



# KnowMIS-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis

Giuseppe D’Aniello<sup>1</sup> · Matteo Gaeta<sup>1</sup> · Ilaria La Rocca<sup>2</sup>

Published online: 9 January 2022  
© The Author(s) 2022

## Abstract

The analysis of the opinions of customers and users has been always of great interest in supporting decision-making in many fields, especially in marketing. Sentiment analysis (SA) is the umbrella term for techniques and approaches that analyze user’s sentiments, emotions, opinions in text or other media. The need for a better understanding of these opinions paved the way to novel approaches that focus on the analysis of the sentiment related to specific features of a product, giving birth to the field of aspect-based sentiment analysis (ABSA). Although the increasing interest in this discipline, there is still confusion regarding the basic concepts of ABSA: terms like sentiment, affect, emotion, opinion, are used as synonyms while they represent different concepts. This often leads to an incorrect analysis of the users’ opinions. This work presents an overview of the state-of-the-art techniques and approaches for ABSA, highlighting the main critical issues related to current trends in this field. Following this analysis, a new reference model for SA and ABSA, namely the KnowMIS-ABSA model, is proposed. The model is grounded on the consideration that sentiment, affect, emotion and opinion are very different concepts and that it is profoundly wrong to use the same metric and the same technique to measure them. Accordingly, we argue that different tools and metrics should be adopted to measure each of the dimensions of an opinion. A qualitative case study, regarding product reviews, is proposed to motivate the advantages of the KnowMIS-ABSA model.

**Keywords** Sentiment analysis · Aspect-based sentiment analysis · Affective reaction · Opinion mining

---

✉ Matteo Gaeta  
mgaeta@unisa.it

Giuseppe D’Aniello  
gidaniello@unisa.it

Ilaria La Rocca  
ilialarocca0@gmail.com

<sup>1</sup> Knowledge Management and Information Systems Lab (KnowMIS), Department of Information and Electrical Engineering and Applied Mathematics, University of Salerno, 84084 Fisciano, SA, Italy

<sup>2</sup> Department of Industrial Engineering, University of Salerno, 84084 Fisciano, SA, Italy

## 1 Introduction

The study and analysis of the sensations, emotions, sentiments, opinions, judgments, and evaluations contained in digital recordings (like text, audio, video) are among the main objectives of the sentiment analysis (SA). Specifically, with the term Sentiment Analysis, we refer to the field of study and research which analyzes sensations, emotions, sentiments, opinion, viewpoints, judgments, evaluations, attitudes, behaviors that people have with respect to entities (like products, services, organizations, individuals, events, topics) and their attributes (Liu 2012). Nowadays, this analysis involves any possible media that people use to express their opinion like writing, speech, movements, facial expression, actions, and so on Yadollahi et al. (2017). However, still today, among all the possible media, the main field of application of SA is the analysis of the natural language. In this field, the SA can be seen as the set of activities aimed at analyzing and extract opinions, judgments, emotions, and sentiments contained within a text, written by a user regarding a specific entity. Currently, the SA is applied to a plethora of different domains like, for instance, marketing (products and services), political elections, tourism and cultural heritage (Casillo et al. 2020, 2019), healthcare, security, social events, education (Chang et al. 2020; Colace et al. 2014), smart city (D'Aniello et al. 2020).

From the point of view of marketing applications, despite the numerous techniques and applications that have been realized, the traditional approaches of sentiment analysis still present some limitations. Indeed, the analysis of sentiment does not allow us to analyze, with a sufficient level of detail, the specific aspects, components, parts, or functionalities of the entity that have been judged and commented on by the end-users. To obtain a finer analysis that could make evident the specific aspects of a product or a service the users liked or disliked the most, novel approaches and techniques are needed. In this case, the opinions of the users have to be analyzed for each specific aspect of the product or service: this kind of analysis is called aspect-based sentiment analysis (ABSA).

