*Journal of Computer Science and Cybernetics, V.37, N.4 (2021), 403–428*
DOI 10.15625/1813-9663/37/4/15892

# SENTIMENT ANALYSIS FOR OPINIONS ON SOCIAL MEDIA: A SURVEY

HUYEN TRANG PHAN[1], NGOC THANH NGUYEN[2*], DOSAM HWANG[1]

[1] *Department of Computer Engineering, Yeungnam University, Gyeongsan, South Korea*

[2] *Department of Applied Informatics, Wroclaw University of Science and Technology, Wroclaw, Poland*

**Abstract.** With the rapid development of the Internet industry, an increasing number of social media platforms have been developed. These social media platforms have become the main channels for communication among most users. Opinions from social media platforms provide the most updated and inclusive information. Sentiments from opinions are a valuable data source for solving many issues. Therefore, sentiment analysis has developed into one of the most popular natural language processing fields. Hence, improving the performance of sentiment analysis methods or discovering new problems related to these methods is essential. In this context, we must be aware of the general information relevant to this area. This survey presents a summary of the necessary stages for building a complete model to be used in sentiment analysis. For each stage, we list the popular techniques that have been widely used in recent years. In addition, discussions and comparisons related to these methods are provided. Additionally, we discuss the challenges and possible research directions for future research in this field.

**Keywords.** Sentiment analysis; Sentiment classification; Types of sentiment analysis; Challenges in sentiment analysis.

## 1. INTRODUCTION

Social media is a platform for sharing data among users, whose number is increasing continuously. Two types of social media platforms exist, i.e., social networks and online communities [30]. Social networks allow individuals to provide and share content with society as well as interact with and reach out to more people. Users' opinions related to entities are one of the contents that are shared most frequently. Users' opinions are typically short texts posted on social networks. These short texts may contain a user's sentiment toward an entity. User sentiment is a user's view of or a user's attitude toward a situation or event (known as a topic) [14].

---

Dedicated to Professor Phan Dinh Dieu on the occasion of his 85th birth anniversary.

---

*Corresponding author.
*E-mail addresses*: huyentrangtin@gmail.com (H.T. Phan); Ngoc-Thanh.Nguyen@pwr.edu.pl (N.T. Nguyen); dshwang@yu.ac.kr (D. Hwang).

Sentiment analysis, which is a research area in natural language processing, categorizes sentiments expressed over a particular topic. Sentiment analysis for opinions on social media has been widely used in various real applications, such as in market price prediction, political elections and box office prediction, movie sales estimation, and decision-making systems. For example, the authors in [99] evaluated the online opinions related to body wash products. Then, They used aspect-level sentiment analysis to extract aspects, which customers are most not satisfied aiming to improve these products. Karyotis et al. [36] applied opinions sentiment analysis of users to extract aspects satisfied user needs aiming to improve products market-driven. Phan et al. [58] introduced a novel method to support decision-making based on assessing the aspect's effectiveness on user satisfaction using aspect-level sentiment analysis. We can see sentiment analysis for opinions being used in many real applications. Therefore, the performance of the sentiment analysis methods strongly affects the quality of these applications. The increase in opinions available in social networks renders sentiment analysis more crucial. Recent studies [24, 35, 55, 60, 62] show that tweets and human sentiment analysis can be useful for predicting crimes, stock markets, election votes, disaster management, etc. In addition, researchers have used tweets as the data source to analyze human sentiment related to the Covid-19 [72, 73, 86] as well as to detect false information [18, 67]. Therefore, a significant amount of research related to sentiment analysis, such as new approaches and performance improvement, has been published annually. Accordingly, researchers have attempted to summarize recent research developments and directions in sentiment analysis. In [30, 68], the authors present a complete review of opinion mining and sentiment analysis to classify existing methods and compare their advantages and disadvantages. The current challenges and solutions should be understood to clarify future directions for sentiment analysis. In addition, a survey paper by Tedmori *et al.* [79] provides a comprehensive overview of sentiment analysis by focusing on the main tasks and applications of sentiment analysis. This paper provides references to new studies, including a detailed discussion of four main sentiment analysis tasks and an overview of diverse sentiment analysis applications. Additionally, Yue *et al.* [96] reviewed sentiment analysis researches by considering multiple perspectives, such as task, granularity, and methodology. They also discussed different data and tools used for sentiment analysis and their strengths and drawbacks. Subsequently, the authors provided challenges and discussions for possible extensions for multimodal sentiment analysis in the future. Chakraborty *et al.* [14] provided the evolution of sentiment analysis from the explosion of data on social media. The authors presented the typical process of crawling data from social media over the years and detected the similarity based on similar choices of users in social networks. Data types that were used the most frequently in recent times were assessed and discussed in different forms. In addition, the authors classified and compared the methods for analyzing sentiments. Finally, the authors presented the limitations of sentiment analysis techniques, demonstrating that those limitations will enable better investigations in the future.

This paper provides a general overview of the necessary stages for building a complete model for opinions sentiment analysis. It differs from other related publications in that a general architecture for opinions sentiment analysis is presented. Subsequently, to analyze a specific stage in architecture, the newest methods are listed and their advantages and disadvantages compared. Publications pertaining to sentiment analysis are reviewed and the methods and techniques reported are compared in terms of their strengths and drawbacks.

Various sentiment analysis methods have been categorized based on different orientations. The contributions of this survey are as follows: (i) Publications pertaining to sentiment analysis are reviewed, and the strengths and drawbacks of the methods presented are explained. Various sentiment analysis methods are categorized based on different orientations. (ii) The types of tools, lexicons, and available benchmark datasets that can be used in sentiment analysis research and their limitations are discussed and categorized based on their usage in specific applications. (iii) The essential prospects and challenges of sentiment analysis are identified and discussed.

The remainder of this paper is organized as follows. In Section 2, we present the background concepts related to sentiment analysis. The general architecture of the sentiment analysis process is described in Section 3. In Section 4, we present an analysis of the most recent investigations pertaining to the applications of sentiment analysis methods in terms of data, techniques, levels, and applications. In Section 5, we discuss the challenges that may exist in the sentiment analysis field. The conclusions are presented in Section 6.

## 2. BACKGROUND CONCEPTS

### 2.1. Opinion definition

In [42], the authors defined an opinion as a quadruple, $\langle g, s, h, t \rangle$, where sentiment target ($g$) is the entity or a part or an attribute of the entity that the sentiment has been expressed upon; the sentiment ($s$) of the opinion regarding the target ($g$) can be a *strong positive, positive, neutral, negative, strong negative* or a rating (e.g., 1–5 stars); opinion holder ($h$) is the person or organization who holds this opinion; and time ($t$) is the time when an opinion is expressed.

**Example 1.** Assume that we have an opinion as follows: "Posted by: Amanda Ivan; Date: November 09, 2020; "I bought a Galaxy note 10-20 days ago (1). I am so satisfied with it (2). The selfie quality is amazing (3). The battery is also good (4). However, I think it is too heavy for me (5)."

From this opinion, some important information can be obtained, as follows:

- Sentiment target: (1) (2), and (5): Galaxy note 10; (3): selfie quality; (4): battery.
- Sentiment toward target: (2): positive; (3): positive; (4): positive; (5): negative.
- Opinion holder: From (1) to (5): Amanda Ivan.
- Time: November 09, 2020.

From the definition of opinion presented in [42], we constructed another definition of opinion by focusing on the main information related to the objective of sentiment analysis. Definitions regarding opinion, topic, and aspect were detailed in [57, 58], and they are described represented briefly as follows.

**Definition 1.** An opinion can be described by a quadruple, $\langle p_{i\ell}, a_{ij}, s_{ij}, s_i \rangle$, where $p_{i\ell}$ is the topic that holds the sentiment appearing in the given opinion, $a_{ij}$ is an aspect of topic $p_{i\ell}$; $s_{ij}$ is the user's sentiment toward aspect $a_{ij}$, and $s_i$ is the user's sentiment toward topic $p_{i\ell}$.

**Example 2.** Using the Definition 1 for Example 1, the following were obtained:

- $p_{i\ell}$: Galaxy note 10.
- $a_{i1}$: selfie quality, $a_{i2}$: battery, $a_{i3}$: weight.

- $s_{i1}$: positive, $s_{i2}$: positive, $s_{i3}$: negative.
- $s_i$: positive.

**Definition 2.** (Topic and aspect [58]) A topic can be any entity, such as a product, service, organization, event, or person. The topic is often a noun or noun phrase, and it is typically described in more detail by clauses or sentences regarding objects pertaining to the topic. A topic can exhibit the aspects mentioned in the opinion. An aspect is a specific detail, a subpart, or an attribute of the topic. The aspect is typically a noun or a noun phrase that is complemented by a phrase, clause, or sentence containing a user's emotion.

### 2.2. Sentiment definition

**Definition 3.** Sentiment is the feeling, attitude, evaluation, or emotion of users toward specific aspects of topics or for the topics. This implies that $s = \{\textit{strong positive}, \textit{positive}, \textit{neutral}, \textit{negative}, \textit{strong negative}\}$ is a set of sentiment orientations $\Rightarrow s_{ij}$ and $s_i \in s$.

**Definition 4.** Sentiment analysis is a process used to determine the sentiment orientation in opinions. Hence, assuming that we have an opinion, one of the following main tasks is performed in sentiment analysis:

- Task 1: Identify all the topics in the opinion. Subsequently, group synonymous topics into the same topic cluster. Finally, extract a unique topic for each cluster.
- Task 2: Identify all aspects of the topics in this opinion Subsequently, group the synonymous aspects into the same aspect cluster. Finally, extract a unique aspect corresponding to each cluster.
- Task 3: Determine the sentiment orientation of an opinion related to a topic.
- Task 4: Determine the sentiment orientation of opinions toward the aspect of a specific topic.
- Task 5: Extract the reasons why users express this sentiment orientation for a specific topic.
- Task 6: Extract the reasons why users express this sentiment orientation toward an aspect of a topic.

### 2.3. Levels of sentiment analysis

Assume that we have a finite set of opinions, $\mathcal{T}$,representing the opinions of users regarding a specific topic (denoted by $p_{il}$), where $\mathcal{T}$ is represented by $\mathcal{T} = \{t_1, t_2, ..., t_n\}$, and $n$ is the number of gathered opinions. Let $\mathcal{P}t$ be a set of positive words, and $\mathcal{N}t$ be a set of negative words. Each opinion must be separated into a set of tokens to determine the necessary elements in the opinions. For $t_i \in \mathcal{T}$: Let $\mathcal{W}_i = \{w_1, w_2, ..., w_g\}$ be a set of words in opinion $t_i$. In this paper, we divide levels of sentiment analysis into three main categories as follows.

