular among product reviews or hotel reviews, as this approach will help them identify various aspects focused by the review writers and help them rectify aspects that have a negative sentiment (Tran et al. 2019). This is useful to both consumers as well as producers. For instance, for a hotel review dataset, implicit aspects may be defined as A= taste, service, value, miscellaneous. For instance, consider a review R= “the food was awesome, but service was slow”, this review consists of two aspects which are food and service, i.e., A = taste, service and the corresponding sentiment words are awesome and slow i.e. S = awesome, slow which be classified as P = positive, negative which when aggregated is neutral. If we consider the sentiment scores based on their positiveness or negativeness, the aggregated polarity may vary.

Aspect level sentiment analysis has many challenges as it to identify the individual aspect(implicit or explicit) and classify as per sense is challenging to mine aspects (Tubishat et al. 2018), Therefore, complex algorithms like LSTM, Bi-LSTM or pre-trained models like BERT, GPT-2 may be used to accomplish the task. The researchers avoid vanilla RNN as it faces many problems like vanishing and exploding gradient descent. It is seen that recently attention-based models are being used in aspect detection. The next step after aspect detection is polarity assignment to those mined aspects. There are multiple approaches to perform the task, Machine learning algorithms may be used to complete the task, or a dictionary-based approach may be used. Assigning the polarity to the aspect an aggregation score may be calculated to find the overall polarity of the sentence. Hard or soft voting is used to determine the sentence’s overall sentiment. Consumer sentiment is assessed concerning qualitative content, quantitative ratings, and cultural factors in order to forecast consumer recommendation decisions (Jain et al. 2021c, d).

4.5.2 Transfer learning

Transfer learning is one of the advances techniques in AI, where a pre-trained model can use its acquired knowledge to transfer to a new model. Transfer learning uses the similarity of data, distribution, and task. The new model directly uses the previously learned features without needing any explicit training data. Training data may be used to fine-tune

the model to a new task. This technique can be used to transfer knowledge of one domain to another domain. This methodology has grown as a transfer learning technique because it can produce great accuracy and results while requiring significantly less training time than training a new model from scratch (Celik et al. 2020). Transfer learning is frequently used in sentiment analysis to classify sentiments from one field to another field. In Meng et al. (2019) developed a multiple-layer CNN based transfer learning approach. They used the weights and biases of a convolutional and pooling layer from a pre-trained model to model. They used the features from pre-trained model and fine-tuned weights of Fully connected layers. This approach can produce good results when large labeled data sets are absent and similarities in the tasks accomplished by the models. In the work of Bartusiak et al. (2015), applied Transfer Learning to propose the sentiment analysis challenge. They used this technique to evaluate the sentiment at the document level in the polish language. They used N-gram and Bi-gram to encode complex words and phrases. They used two different datasets from two different domains to provide evidence that knowledge gained from the training model using dataset of one domain can be used for a dataset of another domain. Sentiment Analysis by using Deep learning and Machine Learning Method as shown in Table 6.

In 2018, Google AI Language Researchers open-sourced a new model for NLP called BERT. It has a breakthrough and has taken the industry of deep learning by storm due to its performance. In the work of Han et al. (2021) Transformer network revolutionized the area of NLP and replaced the usage of LSTM and Bi-LSTM. The main advantage is that Transformers do not suffer from vanishing or exploding gradient problems as they do not use recurrence at all, and also, they are faster and less expensive to train. BERT is an extension of the Transformers model proposed (Vaswani et al. 2017) in the "Attention is all you need" paper. BERT uses transformers, an attention mechanism that learns contextual relationships between words or sub-words in a given text. The input in this model contains the word embeddings and position embeddings, unlike transformers, but also has an extra vector representing the sentence it belongs to handle two or more sentences at a time. BERT consists of encoders based transformers; the encoder part is similar to the transformer encoder. BERT has two models BERT base with 12 encoders stacked with 110 million parameters and BERT large model with 24 encoders stacked with 330 million parameters. BERT model trained in two stages pre-training and fine-tuning. This is the model main advantage as the fine-tuning with the dataset can be done as per the task. Such as sentiment analysis (Singh et al. 2021a), aspect detection (Li et al. 2019b), spam detection (Yaseen et al. 2021), Transformer models for text-based emotion detection (Acheampong et al. 2021), impact of coronavirus(singh2021sentiment). A single sentence or a pair of sentences can be represented as a successive array of tokens using the task-specific BERT architecture (Gao et al. 2019). In the work of Sun et al. (2019) transform ABSA to a sentence-pair classification problem, such as question answering and natural language inference, by constructing an auxiliary sentence from the aspect. BERT pre-trained model has been fine-tuned.

4.6 Multimodal sentiment analysis (MSA)

MSA adds a new level to standard text-based sentiment analysis by incorporating additional modalities such as audio and visual data. Several studies have attempted to discern sentiment analysis in social multimedia using a variety of multimodal inputs, including visual, audio, and textual data (Soleymani et al. 2017). Social multimedia sites such as YouTube, video blogs (vlogs), or spoken evaluations contain expressions of sentiment, such as

Fig. 7 Confusion matrix

ACTUAL	Positive	True Positive	False Negative
	Negative	False Positive	True Negative
		Positive	Negative
		PREDICTED	

a video portraying a person discussing a product or a movie. Typically, spoken transcripts are examined separately from face and voice expressions, and the results of unimodal, text-based sentiment analysis are combined in post to create a “MSA” system. It may be bimodal, consisting of various combinations of two modalities, or trimodal, consisting of three modalities (Stappen et al. 2020). The majority of MSA techniques focus on developing complex fusion processes, ranging from attention-based models to tensor-based fusion.

MSA is a rapidly expanding area of study. A key area of opportunity in this subject is to enhance the mechanism of multimodal fusion. In the work of Majumder et al. (2018) and Poria et al. (2018b) feature fusion technique that is hierarchical in nature, merging the two modalities first and subsequently all three modalities. MSA of human spoken language has developed into a significant subject of research (Liu 2012; Poria et al. 2017). Unlike traditional emotional learning tasks that require the use of single modalities (text, speech), multimodal learning makes use of many sources of information, including language (text/transcripts/ASR), audio/acoustic, and visual modalities.

5 Performance evaluation parameter

The majority of state-of-the-art sentiment analysis makes use of accuracy, F1 score, and precision. Sentiment analysis using deep learning architectures: a review utilizes recall and accuracy as performance metrics. These metrics are as follows:

True Positive(TP): The number of positive reviews that have been correctly classified.

True Negative(TN):The number of negative reviews correctly classified as negative.

False Positive(FP): Number of incorrectly classified positive review.

False Negative(FN): Number of incorrectly classified negative review.

Precision Precision is defined as the ratio of correctly classified positive samples to the total number of samples predicted as positive. This metric can be used to indicate the strength of the prediction. i.e., if a model has 100 percent precision, all the samples evaluated as positive are confidently positive.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall Recall is also known as sensitivity. It is defined as the ratio of actual positive instances out of a total number of positive instances present in the classification. It measures the misclassifications done by the model. Precision and recall are inversely

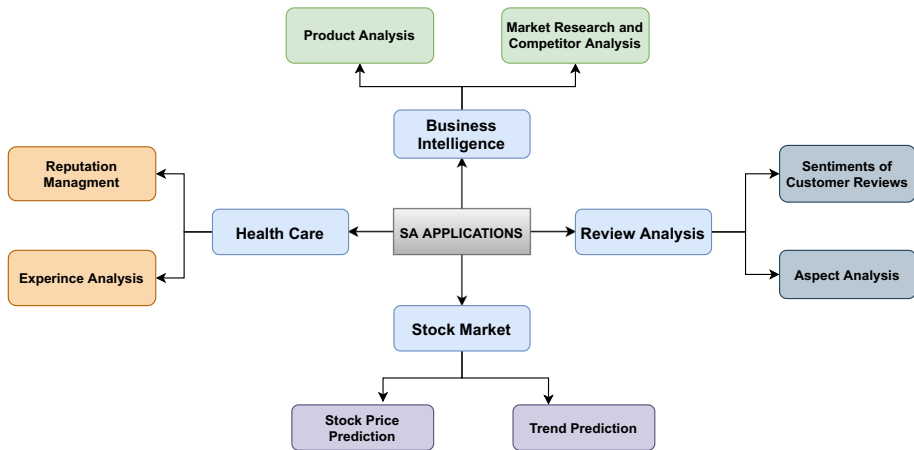


Fig. 8 Applications of sentiment analysis

proportional to each other. Therefore it is impossible to increase both Precision and Recall at the same time. A recall is used in cases where the capture of a class is dominant.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

F1 score F1 score is the harmonic mean of Recall and Precision. It is the most used metric after Accuracy. It is used when we are unable to choose between Precision or Recall. F1 score manages the trade-off between recall and precision.

Accuracy This is the most commonly used metric in all the classification tasks. Accuracy defines how accurate the model is. It is the ratio of correct classification to total predictions done by the model. Accuracy is a good metric to use for sentiment classification for a balanced dataset.

Specificity Specificity Is the opposite of sensitivity. It is not popularly used by researchers but is helpful in few domains. It is the ratio of the total number of correctly classified negative samples to negative classes actually present in the confusion matrix as shown in Fig. 7.