In recent years, ABSA received great attention from the scientific community with a proliferation of original and novel techniques and approaches that have their roots in the SA field. The main tasks of the ABSA consist of: the identification of the aspects of an entity from a text; the extraction of the linguistic expression used in the text to refer to the aspects; the identification of the opinion or sentiment polarity for each aspect (Pontiki et al. 2016). Unfortunately, many of the techniques and approaches proposed so far, consider a different interpretation of what is “sentiment” and how it should be measured. This situation generates a lot of confusion between those that are very different concepts (e.g., sentiment, affective reaction, judgment, emotion) which are often treated in the same way by the same approach. Such a problem has been widely discussed and highlighted also by scholars of other disciplines (social sciences, economy, psychology, etc.) as a serious limitation of the current computational approaches adopted for real-life applications. Considering the affective reactions as sentiments, for instance, has a great impact in terms of wrong evaluations on the loyalty of customers from the viewpoint of marketing applications. To date, the technical aspects and the consolidated approaches used for realizing applications of sentiment analysis seems to receive much attention and to prevail in the research agenda, by accepting on the other hand simplifications with respect to the real differences between the concepts. Such simplifications, which probably could have had a rationale in the pioneer works on sentiment analysis

to be able to tackle the problem, are not anymore needed thanks to the advancements in computer science and in the related disciplines. Rather, continuing to use such simplifications, does not allow a real leap forward and a change of pace of the research in this area.

The main goal of such works is to detect the users' attitude towards an entity or its aspects. Therefore, they reduce the whole variety of human sentiments to these two classes: positive of negative sentiment, due to their specific goals. This approach, in principle, is correct as it answers their specific goal. However, recent advances in artificial intelligence and sentiment analysis research areas, together with the emerging needs of business companies and marketing strategies that require a finer analysis of user opinions, require a finer analysis that cannot reduce the whole human sentiments to only two classes (positive and negative). Sentiments, opinions, emotions, affects, orientations, are very different concepts and should be interpreted and treated separately with different approaches. Their correct interpretation is critical from the point of view of marketing applications because an incorrect interpretation may lead to incorrect results and consequently to wrong actions. Furthermore, the different interpretations used by these approaches make also it very difficult to compare these techniques and to use them in real-world applications. According to us, and as noted in several works in literature (Munezero et al. 2014; Yadollahi et al. 2017) (as will be discussed in Sect. 3), sentiment is a lasting, durable emotional disposition that is developed by the user in a long period of time with respect to an entity or one of its aspects. Therefore, it is very different from an immediate declaration of an affect, which has not matured during the time but rather is a reaction to an entity or an aspect: we refer to these with the term affective reaction. As a consequence, these two different kinds of elements (sentiment, and affective reaction) should be treated in a different way as they will lead to different decisions in marketing contexts. Sentiment will need time to be developed, therefore it could be an indication of loyal customers, and it requires long-term marketing strategies, while affective reactions require completely different, more immediate, strategies. Section 3 provides further details on these definitions.

What seems to lack is a formal definition of the fundamental concepts of ABSA (although some works have proposed important definitions and clarified such concepts, as we will report in the next section). In such a context, this work proposes an overview of the literature in the field of SA and ABSA which identifies the main families of techniques and approaches and the interpretation they gave to the fundamental concepts of ABSA. Furthermore, this work proposes a new model for ABSA, namely the KnowMIS-ABSA model, with the aim of providing a shared model on which further approaches and techniques can be grounded. The KnowMIS-ABSA model recognizes the importance of some recent advancements in the field and tries to harmonize and integrate such results in a new formal representation that can be used for novel ABSA techniques. The model underlines the main differences between the ways an opinion can be expressed (sentiment, reactions, emotion, affect) while providing a common representation that allows dealing with these different dimensions in an integrated way. The model, indeed, allows for representing the results of a sentiment analysis technique at a different level of detail (document, sentence, and aspect level) by using different metrics, always separating the interpretation of the different dimensions of a sentiment (affective reactions, sentiment, opinion). The proposed model is both technology and techniques agnostic, in the sense that it can be adopted and used to represent the information regarding sentiments and reviews in different approaches. Prior to the presentation of the proposed model, in what follows, we describe the main reasons and limitations that motivate this work.

## 1.1 Motivations

In this work, we mainly refer to the analysis of sentiment of opinionated texts in which users express their opinion regarding a product, a service, or an event, by writing a review. A model for ABSA has to define the different dimensions of sentiment, by defining concepts like emotion, affect, opinion. The analysis of the literature review, presented in the next section, evidences the many different interpretations of sentiment. Those definitions lead to different ways of measuring the same thing or, on the other side, to measure different aspects with the same metric. This situation makes evident the need to have a shared definition of the basic concepts of ABSA, which is among the objectives of our work. Alongside the definitions of the basic concepts of ABSA, it is also important to identify the metrics that should be used for each different measure. Considering, for instance, that expressing a sentiment towards a product is very different from just having a positive opinion, it is clear that these two elements should be evaluated and measure in different ways. This approximation between sentiment and opinion, which probably could be acceptable to a computer scientist, is not so in other domains like social sciences, psychology, marketing & communications, neurosciences, etc. We consider this aspect as the real limitation to realize successful and trustful applications for the real world.