#### 2.3.1. Aspect-level sentiment analysis

*Aspect-level sentiment analysis* is a method based on determining the polarity of each word and phrase in a lexicon [96]. A lexicon is a vocabulary of sentiment words with sentiment polarity, strength value, or sentiment score. We described concepts related to aspect-level sentiment analysis in detail in our previous researches, namely [58, 59, 60]. In this study, we briefly represent such as the following definitions.

**Definition 5.** A sentiment relation [58] between words $w_k$ and $w_h$ ($w_k \in \mathcal{W}_i, w_h \in \mathcal{P}t \cup \mathcal{N}t$), denoted by $\Theta$, is defined by the following function

$$\Theta(w_h, w_k) = \begin{cases} 1, & \text{if } w_k \text{ related to } w_h \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

A topic can contain many aspects. For $w_k \in \mathcal{W}_i$: Let $\mathcal{A}_i = \{a_1, a_2, ..., a_m\}$ be a set of aspects related to specific topic $p_{il}$ existing in $\mathcal{T}$, where $a_{ij}$ is an aspect that is assigned a sentiment and is related to the selected topic.

**Definition 6.** Aspect $a_{ij}$ of specified topic $p_{il}$ in opinion $t_i$ [58] is a token $w_k \in \mathcal{W}_i$ that satisfies two conditions simultaneously: $w_k$ must be a noun or noun phrase, and it must be related to at least one sentiment word existing in this tweet. Aspect $a_{ij}$ is expressed as

$$a_{ij} = \{w_k | tag(w_k) = \text{`}\mathcal{NOUN}\text{'}, \exists w_h \in \mathcal{W}_i : \Theta(w_k, w_h) = 1\}. \tag{2}$$

Let $\mathcal{F}$ be a set of fundamental sentiment words; $\mathcal{F}s$ be a set of fuzzy semantic words; $\mathcal{F}p$ be a set of fuzzy sentiment phrases; and $\mathcal{C}p$ be a set of clear sentiment phrases.

- For $fs \in \mathcal{F}s$: let $\mathcal{S}c_{fs}$ be a sentiment score of $fs$.
- For $f \in \mathcal{F}$: let $\mathcal{S}c_f$ be a sentiment score of $f$.
- For $fp \in \mathcal{F}p$: let $\mathcal{S}c_{fp}$ be a sentiment score of $fp$.
- For $cp \in \mathcal{C}p$: let $\mathcal{S}c_{cp}$ be a sentiment score of $cp$.

In which, $\mathcal{S}c_f$, $\mathcal{S}c_{fs}$, $\mathcal{S}c_{fp}$, and $\mathcal{S}c_{cp}$ are computed as proposed by Phan *et al.* in the paper [63].

**Definition 7.** The sentiment score of clear sentiment phrase $cp$, denoted by $\mathcal{S}c_{cp}$, is determined based on the score of the fundamental sentiment words in this $cp$ [63]. The value of $\mathcal{S}c_{cp}$ is computed as follows

$$\mathcal{S}c_{cp} = (-1)^i((-1)^j \mathcal{S}c_f), \tag{3}$$

where

$$i = \begin{cases} 1 & \text{if } (\mathcal{S}c_f \leq 0) \\ 2 & \text{if } (\mathcal{S}c_f > 0), \end{cases} \tag{4}$$

$$j = \begin{cases} 1 & \text{if } (cp \text{ is a CSP of Type II}) \\ 2 & \text{if } (cp \text{ is a CSP of Type I}), \end{cases} \tag{5}$$

where CSP is the acronym for *"clear sentiment phrase"* described in [58].

**Definition 8.** The sentiment score of fuzzy sentiment phrase $fp$, denoted by $\mathcal{S}c_{fp}$, is identified based on the scores of the fundamental sentiment and fuzzy semantic words in the $fp$ [63]. The value of $fp$ is computed as follows

$$\mathcal{S}c_{fp} = (-1)^i((-1)^k \mathcal{S}c_f + (-1)^j \mathcal{S}c_{fs}), \tag{6}$$

where

$$i = \begin{cases} 1, & \text{if } (\mathcal{S}c_f \leq 0) \\ 2, & \text{if } (\mathcal{S}c_f > 0), \end{cases} \tag{7}$$

$$\mathcal{k} = \begin{cases} 1 & \text{if } (\mathit{fp} \text{ is a FSP of Type II}) \\ 2 & \text{if } (\mathit{fp} \text{ is a FSP of Type I}), \end{cases} \tag{8}$$

$$j = \begin{cases} 1 & \text{if } (\mathit{fs} \text{ is a diminisher word}) \\ 2 & \text{if } (\mathit{fs} \text{ is an intensifier word}), \end{cases} \tag{9}$$

where FSP is the acronym for *"fuzzy sentiment phrase"* described in detail in [63].

Therefore, for $w_{\mathcal{k}} \in \mathcal{F}p \cup \mathcal{C}p$, we obtain

$$\mathcal{S}c_{\mathcal{k}} = \begin{cases} \mathcal{S}c_{cp} & \text{if } (w_{\mathcal{k}} \in \mathcal{C}p) \\ \mathcal{S}c_{fp} & \text{if } (w_{\mathcal{k}} \in \mathcal{F}p). \end{cases} \tag{10}$$

**Definition 9.** For $a_{ij} \in \mathcal{A}_i$, the sentiment of aspect $a_{ij}$ in opinion $t_i$ is identified as follows

$$s_{ij} = \begin{cases} \textit{Strong positive}, & \text{if } (\exists w_{\mathcal{k}} : w_{\mathcal{k}} \in (\mathcal{F}p \cup \mathcal{C}p) \wedge \Theta(w_{\mathcal{k}}, a_{ij}) = 1) \wedge (0.6 < \mathcal{S}c_{\mathcal{k}} \leq 1) \\ \textit{Positive}, & \text{if } (\exists w_{\mathcal{k}} : w_{\mathcal{k}} \in (\mathcal{F}p \cup \mathcal{C}p) \wedge \Theta(w_{\mathcal{k}}, a_{ij}) = 1) \wedge (0.2 < \mathcal{S}c_{\mathcal{k}} \leq 0.6) \\ \textit{Neutral}, & \text{if } (\exists w_{\mathcal{k}} : w_{\mathcal{k}} \in (\mathcal{F}p \cup \mathcal{C}p) \wedge \Theta(w_{\mathcal{k}}, a_{ij}) = 1) \wedge (-0.2 \leq \mathcal{S}c_{\mathcal{k}} \leq 0.2) \\ \textit{Negative}, & \text{if } (\exists w_{\mathcal{k}} : w_{\mathcal{k}} \in (\mathcal{F}p \cup \mathcal{C}p) \wedge \Theta(w_{\mathcal{k}}, a_{ij}) = 1) \wedge (-0.6 \leq \mathcal{S}c_{\mathcal{k}} < -0.2) \\ \textit{Strong negative}, & \text{if } (\exists w_{\mathcal{k}} : w_{\mathcal{k}} \in (\mathcal{F}p \cup \mathcal{C}p) \wedge \Theta(w_{\mathcal{k}}, a_{ij}) = 1) \wedge (-1 \leq \mathcal{S}c_{\mathcal{k}} < -0.6). \end{cases} \tag{11}$$

Herein, we list the available lexical resources that are effective and widespread in the sentiment classification of short posts.

SentiWordNet: SentiWordNet [6] is a lexical resource devised to support word-based sentiment analysis applications. Each WordNet synset [44] in this lexicon is assigned to one of three sentiment scores: *positive, negative,* and *neutral.* Because this lexical resource provides a synset-based sentiment analysis, different synsets of the same term may reflect other sentiment scores.

WordNet-Affect: WordNet-Affect [75] is a linguistic resource for word-based sentiment analysis. It is an extension of WordNet, where a label is defined for each affective-related synset concept (e.g., the term "good" is labeled with the concept of positive emotion). Mapping is performed based on a domain-independent hierarchy of affective labels based on WordNet relationships.

MPQA: MPQA [89] is a lexicon of 8,222 terms obtained from various sources. These terms contain words and part-of-speech (POS) tags, labeled in terms of polarity (*positive, negative, neutral*) and intensity (*strong, weak*).

SenticNet: SenticNet [13] is a lexical resource for lexicon-level sentiment analysis. This lexicon is created based on sentic computing [12], a novel multidisciplinary paradigm for sentiment analysis. Unlike the previous lexicons, SenticNet can identify the sentiment score for complex concepts using the intensity of 16 basic emotions. Currently, SenticNet contains 14,000 common sense concepts, with a sentiment score in the range of [-1,1].

SentiStrength: SentiStrength [80] was developed by Thelwall *et al.* via determining the sentiment expressed in a range of recognized nonstandard spellings and other standard textual methods available on MySpace[1]. Subsequently, these authors published an extended

[1]https://myspace.com/

version of SentiStrength, namely SentiStrength viz, SentiStrength 2 [81]. This lexicon provides the sentiment strength, either positive or negative, for each opinion. Then, the average scores of sentiment strengths are calculated for each entity.

Semantics Orientation-Calculator (SO-CAL): The SO-CAL [78] provides the polarity and strength of words. In this lexicon, semantic orientation is determined based on the aggregate-and-average method; that is, the total score of all adjectives is divided by the total number of adjectives in the document.

### 2.3.2. Sentence-level sentiment analysis

*Sentence-level sentiment analysis* is a method based on determining the sentiment polarity of each sentence. This implies means that before analyzing a sentiment, sentences appearing in opinions must be separated.

Assuming that an opinion containing many sentences. Let $\mathcal{S}t_i = \{st_{1i}, st_{2i}, ..., st_{zi}\}$ be a set of sentences in opinion $t_i$. For $st_{\chi i} \in \mathcal{S}t_i$, let $se_{\chi i}$ be the sentiment expressed in sentence $st_{\chi i}$.

**Definition 10.** The sentiment score expressed in sentence $st_{\chi i}$ (denoted by $\mathcal{S}c_{\chi i}$) is computed based on the average sentiment score of fuzzy sentiment phrases (FSPs) and clear sentiment phrases (CSPs) contained in the sentence. The value of $\mathcal{S}c_{\chi i}$ is calculated as follows

$$\mathcal{S}c_{\chi i} = \frac{1}{m + \ell}(\sum_{\hbar=1}^{m} \mathcal{S}c_{\hbar} + \sum_{q=1}^{\ell} \mathcal{S}c_{q}), \tag{12}$$

where, $m$ and $\ell$ are the number of FSPs and CSPs, respectively, in sentence $st_{\chi i}$, $\mathcal{S}c_{\hbar}$ is the sentiment score of $\hbar$-th fuzzy sentiment phrase in $st_{\chi i}$, $\mathcal{S}c_{q}$ is the sentiment score of $q$-th fuzzy sentiment phrase in $st_{\chi i}$, and the method for calculating $\mathcal{S}c_{\hbar}$ and $\mathcal{S}c_{q}$ is described in [58].