Confusion matrix A confusion matrix is a table that is frequently used to evaluate a categorization model's (or "classifier's") effectiveness on a set of training test data values are known. While the confusion matrix Fig. 7. itself is rather straightforward to comprehend, the associated language might be perplexing.

TF-IDF Term Frequency refer to as number of times term present in a document. *TF* which counts the number of times a term word appears in the document Because each document is varied in length, it is likely that a term will appear far more frequently in longer documents than in shorter ones. As a result, the phrase frequency is frequently divided by the document length.

$$Term\ Frequency = \frac{(Number\ of\ times\ term\ t\ present\ in\ a\ document)}{(Total\ number\ of\ terms\ in\ the\ document)} \quad (3)$$

$$\text{Inverse Document Frequency (IDF)} = \frac{(\text{Total Number of document})}{(\text{number of terms } t \text{ in the document})} \quad (4)$$

$$TF - IDF = (TF) * (IDF) \quad (5)$$

6 Applications of sentiment analysis

Sentiment analysis has many applications, ranging from analyzing customer opinion, analyzing patient mental health status based on posts done on social media. Furthermore, technological advances such as Blockchain, IoT, Cloud Computing, and Big Data have broadened the range of applications for Sentiment Analysis, allowing it to be used in practically any discipline. Few most frequently used application in sentiment analysis shown in Fig. 8. A few significant domains and industries where Sentiment Analysis is applied are described below:

6.1 Business analysis

Sentiment analysis in the field of business intelligence offers several benefits. Additionally, firms can utilize sentiment analysis data to improve products, investigate client feedback and develop an innovative marketing strategy. The most typical use of sentiment analysis in the field of business intelligence is analyzing customers impressions of services or products. These studies, however, are not limited to product producers; consumers may use them to review items and make more informed decisions. Sentiment analysis in business intelligence has various benefits. For example, Businesses can use the results of Sentiment Analysis to make product enhancements, examine consumer feedback, or develop a new marketing plan (Han et al. 2019). Sentiment analysis is most frequently used in business intelligence to examine customers perceptions of products or services. However, these analyses are not limited to product producers; consumers may also use them to compare items and make a more informed choice. For eight years, (Bose et al. 2020) service food reviews on *amazon.com*. Emotion lexicon, which classifies them into eight different emotions and two moods (positive and negative). They found that sentiment analysis may be used to identify customer behaviours and hazards and increase customer satisfaction.

6.1.1 Product reviews

As the e-commerce business is burgeoning, so is the number of products sold and reviews given from the customers. Sentiment analysis one them will help customers choose better products (Paré 2003). Phrase level or aspect level (Schouten and Frasinca 2015) sentiment analysis performed on product reviews. Sentiment analysis can determine what the customer thinks about its latest product after launching or examining comments and reviews. Keywords for a specific product feature (food, service, cleanliness) can be chosen, and a sentiment analysis framework (Mackey et al. 2015) can be trained to identify and analyze only the necessary information.

6.1.2 Market research and competitor analysis

Market research is perhaps the most common sentiment analysis application, besides brand image monitoring and consumer opinion investigation. The purpose of sentiment analysis is to determine who is emerging among competitors and how marketing campaigns compare. It can be utilized to acquire a complete picture of a brand's and its competitors consumer base from the ground up. Sentiment analysis may collect data from several platforms Twitter, Facebook, blogs, deliver tangible results, and overcome difficulties in business intelligence.

6.2 Healthcare and medical domain

This is one of the industries where sentiment analysis is being utilized in recent times. Data can be collected from various sources like surveys, Twitter (Carvalho and Plastino 2021), blogs, news articles, reviews, etc. This data can then be analyzed for various use cases, one of them being an evaluation of standards and analysis of new updates in the medical field. Domain experts are researching actively to find more uses of sentiment analysis and other NLP applications (Ebadi et al. 2021). This application helps healthcare service providers collect and evaluate patient moods, epidemics, adverse drug reactions, and diseases to improve healthcare services. In work of Jiménez-Zafra et al. (2019) pointed out the difficulties in applying sentiment analysis in health care because of the specific and unique terminologies used in the domain. In work of Clark et al. (2018) used Twitter tweets concerning patient's experiences as an add-on to analyze public health. Over a year, they generated roughly five million breast cancer-related tweets using Twitter's Streaming API. After pre-processing, the tweets were classified with a standard LR classifier and a CNN model. Positive treatment experiences, rallying support, and expanding public awareness were all linked. In conclusion, applying sentiment analysis to analyze patient-generated data on social media can help determine patients' needs and views.

6.2.1 Reputation management

The application of sentiment analysis in diverse markets is brand monitoring and reputation management. Evaluating how customers view their brand, product, or service is beneficial to fashion companies, marketing agencies, IT companies, hotel chains, media channels, and other businesses. Sentiment analysis tool adds more variety and intelligence to the brand's and their products portrayal. It enables businesses to track how their customers perceive their brands and highlight the precise data about their attitudes. Look for trends and changes, and pay attention to influencers presentations. Altogether, sentiment analysis can be utilized in automating the media surveillance system as well as the alarm system that goes with it. Keep track of the brand's discussions and ratings on various social media platforms.

6.3 Review analysis

Sentiment analysis is extensively used in the domain of Entertainment. Reviews of movie, shows, and short films may be analyzed to determine the viewer's response (Kumar et al. 2019). This not only helps viewers make a better choice but also helps good contents gain popularity. Sentence level (Lin and He 2009) Sentiment Analysis has commonly used in this domain to determine the overall sentiment of the reviews given accurately. The travel industry has sought to improve client experiences by developing machine learning and online consumer recommendation systems based on intelligent, data-driven decision-making techniques (Jain et al. 2021f) also discussed categorizing consumer decisions as positive or negative based on online reviews provided by the valuable consumer (Jain et al. 2021e).

6.3.1 Customer reviews

Sentimental analysis on reviews on hotels and restaurants can help customers choose better and also help the owners improve. Aspect-based sentiment analysis done on hotels and restaurants will help identify the aspect with the most positive reviews and negative reviews, on which Hotels can work and make it better. (Sann and Lai 2020; Al-Smadi et al. 2018) According to sentiment analysis, this is one of the most attractive industries. Sentimental analysis on reviews on hotels and restaurants can help customers choose better and also help the owners improve (Zhao et al. 2019). *ABSA* (Akhtar et al. 2017) done on hotels and restaurants will help identify the aspect with the most positive reviews and negative reviews, on which hotels can work and make it better. The service providers profit the most since they may extract the aspect that receives the most negative feedback and improve on it.

6.3.2 Aspect analysis

Aspect-based sentiment analysis can help businesses make the most use of the massive amounts of data they create. The aspect-based method will enable companies to extract the most important aspects of client feedback and service.

6.4 Stock market

One of the applications of sentiment analysis is stock price prediction. It can be done by analyzing all the news about the stock market and predicting the stock price trends. Data can be collected from various sources like Twitter, news articles, blogs, etc. Sentence level sentiment analysis can be done on these texts, after which the overall polarity of texts will be decided of news of a particular company. In work of Xing et al. (2018) used to determine whether the trend will be rising or decreasing. Positive news tended to lead to an upward trend, whereas negative news tended to lead to a downward trend. Bitcoin and other digital cryptocurrencies relate to a novel technology known as Blockchain. Participants inside the blockchain network verify the digital transactions using peer to peer consensus methods. However, investigations which apply Sentiment Analysis towards the area of blockchain technology are still infrequent, and those that do exist, such as work in Kraaijeveld and De Smedt (2020), have employed sentiment

Table 7 Study of various challenges in sentiment analysis as per structure of data

References	Challenge	Structure of review
Heerschop et al. (2011)	Negation	Semi structure data(adjective)
Lapponi et al. (2012)	Negation + huge lexicon	Semi Structured data
Remus (2013)	Negation + bi polar words	Semi structured data , sentences or topic documents
Chunping et al. (2014)	Domain decencies	Unstructured data, Emotion reviews
Hu et al. (2014)	Short abbreviations	Unstructured data
Kiritchenko et al. (2014)	Emotions	unstructured data
Cao et al. (2015)	Ambiguity	Semi-structured data
Flekova et al. (2015)	Bi-polar words	Unstructured data
Kasmuri and Basiron (2017)	Opinion Spam Detections	Structured data
Shayaa et al. (2018)	Incomplete information and Sparsity	Unstructured data and Structured data
Salah et al. (2019) and Mite-Baidal et al. (2018)	Domain dependency	Unstructured data
Oueslati et al. (2020)	Language dependency	Unstructured data
Zuo et al. (2020)	Implicit language	Unstructured data
Acheampong et al. (2020)	Emotion detection	unstructured data
Kumar and Garg (2020)	Multi-source data fusion	Unstructured data, Structured data
Parvin et al. (2021)	Informal speech, and slangs	Unstructured data
Mendon et al. (2021)	Multiple class classifications	Unstructured data, Structured data

analysis to anticipate the value of digital cryptocurrencies. In the work of Rognone et al. (2020) investigated the influence of news sentiment on cryptocurrencies like bitcoin and other standard currencies volatility, volume, and returns.