Another element that a model for ABSA should propose is the characteristics of the document or media that is considered. Usually, when considering a review, the analysis is focused especially on the text written by the users. However, modern websites as well as social networks, offer many features and tools that allow the end-users to create a very complex and detailed review of a product or a service. Besides the main text which is written in natural language, the user typically can add a rating (for instance on a given scale of five stars); add tags to catalog the review; add some emoticons to express their opinions, and so on. Traditional ABSA techniques usually take into account only some elements of the above-mentioned. More recent approaches consider also rating scales associated with the review (Nakov et al. 2016; Rosenthal et al. 2017), which transforms the sentiment classification problem from a binary classification to ordinal classification. Other approaches consider elements like emoticons (Wang and Castanon 2015), domain-specific language (Jiaxiang 2020) and jargon (Aly and van der Haar 2020), and so on. Therefore, it is evident that a new model, from the point of view of ABSA, for taking into account all these additional aspects to the plain text of the reviews, is needed.

## 1.2 Contributions

The main contributions of this work can be summarized as follows: (1) an overview of the modern approaches and techniques for SA and ABSA; (2) an analysis of current definitions of the basic concepts of SA and ABSA; (3) the KnowMIS-ABSA model, a reference model for ABSA which is grounded on the major results of the recent literature; iv) a qualitative case study, using a public dataset, to highlight the capabilities of the proposed model in the analysis of the sentiment. Let us underline that the KnowMIS-ABSA model is intentionally generic and abstract in order to be used with different techniques (from lexicon-based to learning-based), and it can be applied to the different levels of analysis (from document level to aspect level). The aim of this model is not to propose novel techniques for ABSA, rather it represents a reference to guide the use or the new definitions of techniques taking into account the deep differences between sentiments, affective reactions, and statements.

### 1.3 Organization

The rest of the article is organized as follows. Section 2 contains an overview of the state-of-the-art techniques and approaches to SA and ABSA. Section 3 describes the KnowMIS-ABSA model. Section 4 contains the qualitative case study with a comparison between a traditional approach to ABSA and the new model. Section 5 concludes the work with final discussions and future work.

## 2 Sentiment analysis and aspect-based sentiment analysis: an overview

### 2.1 Sentiment, opinion, emotion and affect

In the last twenty years, the research on Opinion Mining (OM) and Sentiment Analysis (SA) has reached a high level of maturity, with wide adoption in many (also commercial) applications. Despite the great ferment and innovation in this research field, together with the great interest from the business world in such kind of applications, there is still a lot of confusion regarding what is the meaning of OM and SA and what are the differences (if any) between the concepts of opinion, sentiment, emotion, affect, and related terms. Even though one may think that this is just a terminological debate and that there is not a real difference between the above-mentioned concepts, an increasing number of researchers argue that not only such concepts are profoundly different, but they must be treated using completely different approaches as they lead to different kind of results. Specifically, while almost all the scholars agree on the fact that the tasks of Opinion Mining and Sentiment Analysis regards understanding the subjectivity embedded in a text (or other multimedia content), they have different opinions on what is the meaning of opinion and sentiment.

Some researchers consider OM and SA as synonyms. For instance, in the work of Farhadloo and Rolland (2016), SA and OM are considered as synonyms and the tasks of SA and OM are treated in the same way. In particular, the word “sentiment” is associated with a personal experience which leads to having a certain opinion on a specific topic. Haider et al. (2021) affirm that feelings are expressed by opinions and emotions that usually are collectively defined as sentiments. A similar vision is shared also by Liu (2012) and by Saad and Saberi (2017), which considers the two terms as belonging essentially to the same concept. Medhat et al. (2014) also affirm that the term SA and OM are interchangeable but they recognize that for some researchers the terms have a slightly different meaning. They consider the OM as the activity of extraction and analysis of the opinion regarding an entity while the SA is defined as the identification of a sentiment expressed in a text, by considering opinion and sentiment as two separate concepts. In Novielli et al. (2020), the authors define their task of sentiment analysis applied to the Software Engineering field as the identification of six basic emotions: love, joy, anger, sadness, fear, and surprise, where, therefore, sentiment and emotion can be seen as synonyms (or at least very strictly related concepts). Moreover, they identify the need for defining a theoretical model of affect on which the sentiment analysis should be ground, from which it seems to emerge that also affect can be considered as a similar concept to emotion and sentiment.