**Definition 11.** The sentiment expressed in sentence $st_{\chi i}$ is determined based on the sentiment score expressed in this sentence as follows

$$se_{\chi i} = \begin{cases} \textit{Strong positive}, & \text{if } (0.6 < \mathcal{S}c_{\chi i} \leq 1) \\ \textit{Positive}, & \text{if } (0.2 < \mathcal{S}c_{\chi i} \leq 0.6) \\ \textit{Neutral}, & \text{if } (-0.2 \leq \mathcal{S}c_{\chi i} \leq 0.2) \\ \textit{Negative}, & \text{if } (-0.6 \leq \mathcal{S}c_{\chi i} < -0.2) \\ \textit{Strong negative}, & \text{if } (-1 \leq \mathcal{S}c_{\chi i} < -0.6). \end{cases} \tag{13}$$

### 2.3.3. Document-level sentiment analysis

*Document-level sentiment analysis* aims to arrange a feeling report by communicating a positive or negative conclusion or assessment. It considers all the sentence-level plans to place sentiment expressed in every sentence. Document- and sentence-level classifications do not differ significantly because sentences are merely short documents. In this paper, we consider an opinion corresponding to a short document. Therefore, in this study, document-level sentiment analysis is to determine the orientation sentiment expressed on the whole opinion.

**Definition 12.** The sentiment score expressed in opinion $t_i$ (denoted by $\mathcal{S}c_i$) is computed based on the average sentiment score expressed in sentences appearing in opinion $t_i$. The value of $\mathcal{S}c_i$ is expressed as follows

$$\mathcal{S}c_i = \frac{1}{z}\sum_{\chi=1}^{z} \mathcal{S}c_{\chi i}, \tag{14}$$

where $z$ is the number of sentences in opinion $t_i$.

**Definition 13.** The sentiment expressed in opinion $t_i$, denoted by $s_i$, is determined based on the sentiment score expressed in this document as follows

$$s_i = \begin{cases} Strong\ positive, & \text{if } (0.6 < \mathcal{S}c_i \leq 1) \\ Positive, & \text{if } (0.2 < \mathcal{S}c_i \leq 0.6) \\ Neutral, & \text{if } (-0.2 \leq \mathcal{S}c_i \leq 0.2) \\ Negative, & \text{if } (-0.6 \leq \mathcal{S}c_i < -0.2) \\ Strong\ negative, & \text{if } (-1 \leq \mathcal{S}c_i < -0.6). \end{cases} \tag{15}$$

## 3. GENERAL SENTIMENT ANALYSIS ARCHITECTURE

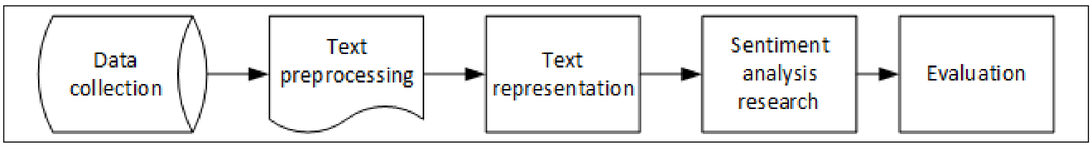


Figure 1: General sentiment analysis architecture

The general architecture of the sentiment analysis model is illustrated in Figure 1. Its components are as follows: First, the input of this model is typically user opinions. Second, opinions are preprocessed by correcting natural language errors and removing unnecessary or noise factors. Third, the opinions written in natural language are converted into numerical vectors. Fourth, the sentiment analysis model is built using existing classification algorithms, models, and methods. Fifth, vectors describing opinions are fed into the model created from the previous step for training. Finally, the results of the sentiment analysis model are evaluated using well-known metrics. If the model does not perform well, some parameters will be adjusted, e.g., the balance of label types in the data, the values of the parameters chosen when training the model. Otherwise, this model will be used to develop real applications. The details of these components are described as follows.

### 3.1. Data collection

Data collection is the primary and first procedure of the sentiment analysis model. Data may be collected by crawling data from sources such as social networks, online libraries, and journals, or by reusing available data published by other authors. Some social networks, such as Twitter and Sina Weibo, availed their application programming interface (API) to obtained public data from their sites. Twitter has provided some APIs such as Twitter REST API[2] to obtain static data such as user profile information and streaming API, and

[2]https://developer.twitter.com/en/docs/twitter-api/v1

Twitter4J API[3] to obtain streaming tweets. Similarly, Facebook and Sina Weibo have availed Facebook Graph API[4] and Tencent API[5], respectively. Accordingly, we used the available dataset, as shown in Table 1.

Table 1: The candidate available datasets for sentiment analysis researches

| Data scope | Data source |
|---|---|
| Film reviews | https://nlp.stanford.edu/sentiment/treebank.html |
| Movie reviews | https://seedmelab.org/ |
| SentiDrugs | http://www.druglib.com/ |
| Mandarin Text Data Only (SMP2019) | https://www.biendata.xyz/competition/smpecisa2019/ |
| Beautyspa, hotel, restaurant | https://github.com/hsqmlzno1/MGAN/tree/master/data |
| Tehran Stock Exchange | https://www.sahamyab.com/stocktwits |
| Altmetric data | https://www.altmetric.com/ |
| Testdata-manual | http://help.sentiment140.com/for-students/ |
| Airline-twitter-sentiment | https://www.crowdflower.com/data/airline-twitter-sentiment/ |
| Spam reviews | https://myleott.com/op-spam.html |
| Women's Clothing E-Commerce Reviews | https://www.kaggle.com/nicapotato/womens-ecommerce-clothing-reviews |
| SemEval-2014 | http://alt.qcri.org/semeval2014/task4/index.php?id=data- and- tools |
| Hindi ABSA dataset | http://www.iitp.ac.in/ ai- nlp- ml/resources.html |
| Amazon product reviews for four different product types: books, DVDs, electronics, and kitchen appliances. | http://www.cs.jhu.edu/∼mdredze/data sets/sentiment/. |
| Hotel reviews | https://webhose.io/ |
| Automobiles reviews | https://www.cvedia.com/ |
| Movie reviews | https://seedmelab.org/ |
| SemEval 2016 Task 5 | https://alt.qcri.org/semeval2016/task5/index.php?id=data-and-tools |
| SemEval 2015 Task 12 | https://alt.qcri.org/semeval2015/task12/index.php?id=data-and-tools |
| SemEval-2014 task 4 datasets (Laptop, Restaurant) | http://alt.qcri.org/semeval2014/task4/ |
| SemEval-2013 | www.cs.york.ac.uk/semeval-2013/task2/ |
| Reviews and forum posts | http://www.howardforums.com/ |
| Internet Movie Database (IMDB) | http://www.cs.cornell.edu/people/pabo/movie-review-data/ |
| Reviews of digital cameras on Taobao | http://www.taobao.com |
| Chinese implicit sentiment analysis task in SMP2019 | https://biendata.com/competition/smpecisa2019/ |
| SemEval 2017 Task 4 | http://alt.qcri.org/semeval2017/task4/ |
| SemEval 2014 and 2015 datasets | https://www.yelp.com/academic dataset<br>https://snap.stanford.edu/data/web-Amazon.html<br>http://nlp.stanford.edu/software/corenlp.shtm<br>https://crfsharp.codeplex.com/ |
| Opinions about products and services | https://tophotels.ru<br>https://tripadvisor.com<br>https://yelp.com |
| Hotel reviews (Chinese review) | https://www.ctrip.com/ |
| Notebook reviews (Chinese review) | https://global.jd.com/ |
| Reviews on Weibo (Chinese review) | https://www.ccir2015.com/ |
| Fine-grained technology product (Chinese review) | https://www.ccir2011.com/ |
| Restaurant and Hotel reviews (Vietnamese review) | https://vlsp.org.vn/resources-vlsp2018 |
| Ntc-sv (Vietnamese reviews about food and restaurant) | https://streetcodevn.com/blog/sav |
| Vreview (Vietnamese reviews about product[1], food[2], and restaurant[2]) | [1]https://www.aivivn.com/contests/1<br>[2]https://forum.machinelearningcoban.com/t/du-lieu-review-cua-foody/203 |

## 3.2. Text preprocessing

The opinions were first simplified by removing punctuation marks, retweet symbols, URL symbols, hashtags, and query terms. Next, one emoji icon was replaced by one describing the text. In addition, opinions on social networks are informal, with many acronyms and spelling errors. This can affect the accuracy of the results. Therefore, spelling corrections must be implemented. Subsequently, these opinions were preprocessed via tokenization and POS tags. Some available tools are often used for text preprocessing, as listed in Table 2.

[3]http://twitter4j.org/en/javadoc.html
[4]http://https//developers.facebook.com/docs/graph-api
[5]http://dev.datasift.com/docs/sources/public-sources/tencentweibo

Table 2: The candidate of available tools for text preprocessing

| Ref | Name of the tool | Purpose |
|---|---|---|
| [5] | GNU Aspell | Spell checker |
| [15] | POS tagger | Twitter POS tagger |
| [37] | TweeboParser | English parser |
| [51] | TweetMotif | Tokenization of tweets |
| [53] | TweetNLP | Twitter natural language processing |
| [54] | Lancaster stemming algorithm | Stemmer |
| [58] | Python-emoji package | Replace emoji by text |
| [58] | Python-based aspell library | Spell checker |
| [65] | Snowball | English stemmer |
| [82] | Stanford Log-linear Part-Of-Speech Tagger | POS tagger |

### 3.3. Text representation

An increasing number of machine-learning and deep-learning algorithms have been used in natural language processing to solve actual problems, particularly in text sentiment analysis. Most of the text data are the input data of the sentiment analysis models. Therefore, before using machine-learning and deep-learning algorithms, it is essential to convert text data into numerical vectors by building text representation models (TRMs). Text representation has been an active area of research in recent years. Building TRMs is an essential intermediate procedure that is emphasized by many researchers. TRMs are mathematical models used to convert text units, such as characters, words, sentences, and documents, as vectors of identifiers, such as index numbers in a corpus vocabulary [34, 70, 94]. TRMs for opinion sentiment analysis can be categorized into three approaches: basic vectorization approaches (one-hot encoding, bags of words, term frequency-inverse document frequency), distributed representations (Word2vec[6], Glove[7], Fasttext[8]), and universal language representations (BERT[9] and ELMo[10]). The advantages and disadvantages of these models are presented in Table 3.