6.5 Voice of customers

Take all user feedback from the call centres, emails, surveys, chats, and web and combine and assess it. Sentiment analysis will allow categorizing and organizing data in order to detect trends and reoccurring issues and worries. Sentiment analysis may aid in identifying an appropriate customer group and subsequent value proposition development, both of which are essential components of a successful business operation. On the other hand, to stay updated and maintain the product in demand, it must have the finger on the pulse of its customers.

6.6 Social media monitoring

Sentiment analysis of social data will monitor client sentiment 24 hours a day, seven days a week, in real-time when anything unpleasant starts to circulate, which can rapidly reply and bolster image when getting favourable mentions. That also obtains consistent, reliable information on clients, which can track progress from season to season for the decision-making process. Because individuals provide their comments without

being asked, social media posts frequently present some of the most honest points of view regarding products, services, and enterprises. They are obliged to express their feelings to the rest of the world.

7 Challenges in sentiment analysis

Sentiment analysis comes with various challenges ranging from computational cost to informal writing and the presence of variations in languages. We look at the sentiment analysis challenges that occur more frequently with certain types of sentiment structure, as shown in Table 7. Few significant challenges faced in sentiment analysis are:

7.1 Structured sentiments

Structured sentiments are found in formal sentiment reviews, they are more focused on formal problems such as books or research. Because the authors are professionals, they are capable of writing thoughts or observations concerning scientific or factual concerns.

7.2 Semi-structured sentiments

Semi-Structured Sentiments fall between structured and unstructured sentiments. These require an awareness of numerous review-related concerns. This style, which is dependent on benefits and drawbacks, is listed separately by the authors, and the pros and cons sections are typically comprised of brief sentences (Birjali et al. 2021; Hussein 2018; Ebrahimi et al. 2017; Mohammad 2017).

7.3 Unstructured sentiment

Unstructured Sentiment is an informal and free-flowing writing type in which the writer is not constrained by any rules (Mukherjee et al. 2013). The text may comprise multiple sentences, each of which could potentially include both pros and cons. For example, unstructured reviews offer more opinion information than their formal counterparts (Levashina et al. 2014). A feature explicitly stated: If a feature occurs in a review sentence's segment/chunk, the feature is referred to as an explicit feature of the product. For instance, in the segment, the image is marvelous. The image is an explicit feature. If a feature f is not explicitly mentioned in the review section but is implied, it is referred to as an implicit feature of the product (Liu et al. 2010; Elith et al. 2011). For instance, in the section, it is extremely pricey, and expensive is a feature sign. In light of the critical nature of sentiment analysis, this study examines the relationship between respondents perspective structures and sentiment analysis issues.

7.4 Methodological challenges

The majority of sentiment analysis in the modern day is data-driven machine learning models adapting a sentiment analysis algorithm developed for product evaluations to evaluate microblog postings is an unanswered question. Additionally, how to deal with ambiguous situations and irony are key difficulties in sentiment analysis. For instance, a sarcastic remark about an object is intended to communicate a negative sentiment; yet, conventional sentiment analysis algorithms frequently miss this meaning. Numerous methods have been proposed (Castro et al. 2019; Medhat et al. 2014) for detecting sarcasm in language. However, the problem is far from resolved, as comedy is very culturally particular, and it is challenging for a machine to understand unique (and frequently fairly detailed) cultural allusions. In the work of Poria et al. (2018a) suggest by incorporating vocal and facial expressions into multimodal sentiment analysis; This can improve its success rate in identifying sarcastic comments. Furthermore, individuals express sentiment for social reasons unrelated to their fundamental dispositions. For instance, a person may transmit positive or negative thoughts to adhere to a specific topic A norm or express and define one's identity. Finally, machine-based sentiment analysis is confined to outward expressions of sentiment, and conclusive information about an individual expressed ideas is lacking.

Sentiment analysis is applicable to different types of data, each of which presents particular challenges. Sentiment analysis of human to machine and human to human interactions requires very similar datasets to those used for emotion recognition. As a result, it has the same limitations in terms of size and unreliable ground truth. In the work of McDuff et al. (2014) have illustrated how webcams may be used to collect a large number of emotional reactions, including sentiment. While this degrades the audiovisual capture quality, it achieves a scale that is not conceivable in the laboratory. Additionally, there is the issue of labeling confidential laboratory data, which prohibits those permitted to examine the data from performing the time-consuming operation of labeling. As a result, they are restricted in terms of the amount of data they can collect in the laboratory and our ability to label huge volumes of data. There are several methods for assessing feelings, but word embedding algorithms such as word2vec and GloVe turn words into meaningful vectors. These methods, on the other hand, ignore the word's sentiment information (Wankhade et al. 2021).

Multimedia information on websites is the second source of multi-modal sentiment data. Social media provides us with a wealth of data that helps us to scale. The issue is that the data acquired vary in terms of quality and context, and the data is limited to specific populations that are more prevalent on the internet. However, because the data is publicly available, crowd sourcing may be utilized to categorize it easily. According to the available data on MSA, people are more prone to communicate positive or negative ideas online, resulting in a scarcity of neutral opinions represented in all MSA studies evaluated.

Sarcasm People tend to use sarcasm when they do not meet their expectations. It is very tough for machines to pick up sarcasm as many factors affect sarcasm, such as tone, situation, background information, etc. Sarcasm is a satirical remark that may look like praising but in reality. Sarcasm is used by people to criticize. Sarcasm is a type of sentiment in which people express implicit information, usually the polar opposite of the message content, in order to emotionally hurt someone or mock something. Sarcasm detection in text mining is one of the most challenging tasks in NLP, but it has lately become an interesting

research subject due to its usefulness in enhancing social media sentiment analysis (Eke et al. 2020).

Informal style of writing Informal style of writing is the biggest challenge to all NLP tasks, including sentiment analysis. People are very casual about writing reviews or texts; they tend to use acronyms, emojis, shortcuts in their text which is very hard to pick up. Acronyms can be handled if they are universal. There are a lot of regional acronyms¹⁴ which change and grow day by day.

Grammatical errors Grammatical errors are very common in informal texts and can be handled, but only to some extent; spelling errors can also be corrected limited. It is very difficult to burgeoning the spelling mistake of users uniquely every time. The accuracy of sentiment analysis and NLP tasks may be improved if these errors can be handled and corrected.

Computational cost To get better accuracy, we need to increase the training data size and complicate the model, which will exponentially increase the computational cost of the model for training; high-end GPU may be required to train a model with a huge corpus. Models like SVM, NB are not computationally costly, but neural networks and attention models have shown that they are computationally costly.

Availability of data As NLP and sentiment analysis is a recently boomed technology, the Availability of data may also be a challenge in some cases. Although data is available in Twitter for sentiment analysis, high-quality training data is challenging for supervised learning algorithms. Training data for ABSA is challenging to find online therefore needs to be prepared manually. The training data of one domain may not be applicable and valuable to other domains. For instance, a model trained on a hotel review dataset is not helpful in predicting sentiments of a stock or mutual fund dataset and vice versa.

Adaptations of language Languages change as they move to different regions and places; although the base language remains the same, many factors influence language, such as language prominence, pronunciation, literacy rate, etc. For instance, consider English language, which is widely spoken worldwide, but it is seen that many English varieties are spoken worldwide based on the regions like Indian, American, British, etc. Lots of words are used differently depending upon the region there are used. For instance, consider the word “*thong*” which means flip-flops or slippers in Australia but means undergarments in the UK. Similarly, different spellings for the same word, such as “*color*” and “*colour*,” mean the same but are spelled differently in different regions. This will create duplicates and may affect the accuracy and computational cost of the model. Language barrier is the hardest of the challenges to NLP. There are thousands of languages spoken worldwide, although NLP techniques are hardly available to 5-10 languages, and resources are widely available for English.

Phrases containing degree adverbs and intensifiers Adverbs such as slightly, barely, and moderately are used to quantify the sentiments. For instance, consider review r_1 = “The food is barely good” and r_2 = “the food is really good”. r_1 is considered neutral or slightly positive, whereas r_2 is considered to be highly positive. The adverbs ‘*barely*’ and ‘*really*’ decide the extent of positiveness and the word ‘*good*’. Similarly, intensifiers also quantify the sentiment of the sentences. Intensifiers like very, too are used to increase the positiveness or negativeness of the token. For instance, “*too good*” is considered to be more positive than “*good*.” Intensifiers and degree adverbs impose a challenge on aggregating the

¹⁴ <https://www.dictionary.com/e/acronyms>.

sentiment values and comparing two sentences of the same sentiment rather than differentiating between two sentences of opposite polarity.

The theoretical challenges employ a variety of approaches to enhance performance when answering the particular sentiment challenges (Hunter et al. 2012). The theoretical kind makes extensive PoS tagging and lexicon-based approaches (Taboada et al. 2011). The second approach is the BoW technique (El-Din 2016). Finally, there is the ME approach. However, the most frequently used technique is the N-gram technique, which is based on phrases and expressions when it comes to technical sentiment challenges (Wilson et al. 2009). As well as the method that is used the least is the lexicon-based approach.