Our vision is different as we consider opinion and sentiment as two profound different concepts. Such an idea is indeed shared by many recent works as Munezero et al. (2014),

Yadollahi et al. (2017). In what follows, we explain why according to us the two concepts are different, comforted by several works that we will analyze and review. Kim and Hovy (2004), for instance, affirm that an opinion can be subjective without necessarily implying a sentiment. Indeed, an opinion could also not convey a sentiment. On the other hand, Yadollahi et al. (2017) propose a partition of SA into two categories: Opinion Mining, which is considered as a specialization of SA related to the analysis of the neutral, positive or negative opinion in the text, and Emotion Mining, which is the analysis of the emotions in the text. Furthermore, we not only consider sentiment and opinion as different concepts, but we think that it is critical to analyze the different facets of the sentiment, by identifying and defining the key elements as emotion, affect, orientation, etc. Even in this regard, recent works tried to analyze and define such elements.

The work of Munezero et al. (2014) proposes an enlightening discussion, funded on physiological and psychological studies, which tries to eventually establish a fixed frame of definitions for sentiment analysis. We consider these definitions as the foundational pillars on which we construct and define a novel reference model for ABSA, described in Sect. 3. In the cited work, opinions are considered as personal interpretations of information about a topic, while sentiments are prompted by emotions. As an example, an opinion could be: (1) “the battery of the smartphone has a good capacity”, while a sentence containing a sentiment could be: (2) “I always loved the longevity of the battery of all the smartphones of this manufacturer, since the first model”. Considering this difference, it appears clear that an opinion can also be expressed without any emotion, for instance in a descriptive way, as in the sentence: (3) “In my experience, the average duration of the battery is 8 hours with a normal use”. Sentiment can be also considered a social construct of emotions that develop over time and are enduring, meaning that the temporal aspects are critical for the sentiment, as is evident from the previous example. Opinions, on the other hand, are just personal interpretations of facts that may or may not be emotionally charged (in the above example, the first sentence contains an opinion with a positive orientation, while the third sentence contains an opinion without orientation or with a neutral orientation). And even when an opinion expresses a certain kind of subjectivity and judgment, it does not necessarily imply that there is a sentiment (Kim and Hovy 2004; Sokolova and Bobicev 2011). Indeed, considering again the work of Novielli et al. (2020), although they seem to consider as similar concept affect, emotion and sentiment, they affirm that a sentence that contains a negative opinion does not always have to contain an emotion as well, providing some examples supporting this claim. From this consideration, the authors conclude that it is important to identify which is the desired output of a SA tool, whether it is to classify an opinion or to identify emotion in the text.

Continuing the literature analysis with the scope of defining the groundings on which our model will be defined in the next section, here it is important to define the concept of emotion, considering that sentiments are usually prompted by emotions (Loia and Senatore 2014). Munezero et al. (2014) define emotion as social expressions of feelings and affect influenced by culture. Feelings are defined as person-centered psychological and physiological sensations and as affective phenomena to which we have conscious access. On the other hand, emotions are typically preconscious phenomena. Lastly, affect is a more abstract concept that may encompass all the other concepts of emotions, feelings, and sentiments. According to these authors and other scholars, affect is typically difficult to express (and therefore to find) in language. Usually, what we can identify in language and text are not the affect, but the emotions (which can be considered as the expression of affect) or the so-called affective reactions. The affective reactions are the conscious representation of affect (for instance, the fact that we consider an object as beautiful or awful; or

that we have an immediate good or bad reaction to a fact). Notice that an affective reaction is completely different from the sentiment, as the sentiment is an enduring phenomenon and a social construct related to emotions; while the affective reaction is usually more limited in time and could be related to the affect and the feelings.

According to this analysis, our vision, sustained by the recent literature, leverages the deep differences among the concepts of sentiment, opinion, emotions, and affective reactions. Consequently, the techniques to use in order to determine their values should be substantially different, as such techniques should look for different aspects of subjectivity. For instance, when we are looking for analyzing the sentiments, we are looking for the polarity (positive or negative) of an enduring experience towards an object, a fact, or an aspect. When we are looking for an opinion, instead, we would like just to understand which is the idea or judgment (which could be also neutral, and which could not contain emotion or sentiment) regarding an object or a fact. In Sect. 3 we will propose our definitions of sentiments, affective reactions, and opinions (conveyed in what we call statements) as the main elements of the proposed model. As discussed, such definitions will be our original synthesis and elaboration of the recent literature.