Table 3: The candidate text representation methods for sentiment analysis researches

| Method | Advantage | Disadvantage |
|---|---|---|
| Continuous BOW (Word2Vec) | Mediocre semantic accuracy; Absence in papers; unpopular in practice | Ignores global vocabulary information;<br>Does not handle out-of-vocabulary words |
| Skipgram (Word2Vec) | Good semantic accuracy; Pre-trained models available | |
| GloVe | Uses global information of vocabulary;<br>Captures the local and global context of words;<br>Good syntactic and semantic accuracy; Pre-trained models available | Does not handle out-of-vocabulary words |
| FastText | Handles out-of-vocabulary words; Good at syntactic tasks<br>Pre-trained models available; Available aligned word vectors | Takes the longer time to train than Skipgram |
| BERT | Handles out-of-vocabulary words; Good at syntactic and semantic accuracy;<br>Pre-trained models available | Many computation-int-ensive at inference time;<br>Can become costly |

For basic vectorization approaches, text representation is implemented by mapping each word in the vocabulary of the text to an integer value. Subsequently, each sentence or

[6]https://code.google.com/archive/p/word2vec/
[7]https://nlp.stanford.edu/projects/glove/
[8]https://fasttext.cc/docs/en/crawl-vectors.html
[9]https://mccormickml.com/2019/05/14/BERT-word-embeddings-tutorial/
[10]https://www.analyticsvidhya.com/blog/2019/03/learn-to-use-elmo-to-extract-features-from-text/

document is represented as a vector such that the dimensionality of the vectors is equal to the size of the vocabulary. As shown in Table 3, these approaches have some limitations, e.g., the created vectors are sparse and high-dimensional representations. Moreover, they cannot solve words from a vocabulary list. Hence, distributed representation approaches have been developed to create dense, low-dimensional representations using neural network architectures. Currently, this is the typically used TRM. However, in this approach, each word still receives one fixed vector, whereas words can have different meanings depending on the context. Therefore, these methods do not capture the information directly. This problem has motivated some authors to present an approach based on universal language representations, in which BERT is currently the most state-of-the-art method. This approach has significantly improved the performance of some fundamental natural language processing (NLP) tasks, such as question answering and named entity recognition.

Although the previous pretrained TRMs are actively used and perform well, information loss is inevitable. Many approaches have been proposed to improve the methods above by considering the necessary information. In [77], the authors improved word representations by adding the global information. In [76], the authors improved the performance of word embeddings models based on Wikipedia categories. Jianqiang *et al.* [34] introduced an approach to enrich word embedding models using some information related to words, such as the GloVe vector, n-grams, and the sentiment score. Hassan *et al.* [39] improved word representations models by adding the semantic and syntactic information. Meanwhile, in [3], the authors considered information related to generic and sentiment-specific words. Rezaeinia *et al.* [71] considered the factors regarding words, such as the POS tag, lexicon, and position. Bojanowski *et al.* [8] added the morphology of words. Gong *et al.* [23] introduced a model for word representations by considering temporal or spatial information. Gou *et al.* [25] proposed a novel method to represent polysemous words into vectors by clustering context words. Phan *et al.* [63] added the lexicon, word type, sentiment polarity, position, and semantics of words.

### 3.4. Sentiment analysis model

Various approaches can be used to build sentiment analysis models based on opinions. The most typically used methods are illustrated in Figure 2. Using a reasonable method increases the performance of the sentiment analysis model. An appropriate method is selected based on the data, purpose of developing a model, etc. The approaches for selecting a fit model more easily are provided in Figure 2, and they are discussed in the following sections.

#### 3.4.1. Machine-learning approach

Machine-learning methods provide good performances and are widely used in sentiment analysis. These methods can be categorized into four main groups: supervised learning, unsupervised learning, semi-supervised learning, and deep learning. Deep learning is a state-of-the-art approach for sentiment analysis that provides good accuracy. Some researchers consider it a particular type of machine-learning method, whereas other researchers regard it as a separate group of algorithms that do not belong to machine-learning methods. In this study, we considered deep learning as a machine-learning algorithm. In machine-learning methods, a set of training data, labeled or not, is typically prepared at first. Subsequently,

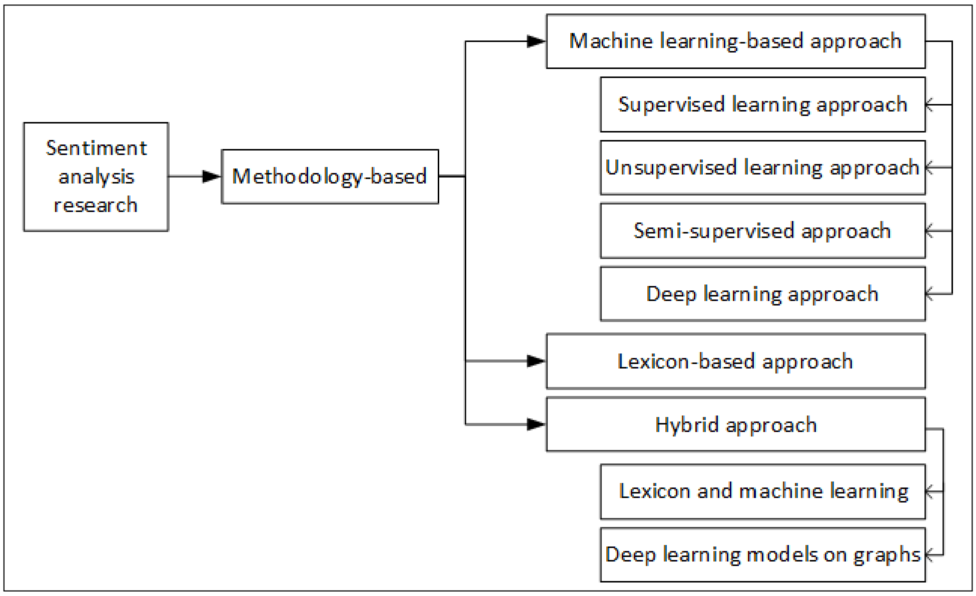


Figure 2: Sentiment analysis approaches from multiple researches

a set of features is extracted from the training data and forwarded to a classification model built using the four groups of approaches mentioned above. After the data are trained, the classifier can be used to classify the sentiment orientation of data that are not annotated in the testing set.

***Supervised learning approach.*** Supervised learning is the first machine-learning approach. Training data that require labeling are used in this method, i.e., each input opinion must be specified by one sentiment label by the annotators. Subsequently, the system attempts to understand a prediction function to estimate the relationship between the input and output by mapping their features. Although this method yields high performance, it often incur considerable time and cost as well as require experienced annotators to label data. Algorithms in the supervised learning approach can be categorized into two main categories [30]: probabilistic classification, including maximum entropy, naive Bayes, and Bayesian network; and non-probabilistic classification, including decision trees, neural networks, support vector machines, and rule-based algorithms.

We selected an algorithm to build the sentiment analysis model based entirely on our data and the problem to be solved. The models belonging to the supervised learning approach are often used in opinions sentiment analysis recently that are listed in Tables 9 and 10. The advantages and disadvantages of each method of the supervised learning-based approach are listed in Table 4 to facilitate the selection of a suitable algorithm.

***Unsupervised learning approach.*** Unsupervised learning is the second approach in machine-learning methods. In this method, unlabeled training data are used. Therefore, the system has to obtain a structure in data that have yet to be classified, and no predefined classes exist without the support of a human. Unlabeled data are abundant, readily available on the web, and relatively easy to obtain; however, they can be used in only a few approaches. Unsupervised learning can solve this problem using fully unlabeled training data. This reduces the time and cost required to assign labels to data. However, this

Table 4: Characteristic of supervised learning algorithms

| Supervised learning algorithm | Advantages | Disadvantages |
|---|---|---|
| Naive Bayesian | Very helpful to extract a subjective sentence<br>Explanation easily<br>No need for so much data to train a model | Assuming the features are independence |
| Bayesian network | Useful to understand complex domains<br>Take little time and cost to build a model | Can deal with limited of continuous variables |
| Maximum entropy | No limitation of features dimension<br>Can combine and add various knowledge sources easily | Over-fitting Problem |
| Support vector machine | Quite easy to train the model<br>Can good generalize in practice and theory<br>Not dependent so much on the feature dimensions | Need to choose an appropriate Kernel function<br>Need to increase in the samples to deceleration<br>Difficult interpretation |
| Neural network | Can deal with noise in data<br>Save time for training the model | Difficult implementation and interpretation;<br>High memory usage |
| K-nearest neighbor | Save the time to train the model<br>Quite easy to build the model | Sensitive to the type of measurement distance |
| Decision tree | Can deal with data noise and understand easily | Cannot apply it with small data |
| Rule-based | Can understand easily | Make overlap in the space of decision-making<br>Low performance with noisy data |

approach does not yield high performance.

Typically used algorithms for analyzing sentiments of the unsupervised learning approach include K-means, fuzzy c-means, agglomerative algorithms, and divisive algorithms [68]. The advantages and disadvantages of each algorithm are listed in Table 5. The models belonging to the unsupervised learning approach are often used in this field recently that are listed in Tables 9 and 10.

Table 5: Characteristic of unsupervised learning algorithms

| Unsupervised learning algorithm | Advantages | Disadvantages |
|---|---|---|
| K-means | Useful for large dataset<br>No need to pre-know the class of an opinion<br>No need free from human participation<br>Not require for high memory | Low accuracy with ambiguities<br>Sensitive with choosing the centroid and the number of clusters<br>Cannot deal with noise and outliers well<br>Cannot train on non-convex clusters<br>High dependence of performance on selecting of centroids |
| Fuzzy c-means | It always converges | Take so much time to train; Low performance with noise data<br>Sensitive to primary prediction and can stop at local minimums |
| Agglomerative algorithm | Not dependent on the number of clusters and centroids selection; Can deal with noise data | Unstable about time complexity for single and complete-link algorithms |
| Divisive algorithm | Not dependent on the number of clusters and centroids selection; Can deal with noise data | Need large memory; Take much time to train |

***Semi-supervised learning approach.*** The semi-supervised learning approach is the third approach in machine-learning methods for sentiment analysis. In this method, both labeled and unlabeled data can be used to train the model; this implies that semi-supervised learning uses a large amount of unlabeled data and a small amount of labeled data to build classifiers. This not only reduces the cost and time required, but also provides higher accuracy. Therefore, this approach is of great interest for opinions sentiment analysis.

The semi-supervised learning approach includes algorithms such as self-training, co-training, multiview learning, graph-based algorithms, as well as generative models [68]. Details regarding the advantages and disadvantages of each of the above algorithms are presented in Table 6. We listed some models belonging to the semi-supervised learning approach, which are often used in this field recently in Tables 9 and 10.