Mixed Code Data Code-mixing is the employment of vocabulary and grammar from different languages in same sentence (Pravalika et al. 2017; Poria et al. 2020). Code Mixing is a linguistic phenomenon that can occur in a multilingual situation where speakers speak multiple languages. This phenomenon is becoming increasingly common as communication between groups of people who speak different languages grows. Code-Mixing: A review of Facebook posts created by Hindi-English users revealed a high code-mixing level in the posts. The problems in the Hindi-English code-mixed text were reported using a PoS tag annotated corpus. (Vijay et al. 2018) a system that can detect the language of the words, normalize them to their standard forms, assign their PoS tag, and split them into chunks to handle the problem of shallow parsing of Hindi-English code-mixed social media content. It's frequent in multilingual societies and presents considerable difficulty to NLP tasks like sentiment analysis. The lack of a formal grammar for code-mixed phrases makes it challenging to identify compositional semantics, which is critical for conducting sentiment analysis using rule-based and machine learning-based techniques. Furthermore, because mixing is up to the individual, there are no predetermined mixing guidelines, which is one of the significant drawbacks (Chatterjere et al. 2020). As a result, in order to conduct sentiment analysis on code-mixed data, new language models must be developed. In work of Chatterjere et al. (2020) and Singh et al. (2018) language modeling challenge for code-mixed Hinglish text was investigated. However, despite the fact that Code Mixing is a significant concern, few research has addressed it as thoroughly as the study of Lal et al. (2019) (English-Hindi) code-mixed data for sentiment analysis, the authors presented a hybrid architecture.

8 Conclusion

This article discussed sentiment analysis and associated techniques. The primary objective of this work is to investigate and complete classification methods with their advantage and disadvantages in sentiment analysis. To begin, several levels of sentiment analysis were discussed, followed by a quick overview of necessary procedures such as data collection and feature selection. Next, methods of sentiment categorization systems were classified and compared in terms of their advantages and disadvantages. Due to their simplicity and excellent accuracy, supervised machine learning methods are often the widely utilized technique in this discipline. Classification using NB and SVM algorithms are commonly used as benchmarks against which newly proposed approaches can be compared. Several of the most common application areas are discussed then the survey examines the significance and consequences of sentiment analysis challenges in sentiment evaluation. The comparison investigates the relationship between the structure of sentiment reviews and the difficulties associated with sentiment analysis. This comparison reveals domain dependence,

which is essential for identifying sentiment issues. The future work will consist of continuously expanding the comparison area with additional findings. The subsequent challenges illustrate that sentiment analysis is still a relatively unexplored subject of study.

References

- Abid F, Alam M, Yasir M, Li C (2019) Sentiment analysis through recurrent variants latterly on convolutional neural network of twitter. *Futur Gener Comput Syst* 95:292–308
- Acheampong FA, Wenyu C, Nunoo-Mensah H (2020) Text-based emotion detection: advances, challenges, and opportunities. *Eng Rep* 2(7):e12189
- Acheampong FA, Nunoo-Mensah H, Chen W (2021) Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artif Intell Rev* 54:5789–5829
- Adomavicius G, Kwon Y (2011) Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Trans Knowl Data Eng* 24(5):896–911
- Ahmad S, Asghar MZ, Alotaibi FM, Awan I (2019) Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Hum Centric Comput Inf Sci* 9(1):1–23
- Ahmad SR, Bakar AA, Yaakub MR (2019) A review of feature selection techniques in sentiment analysis. *Intell Data Anal* 23(1):159–189
- Akhtar MS, Ekbal A, Cambria E (2020) How intense are you? predicting intensities of emotions and sentiments using stacked ensemble [application notes]. *IEEE Comput Intell Mag* 15(1):64–75
- Akhtar N, Zubair N, Kumar A, Ahmad T (2017) Aspect based sentiment oriented summarization of hotel reviews. *Procedia Comput Sci* 115:563–571
- Al Amrani Y, Lazaar M, El Kadiri KE (2018) Random forest and support vector machine based hybrid approach to sentiment analysis. *Procedia Comput Sci* 127:511–520
- Al-Smadi M, Qawasmeh O, Al-Ayyoub M, Jararweh Y, Gupta B (2018) Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *J Comput Sci* 27:386–393
- Alhumoud SO, Al Wazrah AA (2021) Arabic sentiment analysis using recurrent neural networks: a review. *Artif Intell Rev* 55:707–748
- Ali SM, Noorian Z, Bagheri E, Ding C, Al-Obeidat F (2020) Topic and sentiment aware microblog summarization for twitter. *J Intell Inf Syst* 54(1):129–156
- Annett M, Kondrak G (2008) A comparison of sentiment analysis techniques: Polarizing movie blogs. In: *Conference of the Canadian Society for Computational Studies of Intelligence*. Springer, pp 25–35
- Arora A, Chakraborty P, Bhatia M, Mittal P (2021) Role of emotion in excessive use of twitter during COVID-19 imposed lockdown in India. *J Technol Behav Sci* 6(2):370–377
- Baashar Y, Alhussian H, Patel A, Alkawsy G, Alzaharani AI, Alfarraj O, Hayder G (2020) Customer relationship management systems (CRMS) in the healthcare environment: a systematic literature review. *Comput Stand Interfaces* 71:103442
- Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. *Lrec* 2010:2200–2204
- Bai X (2011) Predicting consumer sentiments from online text. *Decis Support Syst* 50(4):732–742
- Bai X, Liu P, Zhang Y (2020) Investigating typed syntactic dependencies for targeted sentiment classification using graph attention neural network. *IEEE/ACM Trans Audio Speech Lang Process* 29:503–514
- Balaji T, Annavarapu CSR, Bablani A (2021) Machine learning algorithms for social media analysis: a survey. *Comput Sci Rev* 40:100395
- Bandara K, Bergmeir C, Smyl S (2020) Forecasting across time series databases using recurrent neural networks on groups of similar series: a clustering approach. *Expert Syst Appl* 140:112896
- Bartusiak R, Augustyniak L, Kajdanowicz T, Kazienko P (2015) Sentiment analysis for polish using transfer learning approach. In: *2015 second european network intelligence conference*. IEEE, pp 53–59
- Basiri ME, Nemati S, Abdar M, Cambria E, Acharya UR (2021) ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Futur Gener Comput Syst* 115:279–294
- Behdenna S, Barigou F, Belalem G (2018) Document level sentiment analysis: a survey. *EAI Endorsed Trans Context-aware Syst Appl* 4(13):e2
- Bergsma S, McNamee P, Bagdouri M, Fink C, Wilson T (2012) Language identification for creating language-specific twitter collections. In: *Proceedings of the second workshop on language in social media*, pp 65–74

- Bhaskar J, Sruthi K, Nedungadi P (2015) Hybrid approach for emotion classification of audio conversation based on text and speech mining. *Procedia Comput Sci* 46:635–643
- Bhatia P, Ji Y, Eisenstein J (2015) Better document-level sentiment analysis from rst discourse parsing. arXiv preprint [arXiv:150901599](https://arxiv.org/abs/150901599)
- Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends. *Knowl-Based Syst* 226:107134
- Bojanowski P, Grave E, Joulin A, Mikolov T (2017) Enriching word vectors with subword information. *Trans Assoc Comput Linguist* 5:135–146
- Bordes A, Glorot X, Weston J, Bengio Y (2014) A semantic matching energy function for learning with multi-relational data. *Mach Learn* 94(2):233–259
- Borg A, Boldt M (2020) Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Syst Appl* 162:113746
- Bose R, Dey RK, Roy S, Sarddar D (2020) Sentiment analysis on online product reviews. In: *Information and communication technology for sustainable development*. Springer, pp 559–569
- Buder J, Rabi L, Feiks M, Badermann M, Zurstiege G (2021) Does negatively toned language use on social media lead to attitude polarization? *Comput Hum Behav* 116:106663
- Calders T, Verwer S (2010) Three naive bayes approaches for discrimination-free classification. *Data Min Knowl Disc* 21(2):277–292
- Cambria E, Das D, Bandyopadhyay S, Feraco A (2017) Affective computing and sentiment analysis. In: *A practical guide to sentiment analysis*. Springer, pp 1–10
- Cambria E, Li Y, Xing FZ, Poria S, Kwok K (2020) Senticnet 6: ensemble application of symbolic and subsymbolic ai for sentiment analysis. In: *Proceedings of the 29th ACM international conference on information & knowledge management*, pp 105–114
- Cao Q, Duan W, Gan Q (2011) Exploring determinants of voting for the “helpfulness” of online user reviews: a text mining approach. *Decis Support Syst* 50(2):511–521
- Cao Y, Zhang P, Xiong A (2015) Sentiment analysis based on expanded aspect and polarity-ambiguous word lexicon. *Int J Adv Comput Sci Appl* 6(2):97–103
- Carvalho J, Plastino A (2021) On the evaluation and combination of state-of-the-art features in twitter sentiment analysis. *Artif Intell Rev* 54(3):1887–1936
- Castro S, Hazarika D, Pérez-Rosas V, Zimmermann R, Mihalcea R, Poria S (2019) Towards multimodal sarcasm detection (an _obviously_ perfect paper). arXiv preprint [arXiv:190601815](https://arxiv.org/abs/190601815)
- Celik Y, Talo M, Yildirim O, Karabatak M, Acharya UR (2020) Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. *Pattern Recogn Lett* 133:232–239
- Chang JR, Liang HY, Chen LS, Chang CW (2020) Novel feature selection approaches for improving the performance of sentiment classification. *J Ambient Intell Humaniz Comput* pp 1–14
- Chatterjere A, Guptha V, Chopra P, Das A (2020) Minority positive sampling for switching points-an anecdote for the code-mixing language modeling. In: *Proceedings of the 12th language resources and evaluation conference*, pp 6228–6236
- Chen CC, Tseng YD (2011) Quality evaluation of product reviews using an information quality framework. *Decis Support Syst* 50(4):755–768
- Chen X, Wang Y, Liu Q (2017) Visual and textual sentiment analysis using deep fusion convolutional neural networks. In: *2017 IEEE international conference on image processing (ICIP)*. IEEE, pp 1557–1561
- Cheng Y, Yao L, Xiang G, Zhang G, Tang T, Zhong L (2020) Text sentiment orientation analysis based on multi-channel CNN and bidirectional GRU with attention mechanism. *IEEE Access* 8:134964–134975
- Chetviorkin I, Loukachevitch N (2012) Extraction of Russian sentiment lexicon for product meta-domain. In: *Proceedings of COLING 2012*, pp 593–610
- Chiew KL, Tan CL, Wong K, Yong KS, Tiong WK (2019) A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Inf Sci* 484:153–166
- Cho H, Lee H (2019) Biomedical named entity recognition using deep neural networks with contextual information. *BMC Bioinform* 20(1):1–11
- Chunping O, Wen Z, Ying Y, Zhiming L, Xiaohua Y (2014) Topic sentiment analysis in Chinese news. *Int J Multimed Ubiquitous Eng* 9(11):385–396
- Clark EM, James T, Jones CA, Alapati A, Ukandu P, Danforth CM, Dodds PS (2018) A sentiment analysis of breast cancer treatment experiences and healthcare perceptions across twitter. arXiv preprint [arXiv:180509959](https://arxiv.org/abs/180509959)
- Cortis K, Davis B (2021) Over a decade of social opinion mining: a systematic review. *Artif Intell Rev* 54:4873–4965

- Crawford M, Khoshgoftaar TM, Prusa JD, Richter AN, Al Najada H (2015) Survey of review spam detection using machine learning techniques. *J Big Data* 2(1):1–24
- Das H, Naik B, Behera H (2020) A java algorithm based wrapper method for optimal feature selection in supervised classification. *J King Saud Univ Comput Inf Sci*
- Dave K, Lawrence S, Pennock DM (2003) Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In: *Proceedings of the 12th international conference on World Wide Web*, pp 519–528
- Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:181004805*
- Donkers T, Loepp B, Ziegler J (2017) Sequential user-based recurrent neural network recommendations. In: *Proceedings of the eleventh ACM conference on recommender systems*, pp 152–160
- Dragoni M, Petrucci G (2017) A neural word embeddings approach for multi-domain sentiment analysis. *IEEE Trans Affect Comput* 8(4):457–470
- Duric A, Song F (2012) Feature selection for sentiment analysis based on content and syntax models. *Decis Support Syst* 53(4):704–711
- Ebadi A, Xi P, Tremblay S, Spencer B, Pall R, Wong A (2021) Understanding the temporal evolution of covid-19 research through machine learning and natural language processing. *Scientometrics* 126(1):725–739
- Ebrahimi M, Yazdavar AH, Sheth A (2017) Challenges of sentiment analysis for dynamic events. *IEEE Intell Syst* 32(5):70–75
- Eke CI, Norman AA, Shuib L, Nweke HF (2020) Sarcasm identification in textual data: systematic review, research challenges and open directions. *Artif Intell Rev* 53(6):4215–4258
- El-Din DM (2016) Enhancement bag-of-words model for solving the challenges of sentiment analysis. *Int J Adv Comput Sci Appl* 7(1):244–252
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of maxent for ecologists. *Divers Distrib* 17(1):43–57
- Ethayarajh K (2019) How contextual are contextualized word representations? Comparing the geometry of BERT, ELMO, and GPT-2 embeddings. *arXiv preprint arXiv:190900512*
- Fan TK, Chang CH (2011) Blogger-centric contextual advertising. *Expert Syst Appl* 38(3):1777–1788
- Fang Z, Zhang Q, Tang X, Wang A, Baron C (2020) An implicit opinion analysis model based on feature-based implicit opinion patterns. *Artif Intell Rev* 53(6):4547–4574
- Ferrari A, Esuli A (2019) An NLP approach for cross-domain ambiguity detection in requirements engineering. *Autom Softw Eng* 26(3):559–598
- Filatova E (2012) Irony and sarcasm: corpus generation and analysis using crowdsourcing. In: *Lrec, Citeseer*, pp 392–398
- Flek L (2020) Returning the N to NLP: towards contextually personalized classification models. In: *Proceedings of the 58th annual meeting of the association for computational linguistics*, pp 7828–7838
- Flekova L, Preotiu-Pietro D, Ruppert E (2015) Analysing domain suitability of a sentiment lexicon by identifying distributionally bipolar words. In: *Proceedings of the 6th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pp 77–84
- Fredriksen-Goldsen KI, Kim HJ (2017) The science of conducting research with LGBT older adults—an introduction to aging with pride: National health, aging, and sexuality/gender study (NHAS)
- Gao Z, Feng A, Song X, Wu X (2019) Target-dependent sentiment classification with BERT. *IEEE Access* 7:154290–154299
- George DR, Rovniak LS, Kraschnewski JL (2013) Dangers and opportunities for social media in medicine. *Clin Obstet Gynecol* 56(3)
- Ghazi D, Inkpen D, Szpakowicz S (2015) Detecting emotion stimuli in emotion-bearing sentences. In: *International conference on intelligent text processing and computational linguistics*. Springer, pp 152–165
- Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53(3–4):325–338
- Hailong Z, Wenyan G, Bo J (2014) Machine learning and lexicon based methods for sentiment classification: a survey. In: *2014 11th web information system and application conference*. IEEE, pp 262–265
- Hajek P, Barushka A, Munk M (2020) Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Comput Appl* 32(23):17259–17274
- Hamdan H, Bellot P, Bechet F (2015) Lsislif: Crf and logistic regression for opinion target extraction and sentiment polarity analysis. In: *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, pp 753–758
- Han K, Xiao A, Wu E, Guo J, Xu C, Wang Y (2021) Transformer in transformer. *arXiv preprint arXiv:210300112*

- Han T, Liu C, Yang W, Jiang D (2019) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. *Knowl-Based Syst* 165:474–487
- Hangya V, Farkas R (2017) A comparative empirical study on social media sentiment analysis over various genres and languages. *Artif Intell Rev* 47(4):485–505
- Hassonah MA, Al-Sayyed R, Rodan A, Ala'M AZ, Aljarah I, Faris H (2020) An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on twitter. *Knowl-Based Syst* 192:105353
- Hassonah MA, Al-Sayyed R, Rodan A, Ala'M AZ, Aljarah I, Faris H (2020) An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on twitter. *Knowl-Based Syst* 192:105353
- Heerschop B, van Iterson P, Hogenboom A, Frasinca F, Kaymak U (2011) Accounting for negation in sentiment analysis. In: 11th Dutch-Belgian information retrieval workshop (DIR 2011), Citeseer, pp 38–39
- Herscovich D, Donatelli L (2021) It's the meaning that counts: the state of the art in NLP and semantics. *KI-Künstliche Intelligenz* pp 1–16
- Ho C, Murad MAA, Doraisamy S, Kadir RA (2014) Extracting lexical and phrasal paraphrases: a review of the literature. *Artif Intell Rev* 42(4):851–894
- Hssina B, Merbouha A, Ezzikouri H, Erritali M (2014) A comparative study of decision tree id3 and c4. 5. *Int J Adv Comput Sci Appl* 4(2):13–19
- Hu N, Bose I, Koh NS, Liu L (2012) Manipulation of online reviews: an analysis of ratings, readability, and sentiments. *Decis Support Syst* 52(3):674–684
- Hu X, Tang J, Gao H, Liu H (2014) Social spammer detection with sentiment information. In: 2014 IEEE international conference on data mining. IEEE, pp 180–189
- Hunter ST, Cushenbery L, Friedrich T (2012) Hiring an innovative workforce: a necessary yet uniquely challenging endeavor. *Hum Resour Manag Rev* 22(4):303–322
- Hussein DMEDM (2018) A survey on sentiment analysis challenges. *J King Saud Univ Eng Sci* 30(4):330–338
- Imani MB, Keyvanpour MR, Azmi R (2013) A novel embedded feature selection method: a comparative study in the application of text categorization. *Appl Artif Intell* 27(5):408–427
- Jain PK, Pamula R, Ansari S, Sharma D, Maddala L (2019) Airline recommendation prediction using customer generated feedback data. In: 2019 4th international conference on information systems and computer networks (ISCON). IEEE, pp 376–379
- Jain PK, Pamula R, Ansari S (2021) A supervised machine learning approach for the credibility assessment of user-generated content. *Wirel Pers Commun* 118(4):2469–2485
- Jain PK, Pamula R, Srivastava G (2021) A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Comput Sci Rev* 41:100413
- Jain PK, Pamula R, Yekun EA (2021c) A multi-label ensemble predicting model to service recommendation from social media contents. *J Supercomput* 1–18
- Jain PK, Quamer W, Pamula R, Saravanan V (2021d) Spsan: sparse self-attentive network-based aspect-aware model for sentiment analysis. *J Ambient Intell Humaniz Comput* 1–18
- Jain PK, Saravanan V, Pamula R (2021) A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. *Trans Asian Low-Resour Lang Inf Process* 20(5):1–15
- Jain PK, Yekun EA, Pamula R, Srivastava G (2021) Consumer recommendation prediction in online reviews using cuckoo optimized machine learning models. *Comput Electr Eng* 95:107397
- Janjua F, Masood A, Abbas H, Rashid I, Khan MMZM (2021) Textual analysis of traitor-based dataset through semi supervised machine learning. *Futur Gener Comput Syst* 125:652–660
- Jiménez-Zafra SM, Martín-Valdivia MT, Molina-González MD, Ureña-López LA (2019) How do we talk about doctors and drugs? sentiment analysis in forums expressing opinions for medical domain. *Artif Intell Med* 93:50–57
- Juraska J, Walker M (2021) Attention is indeed all you need: semantically attention-guided decoding for data-to-text nlg. *arXiv preprint [arXiv:210907043](https://arxiv.org/abs/210907043)*
- Kaity M, Balakrishnan V (2020) Sentiment lexicons and non-english languages: a survey. *Knowl Inf Syst* 1–36
- Kamal A (2013) Subjectivity classification using machine learning techniques for mining feature-opinion pairs from web opinion sources. *arXiv preprint [arXiv:13126962](https://arxiv.org/abs/13126962)*
- Kanapala A, Pal S, Pamula R (2019) Text summarization from legal documents: a survey. *Artif Intell Rev* 51(3):371–402
- Kang H, Yoo SJ, Han D (2012) Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Syst Appl* 39(5):6000–6010