***Deep learning approach.*** In recent years, the deep learning approach has become a state-of-the-art approach in sentiment analysis because of its outstanding performance. This approach is the most suitable for unstructured data, such as text. Deep-learning methods

Table 6: Characteristic of semi-supervised learning algorithms

| Semi-supervised learning | Advantages | Disadvantages |
|---|---|---|
| Self-training | Understand easily and build simplicity<br>Not dependent on the classification model | Sensitive to error and outlier |
| Co-training | Not need the labeled data<br>Can get high performance with very small labeled data | Get high accuracy for only data with multi-viewpoints<br>Can get optimal accuracy with many available features |
| Graph-based | Can deal with multi and binary classifications<br>No need free from human participation<br>Can capture the structural relations more among data<br>Allow to harvest more insights from data | Sensitive to noise<br>Easily change graphs structure when data change<br>Tough to train graphs in solving the problems |
| Multi-view learning | Can handle problems of cross-lingual and various linguistic | Fail for language with low vocabulary |
| Generative models | Can get high accuracy with small labeled training data | Inflexibility; Low efficiency |

require large amounts of labeled input data for training a model. Subsequently, it uses multilayer models to transform low-level features into an abstract feature space. This implies that the deep-learning approach can describe the confidential information of input data more effectively than artificial features. Unlike other methods, deep-learning methods can automatically extract features from input data without effort from experts. Deep-learning methods yield higher accuracy than previously proposed methods. However, these methods incur more time and cost than other methods. Existing deep-learning-based approaches for opinion sentiment analysis include the following networks: autoencoder neural, recurrent neural, convolutional neural, long short-term memory, and recursive neural networks [30].

We selected an algorithm to build the sentiment analysis model based entirely on our data and the problem to be solved. The advantages and disadvantages of each method of the deep-learning approach are listed in Table 7 to facilitate the selection of a suitable algorithm.

Table 7: Characteristic of deep learning algorithms

| Deep learning | Advantages | Disadvantages |
|---|---|---|
| Autoencoder Neural Network | Can be great for feature extraction | Require a lot of data, processing time, hyperparameter tuning, and model validation. |
| Recurrent Neural Network | Can deal inputs of any length;<br>Very useful in any time series prediction;<br>Model size does not change even if the input size increases;<br>Can share the weights across the time steps | Take a long time for computation because of its recurrent nature;<br>Difficulty to train due to sequences of the process are very long if they use relu or tanh function;<br>Easily with problems related to gradient vanishing;<br>Only train word by word. |
| Convolutional Neural Network | Can train a series of words;<br>Can save time to train comparison with RNN model;<br>Can automatically extract features when training. | Take so much cost for computation;<br>Quite slow to train if don't have a good GPU (for complex tasks);<br>Need big training data |
| Long Short Term Memory network | Can learn long-term dependencies;<br>Can handle the problem of the vanishing of gradients | Take so much time and cost to train the model |
| Recursive Neural Network | Suitable to learn a directed acyclic graph structure from data;<br>Take advantages of the recurrent neural network due to it is<br>a generalization of the recurrent neural network. | Input data has to be processed hierarchically in a tree construct. |

### 3.4.2. Lexicon-based approach

The lexicon-based method uses sentiment dictionaries with opinion words and matches them with words appearing in opinions to determine the sentiment polarity of opinions. In sentiment dictionaries, words are assigned sentiment scores or labels describing positive, negative, and objective emotions. Lexicon-based approaches have been widely adopted in conventional textual sentiment analysis because they do not require data training. Some available vocabularies with sentiment scores have been presented for use in lexicon-based sentiment analysis. These vocabularies and their application method are presented in Section 2.3.1. This method is suitable for small data with short, simple text. Although this method requires less time and cost, its performance is low, and it is not suitable for complex texts. The lexicon-based method can be categorized into two approaches: dictionary- and corpus-based approaches. In the dictionary-based approach, a set of terms that are typically obtained and manually annotated is used. This set is created and extended by searching synonyms and antonyms of dictionaries, such as WordNet and SentiWordNet. The primary limitation of this approach is that it cannot manage the context and domain-specific orientations. In the corpus-based approach, dictionaries related to a specific domain are provided. These dictionaries are created using original opinion terms that extend through the discovery of related words using the following techniques: (i) The semantic method using synonyms and antonyms or relationships from the thesaurus, such as WordNet; (ii) the statistical technique using latent semantic analysis.

#### 3.4.3. Hybrid approach

This study presents two main hybrid approaches that are often used in sentiment analysis, namely (i) the Hybrid approach between the Lexicon-based approach and Machine learning methods. (ii) The Hybrid approach by using Deep learning methods over graph structures.

***Lexicon-based approach and Machine learning methods***. The hybrid approach combines machine-learning and lexicon-based methods. This approach not only reflects the weakness of the two methods but also integrates their strengths. Machine-learning and lexicon-based techniques are combined by extracting some features related to the lexicon, syntax, and linguistic, followed by feeding them into machine-learning classifiers. Another method to combine both approaches is by removing words that do not exist in the lexicon from all instances such that only sentiment words will remain. Although the removed terms do not affect the sentiment, they will confuse the classifier and decrease the accuracy. Therefore, when these words are eliminated, the accuracy of the model is expected to increase. The hybrid method has been acknowledged as a better approach than machine-learning and lexicon-based methods.

***Deep learning models on graphs***. As we can know, graph structures have been used in various applications, especially social analysis. Because graphs can capture the structural relations more among data and thus allows for harvesting more insights from data. However, the main limitation of graphs is tough to train graphs in solving the problems. We can solve this limitation by combining graphs with deep learning models. From that, deep learning models on graphs have been introduced and proved superior performance in solving various problems, and most recently in opinions sentiment analysis. Of these, the most common is graph neural networks (GNNs). The models of GNNs can be divided into four main categories, such as recurrent graph neural networks, convolutional graph neural networks (or graph convolutional networks (GCNs)), graph autoencoders, and spatial-temporal graph neural networks. Today, GNNs are quickly noticed due to their great expressive power. One of the categories of GNNs that is the GCNs have most successfully applied in realizing some NLP tasks, such as semantic role labeling, relational classification, text classification, machine translation, and especially the sentiment analysis field. Various GCNs models have been built to apply in opinions sentiment analysis on social media, including [7, 11, 31, 38, 83, 85, 92, 98, 101]. In our assessment via studying related literature, deep learning models on graphs are the most state-of-the-art approaches currently using in opinions sentiment analysis and interested by many researchers.

#### 3.4.4. Comparison

Based on the discussions above, it is evident that machine-learning-based algorithms are the most typically used algorithms for opinions sentiment analysis, in particular for graph-based methods. However, deep-learning methods have garnered increasing attention. It is noteworthy that machine-learning methods have several limitations. The performance of machine-learning methods depends on the number of input opinions, number of labeled opinions, and domain dependence. Meanwhile, lexicon-based methods are used because they require neither labeled data nor much input data. However, the performance of these methods is low and depends significantly on lists of sentiment words. Therefore, hybrid approaches have been proposed to compensate for the shortcomings of machine-learning and lexicon-based approaches. The strengths and weaknesses of the methods applied for opinions sentiment analysis [30, 68] are described in detail in Table 8.

Table 8: Characteristic of sentiment analysis approaches

| Approach | Time training | Accuracy | Advantage | Disadvantage |
|---|---|---|---|---|
| Machine learning | | | | |
| Supervised learning | slow | very high | Useful to classify both binary-class and multi-class<br>Very helpful to extract a subjective issue<br>Can deal with noise data | High dependent on the labeled data<br>Need human participation and have to linguistic pre-analyzation<br>Take much time and cost to train, especially high dimensional data |
| Semi-supervised learning | medium | high | Can handle well for ambiguities opinions<br>Can get highest performance with a few annotations effort | Fail with the unlabeled data containing noise |
| Unsupervised learning | fast | medium | Not require much free from human participation<br>Can get efficient and widely applicable | Not been proven ability to multi-classification<br>Low performance with noise data<br>Unknown the number of clusters in most cases<br>Not stable of performance, sometimes be quite low |
| Deep learning | slow<br>medium if training by GPU | high | Useful to classify both binary-class and multi-class<br>Very helpful to extract a subjective issue<br>Can deal with noise data<br>Can automatically extract features | High dependent on the labeled data<br>Need human participation and have to linguistic pre-analyzation<br>Take much time and cost to train, especially high dimensional data |
| Lexicon-based | | | | |
| Dictionary-based | very fast | relatively low | Not need any labeled training opinions<br>Easy understand and simple implement<br>Achieve better results for less banded domain | No lexicon of words with the specific content-oriented domain<br>Not fit with opinions containing certain semantic dependency<br>Less accurate with opinions containing various domains |
| Corpus-based | very fast | relatively low | Ability to find words with the specific content-oriented<br>Better accurate with opinions containing various domains | High dependence of performance on the lexicon domain<br>Difficult to cover all the opinion words with large texts<br>Can not be used alone |
| Hybrid approach | | | | |
| Machine learning with Lexicon | medium | high | Take advantage of the machine learning approach<br>and the lexicon-based approach regarding the time<br>and accuracy | The accuracy still depends on the lexicon quality |
| Deep learning on graphs | quite fast | high | More valuable and suitable to combine with<br>other neural networks<br>Suitable for solving practical problems | Convolution layers are limited (about 2 and 3 layers)<br>Weights of edges are mainly binary value<br>Using a constant number of layers |

### 3.5. Evaluation methods

The metrics that are typically used to assess sentiment analysis methods include $\mathcal{Precision}$, $\mathcal{Recall}$, $\mathcal{F}_1$, and $\mathcal{Accuracy}$, which are computed as follows

$$\mathcal{Precision} = \frac{\mathcal{TP}}{\mathcal{TP} + \mathcal{FP}}, \tag{16}$$

$$\mathcal{Recall} = \frac{\mathcal{TP}}{\mathcal{TP} + \mathcal{FN}}, \tag{17}$$

$$\mathcal{F}_1 = 2 \times \frac{\mathcal{Precision} \times \mathcal{Recall}}{\mathcal{Precision} + \mathcal{Recall}}, \tag{18}$$

$$\mathcal{Accuracy} = \frac{\mathcal{TP} + \mathcal{TN}}{\mathcal{TP} + \mathcal{FP} + \mathcal{TN} + \mathcal{FN}}, \tag{19}$$

where $\mathcal{TP}$ is an acronym for *"true positive"*, which represents the number of exactly classified items; $\mathcal{FP}$ is an acronym for *"false positive"*, which is the number of misclassified items; $\mathcal{FN}$ is an acronym for *"false negative"*, which is the number of misclassified non-items; $\mathcal{TN}$ is an acronym for *"true negative"*, which is the number of classified non-items.

## 4. APPLICATIONS OF SENTIMENT ANALYSIS

Tables 9 and 10 present some of the most important recent studies in opinion sentiment analysis. The comparison is based on the techniques used, level, applications, and datasets.