- Kasmuri E, Basiron H (2017) Subjectivity analysis in opinion mining—a systematic literature review. *Int J Adv Soft Comput Appl* 9(3):133–159
- Kaufmann M (2012) Jmaxalign: a maximum entropy parallel sentence alignment tool. In: *Proceedings of COLING 2012: demonstration papers*, pp 277–288
- Khairnar J, Kinikar M (2013) Machine learning algorithms for opinion mining and sentiment classification. *Int J Sci Res Publ* 3(6):1–6
- Khan MT, Durrani M, Ali A, Inayat I, Khalid S, Khan KH (2016) Sentiment analysis and the complex natural language. *Complex Adapt Syst Model* 4(1):1–19
- Kiritchenko S, Zhu X, Mohammad SM (2014) Sentiment analysis of short informal texts. *J Artif Intell Res* 50:723–762
- Kitaev N, Klein D (2018) Constituency parsing with a self-attentive encoder. *arXiv preprint [arXiv:1805.1052](https://arxiv.org/abs/1805.1052)*
- Kolchyna O, Souza TT, Treleven P, Aste T (2015) Twitter sentiment analysis: Lexicon method, machine learning method and their combination. *arXiv preprint [arXiv:150700955](https://arxiv.org/abs/150700955)*
- Korkontzelos I, Nikfarjam A, Shardlow M, Sarker A, Ananiadou S, Gonzalez GH (2016) Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. *J Biomed Inform* 62:148–158
- Kosamkar V, Chaudhari SS (2013) Improved intrusion detection system using c4. 5 decision tree and support vector machine. PhD diss, Doctoral dissertation, Mumbai University
- Kraaijeveld O, De Smedt J (2020) The predictive power of public twitter sentiment for forecasting cryptocurrency prices. *J Int Finan Markets Inst Money* 65:101188
- Kumar A, Garg G (2020) Systematic literature review on context-based sentiment analysis in social multimedia. *Multimed Tools Appl* 79(21):15349–15380
- Kumar A, Teeja MS (2012) Sentiment analysis: a perspective on its past, present and future. *Int J Intell Syst Appl* 4(10):1
- Kumar KN, Uma V (2021) Intelligent sentiment-based lexicon for context-aware sentiment analysis: optimized neural network for sentiment classification on social media. *J Supercomput* 77:12801–12825
- Kumar S, Yadava M, Roy PP (2019) Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Inf Fusion* 52:41–52
- Lakkaraju H, Socher R, Manning C (2014) Aspect specific sentiment analysis using hierarchical deep learning. In: *NIPS Workshop on deep learning and representation learning*, pp 1–9
- Lal YK, Kumar V, Dhar M, Shrivastava M, Koehn P (2019) De-mixing sentiment from code-mixed text. In: *Proceedings of the 57th annual meeting of the association for computational linguistics: student research workshop*, pp 371–377
- Lapponi E, Read J, Øvrelid L (2012) Representing and resolving negation for sentiment analysis. In: *2012 IEEE 12th international conference on data mining workshops. IEEE*, pp 687–692
- Lata K, Singh P, Dutta K (2020) A comprehensive review on feature set used for anaphora resolution. *Artif Intell Rev* 54:2917–3006
- Levashina J, Hartwell CJ, Morgeson FP, Campion MA (2014) The structured employment interview: narrative and quantitative review of the research literature. *Pers Psychol* 67(1):241–293
- Li F, Wang W, Xu J, Yi J, Wang Q (2019) Comparative study on vulnerability assessment for urban buried gas pipeline network based on SVM and ANN methods. *Process Saf Environ Prot* 122:23–32
- Li X, Bing L, Zhang W, Lam W (2019b) Exploiting BERT for end-to-end aspect-based sentiment analysis. *arXiv preprint [arXiv:191000883](https://arxiv.org/abs/191000883)*
- Li YM, Li TY (2013) Deriving market intelligence from microblogs. *Decis Support Syst* 55(1):206–217
- Ligthart A, Catal C, Tekinerdogan B (2021) Systematic reviews in sentiment analysis: a tertiary study. *Artif Intell Rev* 54:4997–5053
- Lin C, He Y (2009) Joint sentiment/topic model for sentiment analysis. In: *Proceedings of the 18th ACM conference on information and knowledge management*, pp 375–384
- Ling M, Chen Q, Sun Q, Jia Y (2020) Hybrid neural network for Sina Weibo sentiment analysis. *IEEE Trans Comput Soc Syst* 7(4):983–990
- Liu B (2012) Sentiment analysis and opinion mining. *Synth Lect Hum Lang Technol* 5(1):1–167
- Liu B, Zhang L (2012) A survey of opinion mining and sentiment analysis. In: Aggarwal C, Zhai C (eds) *Mining text data*. Springer, Boston, pp 415–463
- Liu B et al (2010) Sentiment analysis and subjectivity. *Handb Nat Lang Process* 2(2010):627–666
- Liu P, Qiu X, Huang X (2016) Recurrent neural network for text classification with multi-task learning. *arXiv preprint [arXiv:160505101](https://arxiv.org/abs/160505101)*
- Lu B, Ott M, Cardie C, Tsou BK (2011) Multi-aspect sentiment analysis with topic models. In: *2011 IEEE 11th international conference on data mining workshops. IEEE*, pp 81–88