Table 9: Summary of models published for sentiment analysis in opinions on social media, recently

| Refer | Data | Technique | Level-based | Application |
|---|---|---|---|---|
| [1] | Women's Clothing E-Commerce Reviews | RNN + LSTM | Sentence | Understanding customer sentiments |
| [2] | SemEval-2014<br>Hindi ABSA dataset | CNN+BiLSTM | Aspect | Aspect term extraction and<br>aspect sentiment classification |
| [4] | SemEval 2013, SemEval 2014<br>Vader, STS-Gold, IMDB, and PL04 | Deep learning | Document | Improving the performance of sentiment analysis |
| [7] | Health-care reform<br>Stanford, IMDB, Michigan, SemEval | Deep graph-based text representation | Aspect | Aspect-based sentiment analysis |
| [9] | Amazon product reviews for four different product types:<br>books, DVDs, electronics, and kitchen appliances. | Sentiment Sensitive Embeddings | Sentence | Unsupervised Cross-domain Sentiment<br>Classification |
| [10] | Amazon product reviews for four different product types:<br>books, DVDs, electronics, and kitchen appliances. | Semi-supervised learning | Document | Cross-domain sentiment classification |
| [11] | SemEval 2015 and 2016 datasets | Hier-GCN | Aspect | Aspect-Category based Sentiment Analysis |
| [16] | Movie review, internet movie review,:<br>and Stanford sentiment treebank. | BiLSTM | Document | Sentiment analysis method<br>of capsule network |
| [17] | Amazon product review | Wasserstein based Transfer Network | Sentence | Cross-domain sentiment classification |
| [18] | News related to COVID-19 | Machine learning algorithms | Document | Detect misleading information on COVID-19 |
| [19] | Amazon Review (French, German, Japanese)<br>Yelp and Hotel Review (Chinese)<br>Social Media Posts (Arabic) | Unsupervised Machine Translation | Sentence | Cross-Lingual Text Classification |
| [20] | Product review and forum discussion sentences on<br>digital cameras, DVD players, MP3 players, Intel<br>vs AMD, Coke vs Pepsi, and Microsoft vs Google | Lexicon-based | Sentence | Mining opinions from comparative sentences |
| [21] | Movie Reviews (MR) dataset | CNN+LSTM | Document | Analyzing sentiments and<br>classification of the opinions |
| [22] | Online user reviews in both Greek and English languages | Aspect embeddings | Document | Detection of sentiments expressed<br>for certain entity |
| [26] | SemEval 2013 Task 2 | Hybrid approach | Aspect | Twitter sentiment analysis |
| [27] | SentiDrugs demonstrate | PM-DBiGRU model | Aspect | Aspect-Level Drug Reviews Sentiment Analysis |

## 5. DISCUSSIONS AND CHALLENGES

Next, we list the issues in sentiment analysis.

Opinions posted on social media are written in an informal style. That means these opinions contain a lot of grammatical errors, noisy, misspelled, and unstructured. Therefore, it takes a lot of time and cost for data preprocessing and converting unstructured data into a structured format to get a usable standard data set. This is also a challenge that we need to focus on in the future. Because according to our assessment, the data set significantly affects the performance of the models.

For aspect-level sentiment analysis, its main challenge is determining the sentiment relationship between one word or phrase expressing aspect and words describing the sentiment of the word or phrase indicating aspect correctly. Until now, this challenge has not yet been solved completely.

Experimentally we have found that the same model for sentiment analysis applied to different data in terms of quality can give various performances. Thus, the effectiveness of the sentiment analysis models depend heavily on the quality of the training dataset. Currently, the preparation of training data is considered a challenge in sentiment analysis of social media opinions.

Building cross-lingual sentiment analysis systems that can support all of the languages with high accuracy is still a challenge. This problem needs to solve in the future.

Cross-domain sentiment analysis remains a significant challenge. Three issues must be addressed in cross-domain sentiment analysis. First, the same sentiment word can express different sentiments for different domains. For example, *"This drug helps me so healthy"*

Table 10: Summary of models published for sentiment analysis in opinions on social media, recently (cont)

| Refer | Data | Technique | Level-based | Application |
|---|---|---|---|---|
| [28] | Tweets | SentiStrength tool | Sentence | Predicting literature's early impact |
| [29] | Online textual comments and Tehran Stock Exchange data | Hybrid approach | Sentence | Tehran Stock Exchange Prediction |
| [31] | Laptop and Restaurant reviews | SA-GCN | Aspect | Aspect-level Sentiment classification |
| [32] | Hotel reviews; Automobiles reviews; Movie reviews | Hybrid approach | Aspect | Aspect sentiment analysis |
| [33] | Cryptocurrency names and concurrent price data | Multilinear regression model | Sentence | Predict the prices of Bitcoin and Litecoin |
| [34] | Tweets | Deep CNN | Sentence | Twitter sentiment analysis |
| [35] | Tweets | BiLSTM+CNN | Sentence | Tweet classification and rescue scheduling |
| [38] | SemEval-2014 task | Interactive gated convolutional network | Aspect | Aspect sentiment analysis |
| [40] | Vietnamese reviews of restaurant and hotel | Transfer learning method | Aspect | Aspect-based sentiment analysis |
| [41] | Multi-domain dataset named YelpAspect | BiLSTM, Context2Aspect attention and Coarse2fine attention | Aspect | Aspect sentiment analysis |
| [43] | Twitter data | Supervised learning: Naive Bayes, Neural net, SVM | Sentence | Detection of Schizophrenia |
| [45] | Bing Liu datasets (Computer , Router, Speaker) SemEval-2014 task 4 datasets (Laptop, Restaurant) Canon G3, Creative Zen, Nokia 6610 | Adaptive aspect-based lexicons | Aspect | Aspect sentiment analysis |
| [46] | Conditional sentences from 5 different user forums: Cellphone, Automobile, LCD TV, Audio systems and Medicine | Supervised learning models | Sentence | Sentiment analysis of conditional sentences |
| [47] | Comments of electronic products | Combination of Vietnamese language structure, NLP techniques, self-attention network, and Transformer architecture | Sentence | Vietnamese sentiment analysis |
| [48] | Twitter Laptop and Restaurant reviews | Interactive Lexicon-Aware Word-Aspect Attention Network | Aspect | Aspect sentiment analysis |
| [49] | Ntc-sv and Vreview | Fine-Tuning BERT | Sentence | Vietnamese reviews sentiment analysis |
| [50] | Review data—Yelp and Amazon Clothing | Seq2Seq model +Aspect Conditional Masked Language Model | Aspect | Recommendation justification |
| [52] | Collection of suicide-related tweets | Support Vector Machine, Logistic Regression | Document | Detecting suicidality on Twitter |
| [56] | Tweets containing conditional sentences | CNN+MLP | Sentence | Detecting and analyzing the sentiment of tweets containing conditional sentences |
| [58] | Tweets regarding Phone topic | BiLSTM | Aspect | Measuring the user satisfaction level on Twitter |
| [59] | Tweets containing fuzzy sentiment phrase | Lexicon-based | Sentence | Detecting and analyzing the sentiment of tweets containing fuzzy sentiment phrases |
| [60] | Tweets | CNN+BiLSTM | Aspect | Decision-making support method |
| [61] | Tweets | Feature ensemble model +CNN | Document | Detect the degree of risk in an online market |
| [62] | Tweets | Feature ensemble model +CNN | Document | Detect psychological tendencies and causations |
| [63] | Tweets | Feature ensemble model+CNN | Sentence | Sentiment analysis of tweets containing fuzzy sentiment |
| [64] | UIT-ViSFD | Bi-LSTM and fastText word embeddings | Aspect | Vietnamese aspect-based sentiment |
| [66] | Internet Movie Database (IMDb) | Reinforcement learning | Document | Replicating human interpretations |
| [69] | Tweets | Density-based clustering algorithms | Sentence | Tweets sentiment analysis |
| [70] | Movie reviews | Hybrid CNN-LSTM model | Document | Improving accuracy of movie reviews sentiment analysis |
| [72] | Tweets | Machine learning algorithms | Document | Tweets classification |
| [74] | SemEval 2014 Task 4 | Feature distillation network | Aspect | Aspect sentiment analysis |
| [83] | Restaurant,Laptop, and MAMS datasets | Gated Graph Convolutional Networks and Syntax-based Regulation | Aspect | Aspect-level Sentiment classification |
| [84] | Reviews of digital cameras on Taobao | Association constrained LDA (AC-LDA) | Aspect and document | Joint extraction of product aspects and opinions |
| [85] | Restaurant,Laptop, and Twitter datasets | Relational graph attention network (R-GAT) | Aspect | Aspect-level Sentiment classification |
| [86] | Reviews on social media in China | BERT model | Document | Negative sentiment analysis on social media |
| [87] | SemEval 2014 | Recursive neural network and Conditional random field | Aspect | Aspect sentiment analysis |
| [88] | Chinese implicit sentiment analysis task in SMP2019 SemEval 2017 Task 4 | Bilstm with multi-polarity orthogonal attention | Sentence | Implicit sentiment analysis |
| [90] | SemEval-2014 Task4 Twitter dataset | LSTM + attention mechanisms | Aspect | Aspect sentiment analysis |
| [91] | Twitter data of an entire year from September5th, 2013 to September 5th, 2014 | Semi-supervised learning | Document | The effect of climate and seasonality on the prevalence of depression in Twitter users |
| [92] | 20-Newsgroups Ohsumed, R52 and R8 of Reuters 21578 and Movie Review | Text Graph Convolutional Network (TextGCN) | Document | Text classification |
| [93] | SemEval 2014 | CNN + Dependency tree | Aspect | Aspect term extraction |
| [95] | Product review and the SemEval 2014 and 2015 datasets | Unsupervised Word and Dependency Path Embeddings | Aspect | Aspect term extraction |
| [97] | Opinions about products and services | Supervised learning methods | Aspect | Decision support approach |
| [98] | Twitter, Semevals 2014, 2015, and 2016 | Aspect-specific GCNs | Aspect | Aspect-based sentiment analysis |
| [100] | Four Chinese review dataset: Hotel, notebook, Weibo fine-grained technology product Four Amazon online English review dataset: | Multi-source domain adaptation and joint learning | Document | Cross-domain sentiment classification |
| [101] | Twitter, Semevals 2014, 2015, and 2016 | Syntax- and Knowledge-based GCN | Aspect | Aspect-based sentiment analysis |

expresses the positive sentiment of the user. Meanwhile, *"That thief is so healthy"* expresses the negative sentiment of the users. Second, we should consider the differences in sentiment vocabularies across different domains. Finally, an objective method to assign a strength score to each sentiment word should be identified [30, 68].

The superiority of methods are separately applied when analyzing sentiment analysis has almost been leveraged to improve the performance of social media sentiment analysis methods. The current trend is to find ways to combine methods to use advantages of one method to reduce the limitations of the integrated approach. Therefore, finding the suitable methods to combine and optimize the parameters when combining these methods becomes a challenge that we need to solve in the future.

For approaches using deep learning algorithms on graph structures, we need to care about handling when the number of nodes on the graphs is too much. In addition, how to construct GNNs or GCNs that may contain different types of nodes and edges or nodes and edges containing much information related to sentiment is also a current challenge in improving the performance of sentiment analysis methods.

As we all know, human sentiment can change over time. Building models that can capture the streaming of sentiment according to time is also a challenge that needs to be solved for the sentiment analysis field in the future.

Finally, issues pertaining to contextual sentiment analysis, intrinsic feature-based sentiment analysis, sense level subjectivity classification, and sentiment of streaming text remain to be solved, whereas a suitable automatic sentiment analysis system for different media must be identified.

## 6. CONCLUSIONS

A comprehensive review of investigations performed for sentiment analysis from 2018 to 2020 was provided herein. First, the background concepts related to sentiment analysis was introduced. Subsequently, the general architecture of the sentiment analysis process was described, and the procedures involved in this process were reviewed in detail. Next, discussions were presented as well as challenges that must be addressed in the future.

One of the main contributions of this survey is that a list of available public datasets and word embeddings for sentiment analysis was provided. In addition, the weaknesses and strengths of most approaches of sentiment analysis in opinions were compared. Additionally, the methods presented in the most recent publications were discussed in terms of data, techniques, applications, etc. The most important contribution of the survey is that invaluable and interesting discussions and challenges were presented. This survey will enable new researchers and students who can capture and identify the most typical challenges precisely to devise better solutions in the future.

## REFERENCES

[1] A. F. Agarap and P. Grafilon, "Statistical analysis on e-commerce reviews, with sentiment classification using bidirectional recurrent neural network (rnn)," *arXiv preprint arXiv:1805.03687*, 2018.

[2] M. S. Akhtar, T. Garg, and A. Ekbal, "Multi-task learning for aspect term extraction and aspect sentiment classification," *Neurocomputing*, 2020.

[3] N. Al-Twairesh and H. Al-Negheimish, “Surface and deep features ensemble for sentiment analysis of arabic tweets,” *IEEE Access*, vol. 7, pp. 84 122–84 131, 2019.

[4] O. Araque, I. Corcuera-Platas, J. F. Sánchez-Rada, and C. A. Iglesias, “Enhancing deep learning sentiment analysis with ensemble techniques in social applications,” *Expert Systems with Applications*, vol. 77, pp. 236–246, 2017.

[5] K. Atkinson, “Gnu aspell 0.60. 4,” 2006.

[6] S. Baccianella, A. Esuli, and F. Sebastiani, “Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining.” in *Lrec*, vol. 10, no. 2010, 2010, pp. 2200–2204.

[7] K. Bijari, H. Zare, E. Kebriaei, and H. Veisi, “Leveraging deep graph-based text representation for sentiment polarity applications,” *Expert Systems with Applications*, vol. 144, 2020. [Online]. Available: http//doi.org/10.1016/j.eswa.2019.113090

[8] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching word vectors with subword information,” *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 135–146, 2017.

[9] D. Bollegala, T. Mu, and J. Y. Goulermas, “Cross-domain sentiment classification using sentiment sensitive embeddings,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 398–410, 2015.

[10] D. Bollegala, D. Weir, and J. Carroll, “Cross-domain sentiment classification using a sentiment sensitive thesaurus,” *IEEE transactions on knowledge and data engineering*, vol. 25, no. 8, pp. 1719–1731, 2012.

[11] H. Cai, Y. Tu, X. Zhou, J. Yu, and R. Xia, “Aspect-category based sentiment analysis with hierarchical graph convolutional network,” in *Proceedings of the 28th International Conference on Computational Linguistics*, 2020, pp. 833–843.

[12] E. Cambria and A. Hussain, “Sentic computing,” *Cognitive Computation*, vol. 7, no. 2, pp. 183–185, 2015.

[13] E. Cambria, D. Olsher, and D. Rajagopal, “Senticnet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis,” in *Proceedings of the twenty-eighth AAAI conference on artificial intelligence*, 2014, pp. 1515–1521.

[14] K. Chakraborty, S. Bhattacharyya, and R. Bag, “A survey of sentiment analysis from social media data,” *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 450–464, 2020.

[15] L. Derczynski, A. Ritter, S. Clark, and K. Bontcheva, “Twitter part-of-speech tagging for all: Overcoming sparse and noisy data,” in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, 2013, pp. 198–206.

[16] Y. Dong, Y. Fu, L. Wang, Y. Chen, Y. Dong, and J. Li, “A sentiment analysis method of capsule network based on bilstm,” *IEEE Access*, vol. 8, pp. 37 014–37 020, 2020.

[17] Y. Du, M. He, L. Wang, and H. Zhang, “Wasserstein based transfer network for cross-domain sentiment classification,” *Knowledge-Based Systems*, vol. 204, pp. 106–162, 2020.

[18] M. K. Elhadad, K. F. Li, and F. Gebali, “Detecting misleading information on covid-19,” *IEEE Access*, vol. 8, pp. 165 201–165 215, 2020.

[19] H. Fei and P. Li, "Cross-lingual unsupervised sentiment classification with multi-view transfer learning," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 5759–5771.

[20] M. Ganapathibhotla and B. Liu, "Mining opinions in comparative sentences," in *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, 2008, pp. 241–248.

[21] M. Ghorbani, M. Bahaghighat, Q. Xin, and F. Özen, "Convlstmconv network: a deep learning approach for sentiment analysis in cloud computing," *Journal of Cloud Computing*, vol. 9, no. 1, pp. 1–12, 2020.

[22] M. Giatsoglou, M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, and K. C. Chatzisavvas, "Sentiment analysis leveraging emotions and word embeddings," *Expert Systems with Applications*, vol. 69, pp. 214–224, 2017.

[23] H. Gong, S. Bhat, and P. Viswanath, "Enriching word embeddings with temporal and spatial information," *arXiv preprint arXiv:2010.00761*, 2020.

[24] P. Grover, A. K. Kar, Y. K. Dwivedi, and M. Janssen, "Polarization and acculturation in us election 2016 outcomes–can twitter analytics predict changes in voting preferences," *Technological Forecasting and Social Change*, vol. 145, pp. 438–460, 2019.

[25] S. Guo and N. Yao, "Polyseme-aware vector representation for text classification," *IEEE Access*, vol. 8, pp. 135 686–135 699, 2020.

[26] I. Gupta and N. Joshi, "Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1611–1625, 2019.

[27] Y. Han, M. Liu, and W. Jing, "Aspect-level drug reviews sentiment analysis based on double bigru and knowledge transfer," *IEEE Access*, vol. 8, pp. 21 314–21 325, 2020.

[28] S.-U. Hassan, N. R. Aljohani, N. Idrees, R. Sarwar, R. Nawaz, E. Martínez-Cámara, S. Ventura, and F. Herrera, "Predicting literature's early impact with sentiment analysis in twitter," *Knowledge-Based Systems*, vol. 192, 2020. [Online]. Available: https://doi.org/10.1016/j.knosys.2019.105383

[29] A. Hatefi Ghahfarrokhi and M. Shamsfard, "Tehran stock exchange prediction using sentiment analysis of online textual opinions," *Intelligent Systems in Accounting, Finance and Management*, vol. 27, no. 1, pp. 22–37, 2020.

[30] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artificial Intelligence Review*, pp. 1–51, 2019.

[31] X. Hou, J. Huang, G. Wang, K. Huang, X. He, and B. Zhou, "Selective attention based graph convolutional networks for aspect-level sentiment classification," *arXiv preprint arXiv:1910.10857*, 2019.

[32] A. Ishaq, S. Asghar, and S. A. Gillani, "Aspect-based sentiment analysis using a hybridized approach based on cnn and ga," *IEEE Access*, vol. 8, pp. 135 499–135 512, 2020.

[33] A. Jain, S. Tripathi, H. DharDwivedi, and P. Saxena, "Forecasting price of cryptocurrencies using tweets sentiment analysis," in *2018 Eleventh International Conference on Contemporary Computing (IC3)*. IEEE, 2018, pp. 1–7.

[34] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, “Deep convolution neural networks for twitter sentiment analysis,” *IEEE Access*, vol. 6, pp. 23 253–23 260, 2018.

[35] M. Y. Kabir and S. Madria, “A deep learning approach for tweet classification and rescue scheduling for effective disaster management,” in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2019, pp. 269–278.

[36] C. Karyotis, F. Doctor, R. Iqbal, A. James, and V. Chang, “A fuzzy computational model of emotion for cloud based sentiment analysis,” *Information Sciences*, vol. 433, pp. 448–463, 2018.

[37] L. Kong, N. Schneider, S. Swayamdipta, A. Bhatia, C. Dyer, and N. A. Smith, “A dependency parser for tweets,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1001–1012.

[38] A. Kumar, V. T. Narapareddy, V. A. Srikanth, L. B. M. Neti, and A. Malapati, “Aspect-based sentiment classification using interactive gated convolutional network,” *IEEE Access*, vol. 8, pp. 22 445–22 453, 2020.

[39] F. M. Kundi, S. Ahmad, A. Khan, and M. Z. Asghar, “Detection and scoring of internet slangs for sentiment analysis using sentiwordnet,” *Life Science Journal*, vol. 11, no. 9, pp. 66–72, 2014.

[40] N. C. Le, N. T. Lam, S. H. Nguyen, and D. T. Nguyen, “On vietnamese sentiment analysis: A transfer learning method,” in *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*. IEEE, 2020, pp. 1–5.

[41] Z. Li, Y. Wei, Y. Zhang, X. Zhang, and X. Li, “Exploiting coarse-to-fine task transfer for aspect-level sentiment classification,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 4253–4260.

[42] B. Liu, “Sentiment analysis and opinion mining,” *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[43] K. McManus, E. K. Mallory, R. L. Goldfeder, W. A. Haynes, and J. D. Tatum, “Mining twitter data to improve detection of schizophrenia,” *AMIA Summits on Translational Science Proceedings*, vol. 2015, p. 122, 2015.

[44] G. A. Miller, “Wordnet: a lexical database for english,” *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[45] M. E. Mowlaei, M. S. Abadeh, and H. Keshavarz, “Aspect-based sentiment analysis using adaptive aspect-based lexicons,” *Expert Systems with Applications*, vol. 148, p. 113234, 2020.

[46] R. Narayanan, B. Liu, and A. Choudhary, “Sentiment analysis of conditional sentences,” in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 2009, pp. 180–189.