- Mackey TK, Miner A, Cuomo RE (2015) Exploring the e-cigarette e-commerce marketplace: identifying internet e-cigarette marketing characteristics and regulatory gaps. *Drug Alcohol Depend* 156:97–103
- Majumder N, Hazarika D, Gelbukh A, Cambria E, Poria S (2018) Multimodal sentiment analysis using hierarchical fusion with context modeling. *Knowl-Based Syst* 161:124–133
- Maks I, Vossen P (2012) A lexicon model for deep sentiment analysis and opinion mining applications. *Decis Support Syst* 53(4):680–688
- McDuff D, El Kaliouby R, Cohn JF, Picard RW (2014) Predicting ad liking and purchase intent: large-scale analysis of facial responses to ads. *IEEE Trans Affect Comput* 6(3):223–235
- Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J* 5(4):1093–1113
- Mendon S, Dutta P, Behl A, Lessmann S (2021) A hybrid approach of machine learning and lexicons to sentiment analysis: enhanced insights from twitter data of natural disasters. *Inf Syst Front* 23:1145–1168
- Meng J, Long Y, Yu Y, Zhao D, Liu S (2019) Cross-domain text sentiment analysis based on cnn_ft method. *Information* 10(5):162
- Mezquita Y, Alonso RS, Casado-Vara R, Prieto J, Corchado JM (2020) A review of k-nn algorithm based on classical and quantum machine learning. In: *International symposium on distributed computing and artificial intelligence*. Springer, pp 189–198
- Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. arXiv preprint [arXiv:13013781](https://arxiv.org/abs/1301.3781)
- Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2018) Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform* 19(6):1236–1246
- Mite-Baidal K, Delgado-Vera C, Solís-Avilés E, Espinoza AH, Ortiz-Zambrano J, Varela-Tapia E (2018) Sentiment analysis in education domain: a systematic literature review. In: *International conference on technologies and innovation*. Springer, pp 285–297
- Mohammad SM (2017) Challenges in sentiment analysis. In: *A practical guide to sentiment analysis*. Springer, pp 61–83
- Moraes R, Valiati JF, Neto WPG (2013) Document-level sentiment classification: an empirical comparison between SVM and ANN. *Expert Syst Appl* 40(2):621–633
- Moraes R, Valiati JF, Neto WPG (2013) Document-level sentiment classification: an empirical comparison between SVM and ANN. *Expert Syst Appl* 40(2):621–633
- Moreo A, Romero M, Castro J, Zurita JM (2012) Lexicon-based comments-oriented news sentiment analyzer system. *Expert Syst Appl* 39(10):9166–9180
- Mowlaei ME, Abadeh MS, Keshavarz H (2020) Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Syst Appl* 148:113234
- Mukherjee A, Venkataraman V, Liu B, Glance N (2013) What yelp fake review filter might be doing? In: *Proceedings of the international AAAI conference on web and social media*, vol 7
- Naseem U, Razzak I, Musial K, Imran M (2020) Transformer based deep intelligent contextual embedding for twitter sentiment analysis. *Futur Gener Comput Syst* 113:58–69
- Ortigosa-Hernández J, Rodríguez JD, Alzate L, Lucania M, Inza I, Lozano JA (2012) Approaching sentiment analysis by using semi-supervised learning of multi-dimensional classifiers. *Neurocomputing* 92:98–115
- Oueslati O, Cambria E, HajHmida MB, Ounelli H (2020) A review of sentiment analysis research in Arabic language. *Futur Gener Comput Syst* 112:408–430
- Paré DJ (2003) Does this site deliver? B2B e-commerce services for developing countries. *Inf Soc* 19(2):123–134
- Park HW, Park S, Chong M (2020) Conversations and medical news frames on twitter: infodemiological study on covid-19 in South Korea. *J Med Internet Res* 22(5):e18897
- Park S, Kim Y (2016) Building thesaurus lexicon using dictionary-based approach for sentiment classification. In: *2016 IEEE 14th international conference on software engineering research, management and applications (SERA)*, pp 39–44. <https://doi.org/10.1109/SERA.2016.7516126>
- Parvin SA, Sumathi M, Mohan C (2021) Challenges of sentiment analysis—a survey. In: *2021 5th International conference on trends in electronics and informatics (ICOEI)*. IEEE, pp 781–786
- Patel HH, Prajapati P (2018) Study and analysis of decision tree based classification algorithms. *Int J Comput Sci Eng* 6(10):74–78
- Patil N, Lathi R, Chitre V (2012) Customer card classification based on c5. 0 & cart algorithms. *Int J Eng Res Appl* 2(4):164–167
- Peng M, Zhang Q, Jiang Yg, Huang XJ (2018) Cross-domain sentiment classification with target domain specific information. In: *Proceedings of the 56th annual meeting of the association for computational linguistics (Volume 1: Long Papers)*, pp 2505–2513

- Peng Y, Yan S, Lu Z (2019) Transfer learning in biomedical natural language processing: an evaluation of BERT and ELMo on ten benchmarking datasets. arXiv preprint [arXiv:190605474](https://arxiv.org/abs/190605474)
- Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
- Pham TH, Le-Hong P (2017) End-to-end recurrent neural network models for vietnamese named entity recognition: word-level vs. character-level. In: International conference of the Pacific Association for Computational Linguistics. Springer, pp 219–232
- Piryani R, Madhavi D, Singh VK (2017) Analytical mapping of opinion mining and sentiment analysis research during 2000–2015. *Inf Process Manag* 53(1):122–150
- Plank B, Søgaard A, Goldberg Y (2016) Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss. arXiv preprint [arXiv:160405529](https://arxiv.org/abs/160405529)
- Poria S, Cambria E, Winterstein G, Huang GB (2014) Sentic patterns: dependency-based rules for concept-level sentiment analysis. *Knowl-Based Syst* 69:45–63
- Poria S, Chaturvedi I, Cambria E, Hussain A (2016) Convolutional MKL based multimodal emotion recognition and sentiment analysis. In: 2016 IEEE 16th international conference on data mining (ICDM). IEEE, pp 439–448
- Poria S, Cambria E, Hazarika D, Mazumder N, Zadeh A, Morency LP (2017) Multi-level multiple attentions for contextual multimodal sentiment analysis. In: 2017 IEEE international conference on data mining (ICDM). IEEE, pp 1033–1038
- Poria S, Hussain A, Cambria E (2018a) Combining textual clues with audio-visual information for multimodal sentiment analysis. In: *Multimodal sentiment analysis*. Springer, pp 153–178
- Poria S, Majumder N, Hazarika D, Cambria E, Gelbukh A, Hussain A (2018) Multimodal sentiment analysis: addressing key issues and setting up the baselines. *IEEE Intell Syst* 33(6):17–25
- Poria S, Hazarika D, Majumder N, Mihalcea R (2020) Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. *IEEE Trans Affect Comput*
- Pravalika A, Oza V, Meghana N, Kamath SS (2017) Domain-specific sentiment analysis approaches for code-mixed social network data. In: 2017 8th international conference on computing, communication and networking technologies (ICCCNT). IEEE, pp 1–6
- Qiu G, He X, Zhang F, Shi Y, Bu J, Chen C (2010) DASA: dissatisfaction-oriented advertising based on sentiment analysis. *Expert Syst Appl* 37(9):6182–6191
- Quinlan JR (2014) C4. 5: programs for machine learning. Elsevier, Amsterdam
- Rana TA, Cheah YN (2016) Aspect extraction in sentiment analysis: comparative analysis and survey. *Artif Intell Rev* 46(4):459–483
- Rao D, Ravichandran D (2009) Semi-supervised polarity lexicon induction. In: Proceedings of the 12th conference of the European chapter of the ACL (EACL 2009), pp 675–682
- Rao G, Huang W, Feng Z, Cong Q (2018) LSTM with sentence representations for document-level sentiment classification. *Neurocomputing* 308:49–57
- Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowl-Based Syst* 89:14–46
- Razon A, Barnden J (2015) A new approach to automated text readability classification based on concept indexing with integrated part-of-speech n-gram features. In: Proceedings of the international conference recent advances in natural language processing, pp 521–528
- Remus R (2013) Modeling and representing negation in data-driven machine learning-based sentiment analysis. In: ESSEM@ AI* IA, pp 22–33
- Revathy R, Lawrance R (2017) Comparative analysis of c4. 5 and c5. 0 algorithms on crop pest data. *Int J Innovative Res Comput Commun Eng* 5(1):50–58
- Ritter A, Etzioni O, Clark S (2012) Open domain event extraction from twitter. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 1104–1112
- Rizos G, Hemker K, Schuller B (2019) Augment to prevent: short-text data augmentation in deep learning for hate-speech classification. In: Proceedings of the 28th ACM international conference on information and knowledge management, pp 991–1000
- Rognone L, Hyde S, Zhang SS (2020) News sentiment in the cryptocurrency market: an empirical comparison with forex. *Int Rev Financ Anal* 69:101462
- Ruffer N, Knitza J, Krusche M (2020) # Covid4Rheum: an analytical twitter study in the time of the COVID-19 pandemic. *Rheumatol Int* 40(12):2031–2037
- Rui H, Liu Y, Whinston A (2013) Whose and what chatter matters? The effect of tweets on movie sales. *Decis Support Syst* 55(4):863–870
- Salah Z, Al-Ghuwairi ARF, Baarah A, Aloqaily A, Qadoui B, Alhayek M, Alhijawi B (2019) A systematic review on opinion mining and sentiment analysis in social media. *Int J Bus Inf Syst* 31(4):530–554