[47] H. D. Nguyen, T. Huynh, S. N. Hoang, V. T. Pham, and I. Zelinka, “Language-oriented sentiment analysis based on the grammar structure and improved self-attention network.” in *ENASE 2020 - 15th International Conference on Evaluation of Novel Approaches to Software Engineering*, 2020, pp. 339–346.

[48] H.-T. Nguyen and L.-M. Nguyen, “Ilwaanet: An interactive lexicon-aware word-aspect attention network for aspect-level sentiment classification on social networking,” *Expert Systems with Applications*, vol. 146, p. 113065, 2020.

[49] Q. T. Nguyen, T. L. Nguyen, N. H. Luong, and Q. H. Ngo, "Fine-tuning bert for sentiment analysis of vietnamese reviews," in *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*. IEEE, 2020, pp. 302–307.

[50] J. Ni, J. Li, and J. McAuley, "Justifying recommendations using distantly-labeled reviews and fine-grained aspects," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019, pp. 188–197.

[51] B. O'Connor, M. Krieger, and D. Ahn, "Tweetmotif: exploratory search and topic summarization for twitter." in *ICWSM*, 2010, pp. 384–385.

[52] B. O'dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, "Detecting suicidality on twitter," *Internet Interventions*, vol. 2, no. 2, pp. 183–188, 2015.

[53] O. Owoputi, B. O'Connor, C. Dyer, K. Gimpel, N. Schneider, and N. A. Smith, "Improved part-of-speech tagging for online conversational text with word clusters," in *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2013, pp. 380–390.

[54] C. D. Paice, "Another stemmer," in *ACM Sigir Forum*, vol. 24, no. 3. ACM New York, NY, USA, 1990, pp. 56–61.

[55] H. T. Phan, N. T. Nguyen, D. Hwang *et al.*, "A new approach for predicting an important user on a topic on twitter," in *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE, 2020, pp. 1–6.

[56] H. T. Phan, N. T. Nguyen, V. C. Tran, and D. Hwang, "A method for detecting and analyzing the sentiment of tweets containing conditional sentences," in *Asian Conference on Intelligent Information and Database Systems*. Springer, 2019, pp. 177–188.

[57] ——, "A sentiment analysis method of objects by integrating sentiments from tweets," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 6, pp. 7251–7263, 2019.

[58] ——, "An approach for a decision-making support system based on measuring the user satisfaction level on twitter," *Information Sciences*, 2021. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025521000098

[59] H. T. Phan, N. T. Nguyen, T. Van Cuong, and D. Hwang, "A method for detecting and analyzing the sentiment of tweets containing fuzzy sentiment phrases," in *2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*. IEEE, 2019, pp. 1–6.

[60] H. T. Phan, V. C. Tran, N. T. Nguyen, and D. Hwang, "Decision-making support method based on sentiment analysis of objects and binary decision tree mining," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2019, pp. 753–767.

[61] ——, "Detecting the degree of risk in online market based on satisfaction of twitter users," in *International Conference on Computational Collective Intelligence*. Springer, 2020, pp. 58–70.

[62] ——, "A framework for detecting user's psychological tendencies on twitter based on tweets sentiment analysis," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2020, pp. 357–372.

[63] ——, "Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model," *IEEE Access*, vol. 8, pp. 14 630–14 641, 2020.

[64] L. L. Phan, P. H. Pham, K. T.-T. Nguyen, T. T. Nguyen, S. K. Huynh, L. T. Nguyen, T. Van Huynh, and K. Van Nguyen, "Sa2sl: From aspect-based sentiment analysis to social listening system for business intelligence," *arXiv preprint arXiv:2105.15079*, 2021.

[65] M. F. Porter, "Snowball: A language for stemming algorithms," 2001.

[66] N. Pröllochs, S. Feuerriegel, B. Lutz, and D. Neumann, "Negation scope detection for sentiment analysis: A reinforcement learning framework for replicating human interpretations," *Information Sciences*, 2020.

[67] C. M. Pulido, B. Villarejo-Carballido, G. Redondo-Sama, and A. Gómez, "Covid-19 infodemic: More retweets for science-based information on coronavirus than for false information," *International Sociology*, 2020. [Online]. Available: https://doi.org/10.1177/0268580920914755

[68] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," *Knowledge-based systems*, vol. 89, pp. 14–46, 2015.

[69] H. Rehioui and A. Idrissi, "New clustering algorithms for twitter sentiment analysis," *IEEE Systems Journal*, vol. 14, no. 1, pp. 530–537, 2020.

[70] A. U. Rehman, A. K. Malik, B. Raza, and W. Ali, "A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis," *Multimedia Tools and Applications*, vol. 78, no. 18, pp. 26 597–26 613, 2019.

[71] S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings," *Expert Systems with Applications*, vol. 117, pp. 139–147, 2019.

[72] J. Samuel, G. Ali, M. Rahman, E. Esawi, Y. Samuel *et al.*, "Covid-19 public sentiment insights and machine learning for tweets classification," *Information*, vol. 11, no. 6, p. 314, 2020.

[73] A. Sarker, S. Lakamana, W. Hogg-Bremer, A. Xie, M. A. Al-Garadi, and Y.-C. Yang, "Self-reported covid-19 symptoms on twitter: An analysis and a research resource," *Journal of the American Medical Informatics Association*, vol. 27, pp. 1310–1315, 2020. [Online]. Available: https://doi.org/10.1093/jamia/ocaa116

[74] K. Shuang, Q. Yang, J. Loo, R. Li, and M. Gu, "Feature distillation network for aspect-based sentiment analysis," *Information Fusion*, vol. 61, pp. 13–23, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253519303951

[75] C. Strapparava, A. Valitutti *et al.*, "Wordnet affect: an affective extension of wordnet." in *Lrec*, vol. 4, no. 1083-1086. Citeseer, 2004, p. 40.

[76] L. Svoboda and T. Brychcın, "Improving word meaning representations using wikipedia categories," *Neural Network World*, vol. 523, p. 534, 2018.

[77] L. Svoboda and T. Brychcín, "Enriching word embeddings with global information and testing on highly inflected language," *Computación y Sistemas*, vol. 23, no. 3, 2019.

[78] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.

[79] S. Tedmori and A. Awajan, "Sentiment analysis main tasks and applications: A survey." *Journal of Information Processing Systems*, vol. 15, no. 3, pp. 500–519, 2019. [Online]. Available: https://doi.org/10.3745/JIPS.04.0120

[80] M. Thelwall and K. Buckley, "Topic-based sentiment analysis for the social web: The role of mood and issue-related words," *Journal of the American Society for Information Science and Technology*, vol. 64, no. 8, pp. 1608–1617, 2013.

[81] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, 2012.

[82] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, "Feature-rich part-of-speech tagging with a cyclic dependency network," in *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 2003, pp. 252–259.

[83] A. P. B. Veyseh, N. Nour, F. Dernoncourt, Q. H. Tran, D. Dou, and T. H. Nguyen, "Improving aspect-based sentiment analysis with gated graph convolutional networks and syntax-based regulation," *arXiv preprint arXiv:2010.13389*, 2020.

[84] C. Wan, Y. Peng, K. Xiao, X. Liu, T. Jiang, and D. Liu, "An association-constrained lda model for joint extraction of product aspects and opinions," *Information Sciences*, vol. 519, pp. 243–259, 2020.

[85] K. Wang, W. Shen, Y. Yang, X. Quan, and R. Wang, "Relational graph attention network for aspect-based sentiment analysis," *arXiv preprint arXiv:2004.12362*, 2020.

[86] T. Wang, K. Lu, K. P. Chow, and Q. Zhu, "Covid-19 sensing: Negative sentiment analysis on social media in china via bert model," *Ieee Access*, vol. 8, pp. 138 162–138 169, 2020.

[87] W. Wang, S. J. Pan, D. Dahlmeier, and X. Xiao, "Recursive neural conditional random fields for aspect-based sentiment analysis," *arXiv preprint arXiv:1603.06679*, 2016.

[88] J. Wei, J. Liao, Z. Yang, S. Wang, and Q. Zhao, "Bilstm with multi-polarity orthogonal attention for implicit sentiment analysis," *Neurocomputing*, vol. 383, pp. 165–173, 2020.

[89] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources and Evaluation*, vol. 39, no. 2-3, pp. 165–210, 2005.

[90] Z. Wu, Y. Li, J. Liao, D. Li, X. Li, and S. Wang, "Aspect-context interactive attention representation for aspect-level sentiment classification," *IEEE Access*, vol. 8, pp. 29 238–29 248, 2020.

[91] W. Yang, L. Mu, and Y. Shen, "Effect of climate and seasonality on depressed mood among twitter users," *Applied Geography*, vol. 63, pp. 184–191, 2015.

[92] L. Yao, C. Mao, and Y. Luo, "Graph convolutional networks for text classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 7370–7377.

[93] H. Ye, Z. Yan, Z. Luo, and W. Chao, "Dependency-tree based convolutional neural networks for aspect term extraction," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining.* Springer, 2017, pp. 350–362.

[94] Z. Ye, F. Li, and T. Baldwin, "Encoding sentiment information into word vectors for sentiment analysis," in *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 997–1007.

[95] Y. Yin, F. Wei, L. Dong, K. Xu, M. Zhang, and M. Zhou, "Unsupervised word and dependency path embeddings for aspect term extraction," *arXiv preprint arXiv:1605.07843*, 2016.

[96] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, “A survey of sentiment analysis in social media,” *Knowledge and Information Systems*, pp. 1–47, 2019.

[97] N. Yussupova, M. Boyko, D. Bogdanova, and A. Hilbert, “A decision support approach based on sentiment analysis combined with data mining for customer satisfaction research,” *Int. J. Adv. Intell. Syst*, vol. 8, pp. 145–158, 2015.

[98] C. Zhang, Q. Li, and D. Song, “Aspect-based sentiment classification with aspect-specific graph convolutional networks,” *arXiv preprint arXiv:1909.03477*, 2019.

[99] W. Zhang, H. Xu, and W. Wan, “Weakness finder: Find product weakness from chinese reviews by using aspects based sentiment analysis,” *Expert Systems with Applications*, vol. 39, no. 11, pp. 10 283–10 291, 2012.

[100] C. Zhao, S. Wang, and D. Li, “Multi-source domain adaptation with joint learning for cross-domain sentiment classification,” *Knowledge-Based Systems*, vol. 191, 2020. [Online]. Available: https://doi.org/10.1016/j.knosys.2019.105254

[101] J. Zhou, J. X. Huang, Q. V. Hu, and L. He, “SK-GCN: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification,” *Knowledge-Based Systems*, vol. 205, 2020. [Online]. Available: https://doi.org/10.1016/j.knosys.2020.106292

*Received on February 23, 2021*
*Accepted on August 21, 2021*