- Sánchez-Rada JF, Iglesias CA (2019) Social context in sentiment analysis: formal definition, overview of current trends and framework for comparison. *Inf Fusion* 52:344–356
- Sann R, Lai PC (2020) Understanding homophily of service failure within the hotel guest cycle: applying NLP-aspect-based sentiment analysis to the hospitality industry. *Int J Hosp Manag* 91:102678
- Saunders D (2021) Domain adaptation for neural machine translation. PhD thesis, University of Cambridge
- Schouten K, Frasincar F (2015) Survey on aspect-level sentiment analysis. *IEEE Trans Knowl Data Eng* 28(3):813–830
- Sharma A, Lyons J, Dehzingi A, Paliwal KK (2013) A feature extraction technique using bi-gram probabilities of position specific scoring matrix for protein fold recognition. *J Theor Biol* 320:41–46
- Shayaa S, Jaafar NI, Bahri S, Sulaiman A, Wai PS, Chung YW, Piprani AZ, Al-Garadi MA (2018) Sentiment analysis of big data: methods, applications, and open challenges. *IEEE Access* 6:37807–37827
- Singh JP, Irani S, Rana NP, Dwivedi YK, Saumya S, Roy PK (2017) Predicting the “helpfulness” of online consumer reviews. *J Bus Res* 70:346–355
- Singh K, Sen I, Kumaraguru P (2018) A twitter corpus for Hindi-English code mixed POS tagging. In: *Proceedings of the sixth international workshop on natural language processing for social media*, pp 12–17
- Singh M, Jakhar AK, Pandey S (2021) Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Soc Netw Anal Min* 11(1):1–11
- Singh RK, Sachan MK, Patel R (2021) 360 degree view of cross-domain opinion classification: a survey. *Artif Intell Rev* 54(2):1385–1506
- Singh S, Gupta P (2014) Comparative study id3, cart and c4. 5 decision tree algorithm: a survey. *Int J Adv Inf Sci Technol* 27(27):97–103
- Socher R, Perelygin A, Wu J, Chuang J, Manning CD, Ng AY, Potts C (2013) Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the 2013 conference on empirical methods in natural language processing*, pp 1631–1642
- Soleymani M, Garcia D, Jou B, Schuller B, Chang SF, Pantic M (2017) A survey of multimodal sentiment analysis. *Image Vis Comput* 65:3–14
- Stappen L, Schuller B, Lefter I, Cambria E, Kompatsiaris I (2020) Summary of MuSe 2020: multimodal sentiment analysis, emotion-target engagement and trustworthiness detection in real-life media. In: *Proceedings of the 28th ACM international conference on multimedia*, pp 4769–4770
- Straka M, Hajic J, Straková J (2016) UDPipe: trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, pos tagging and parsing. In: *Proceedings of the tenth international conference on language resources and evaluation (LREC'16)*, pp 4290–4297
- Subhashini L, Li Y, Zhang J, Atukorale AS, Wu Y (2021) Mining and classifying customer reviews: a survey. *Artif Intell Rev* 54:6343–6389
- Sun C, Huang L, Qiu X (2019) Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. *arXiv preprint arXiv:190309588*
- Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. *Comput Linguist* 37(2):267–307
- Thet TT, Na JC, Khoo CS (2010) Aspect-based sentiment analysis of movie reviews on discussion boards. *J Inf Sci* 36(6):823–848
- Tian Y, Galery T, Dulcinati G, Molimpakis E, Sun C (2017) Facebook sentiment: reactions and emojis. In: *Proceedings of the fifth international workshop on natural language processing for social media*, pp 11–16
- Tran T, Ba H, Huynh VN (2019) Measuring hotel review sentiment: an aspect-based sentiment analysis approach. In: *International symposium on integrated uncertainty in knowledge modelling and decision making*. Springer, pp 393–405
- Tripathy A, Agrawal A, Rath SK (2015) Classification of sentimental reviews using machine learning techniques. *Procedia Comput Sci* 57:821–829
- Tubishat M, Idris N, Abushariah MA (2018) Implicit aspect extraction in sentiment analysis: review, taxonomy, opportunities, and open challenges. *Inf Process Manag* 54(4):545–563
- Turney PD, Littman ML (2003) Measuring praise and criticism: inference of semantic orientation from association. *ACM Trans Inf Syst* 21(4):315–346
- Uysal AK, Murphey YL (2017) Sentiment classification: feature selection based approaches versus deep learning. In: *2017 IEEE international conference on computer and information technology (CIT)*. IEEE, pp 23–30
- Valdivia A, Luzión MV, Herrera F (2017) Neutrality in the sentiment analysis problem based on fuzzy majority. In: *2017 IEEE international conference on fuzzy systems (FUZZ-IEEE)*. IEEE, pp 1–6
- Valdivia A, Luzón MV, Cambria E, Herrera F (2018) Consensus vote models for detecting and filtering neutrality in sentiment analysis. *Inf Fusion* 44:126–135

- Valencia F, Gómez-Espinoza A, Valdés-Aguirre B (2019) Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy* 21(6):589
- Van de Camp M, Van den Bosch A (2012) The socialist network. *Decis Support Syst* 53(4):761–769
- Varelas G, Voutsakis E, Raftopoulou P, Petrakis EG, Milios EE (2005) Semantic similarity methods in wordnet and their application to information retrieval on the web. In: Proceedings of the 7th annual ACM international workshop on Web information and data management, pp 10–16
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. arXiv preprint [arXiv:170603762](https://arxiv.org/abs/1706.03762)
- Vateekul P, Koomsubha T (2016) A study of sentiment analysis using deep learning techniques on Thai twitter data. In: 2016 13th international joint conference on computer science and software engineering (JCSSE). IEEE, pp 1–6
- Vechtomova O (2017) Disambiguating context-dependent polarity of words: an information retrieval approach. *Inf Process Manag* 53(5):1062–1079
- Venugopalan M, Gupta D (2015) Exploring sentiment analysis on twitter data. In: 2015 eighth international conference on contemporary computing (IC3). IEEE, pp 241–247
- Vijay D, Bohra A, Singh V, Akhtar SS, Shrivastava M (2018) Corpus creation and emotion prediction for Hindi-English code-mixed social media text. In: Proceedings of the 2018 conference of the North American chapter of the Association for Computational Linguistics: student research workshop, pp 128–135
- Wadawadagi R, Pagi V (2020) Sentiment analysis with deep neural networks: comparative study and performance assessment. *Artif Intell Rev* 53:6155–6195
- Wang G, Sun J, Ma J, Xu K, Gu J (2014) Sentiment classification: the contribution of ensemble learning. *Decis Support Syst* 57:77–93
- Wang Z, Ho SB, Cambria E (2020) Multi-level fine-scaled sentiment sensing with ambivalence handling. *Int J Uncertain Fuzziness Knowl-Based Syst* 28(04):683–697
- Wankhade M, Annavarapu CSR, Verma MK (2021) CBVoSD: context based vectors over sentiment domain ensemble model for review classification. *J Supercomput* 1–37
- Weerasooriya T, Perera N, Liyanage S (2016) A method to extract essential keywords from a tweet using NLP tools. In: 2016 sixteenth international conference on advances in ICT for emerging regions (ICTer). IEEE, pp 29–34
- Wilson T, Wiebe J, Hoffmann P (2009) Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis. *Comput Linguist* 35(3):399–433
- Wu D, Chi M (2017) Long short-term memory with quadratic connections in recursive neural networks for representing compositional semantics. *IEEE Access* 5:16077–16083
- Wu P, Li X, Shen S, He D (2020) Social media opinion summarization using emotion cognition and convolutional neural networks. *Int J Inf Manag* 51:101978
- Xia H, Yang Y, Pan X, Zhang Z, An W (2020) Sentiment analysis for online reviews using conditional random fields and support vector machines. *Electron Commer Res* 20(2):343–360
- Xia Y, Cambria E, Hussain A, Zhao H (2015) Word polarity disambiguation using Bayesian model and opinion-level features. *Cognit Comput* 7(3):369–380
- Xing FZ, Cambria E, Welsch RE (2018) Natural language based financial forecasting: a survey. *Artif Intell Rev* 50(1):49–73
- Yadav A, Vishwakarma DK (2020) Sentiment analysis using deep learning architectures: a review. *Artif Intell Rev* 53(6):4335–4385
- Yan-Yan Z, Bing Q, Ting L (2010) Integrating intra-and inter-document evidences for improving sentence sentiment classification. *Acta Autom Sinica* 36(10):1417–1425
- Yang B, Cardie C (2014) Context-aware learning for sentence-level sentiment analysis with posterior regularization. In: Proceedings of the 52nd annual meeting of the association for computational linguistics (Volume 1: Long Papers), pp 325–335
- Yaseen Q et al (2021) Spam email detection using deep learning techniques. *Procedia Comput Sci* 184:853–858
- Yousif A, Niu Z, Tarus JK, Ahmad A (2019) A survey on sentiment analysis of scientific citations. *Artif Intell Rev* 52(3):1805–1838
- Yuan Z, Wu S, Wu F, Liu J, Huang Y (2018) Domain attention model for multi-domain sentiment classification. *Knowl-Based Syst* 155:1–10
- Yue L, Chen W, Li X, Zuo W, Yin M (2019) A survey of sentiment analysis in social media. *Knowl Inf Syst* 60(2):617–663
- Zhang Z, Wang L, Zou Y, Gan C (2018) The optimally designed dynamic memory networks for targeted sentiment classification. *Neurocomputing* 309:36–45

- Zhao W, Guan Z, Chen L, He X, Cai D, Wang B, Wang Q (2017) Weakly-supervised deep embedding for product review sentiment analysis. *IEEE Trans Knowl Data Eng* 30(1):185–197
- Zhao Y, Xu X, Wang M (2019) Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *Int J Hosp Manag* 76:111–121
- Zhu X, Goldberg AB (2009) Introduction to semi-supervised learning. *Synth Lect Artif Intell Mach Learn* 3(1):1–130
- Zuo E, Zhao H, Chen B, Chen Q (2020) Context-specific heterogeneous graph convolutional network for implicit sentiment analysis. *IEEE Access* 8:37967–37975
- Zvarevashe K, Olugbara OO (2018) A framework for sentiment analysis with opinion mining of hotel reviews. In: 2018 Conference on information communications technology and society (ICTAS). IEEE, pp 1–4

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.