## 2.2 Sentiment analysis: levels of analysis

The analysis of the sentiment, or in general of the opinion, can be performed at different levels of detail, from the most generic to the most specific: document-level, sentence-level, aspect-level, and concept-level (Hemmatian and Sohrabi 2019; Pontiki et al. 2016; Nakov et al. 2016; Rosenthal et al. 2017). The document-level analysis [sometimes referred to as text-level (Pontiki et al. 2016)] aims at understanding the polarity of a whole document [e.g., reviews (Behdenna et al. 2018), news article (Shirsat et al. 2017), a post, a tweet (Gurini et al. 2013)]. Therefore, the information is quite general since it summarizes the polarity of many sentences as a unique positive or negative score [usually measured on a two-point or five-point scale (Nakov et al. 2016)]. Lexicon-based approaches for document-level analysis typically identify the polarity of each term and then aggregate their polarity scores to classify the sentiment of the whole document. Considering the level of importance of each word in this kind of approach could improve the accuracy of the classification. Asif et al. (2020) propose an approach to identify the sentiment of posts and comments on social networks by considering the sentiment polarity of each term and by weighting such polarity using a degree of importance evaluating the  $tf - idf$  measure of the term.

A finer analysis is provided by sentence-level SA (Appel et al. 2016). The aim is to understand if a sentence has a positive or negative opinion. Nowadays, there is not a strong difference between the techniques used for document or sentence level, although the latter problem is usually harder, especially for short sentences and sentences without a surrounding context. In fact, as (Neviarouskaya et al. 2007) and (Yadollahi et al. 2017) pointed out, the same sentence used in two different contexts could express two opposite sentiments. Usually, a sentence level approach can be used also to analyze the sentiment of a complete document by analyzing its sentences and using an aggregation operator to combine the polarity of each sentence to obtain the global polarity of the document (Fang and Zhan 2015; Rao et al. 2018).

An analysis at the document or sentence-level provides useful insights regarding the opinion of users about specific entities, but they do not allow to understand which elements, features, or aspects of such entities have influenced the (positive or negative) opinion of the

users. To perform such a kind of analysis, aspect-level (or aspect-based) sentiment analysis should be performed (Liu 2012). This level of analysis clearly provides a more accurate result if compared with document and sentence-level analysis. An interesting description of the main tasks and subtasks of ABSA has been proposed in the SemEval competition (the International Workshop on Semantic Evaluation) and in particular in the work of Pontiki et al. (2016). In what follows, we will refer to many of these tasks to both analyze the literature and propose the process in Sect. 3.1. According to this work, typically, an aspect-based task analysis technique foresees the following tasks: (1) identification of the opinionated/subjective sentences from the document; (2) identify the entity (also called aspect category) for each sentence; (3) identify the aspect of the entity to which the sentence refers to; (4) extract the linguistic expression used in the sentence to refer to the aspect; (5) for each identified aspect, identify the polarity (e.g., on a binary scale, five-point scale, or real number in a range)(Pontiki et al. 2016).

Let us underline that the task of ABSA can be performed at the aspect-level, where each portion of sentence referred to the single aspect is considered; at the sentence level, where a whole sentence is analyzed in order to understand the opinion with respect to each aspect contained in it (there could be more than one fragment of the sentence referring to the same aspects); and at document-level or text-level, where the whole document is processed and a overall analysis for each aspect is performed (Pontiki et al. 2016; Nakov et al. 2016; Rosenthal et al. 2017).

Recently, a finer level of analysis has been introduced by Cambria (Cambria 2013; Cambria et al. 2013) with the concept-level sentiment analysis. The idea is to have machines able to deeply understand the natural language and in particular, to understand emotions, opinions, sentiments. Such kind of approaches is based on semantic analysis techniques able to identify and analyze the concepts in a text. SenticNet Cambria et al. (2020) is a lexical resource that has been proposed to assign emotion labels and mood tags to support concept-level sentiment analysis with interesting results.

### 2.3 Aspect-based sentiment analysis

The first task of ABSA is the aspect extraction (or identification). Several works, like Zhang and Liu (2014), Rana and Cheah (2016), Hemmatian and Sohrabi (2019), proposed the analysis and classification of the techniques for aspect extraction. Specifically Rana and Cheah (2016) proposes a comprehensive analysis of recent techniques for aspect extraction, identifying three main categories: unsupervised, semi-supervised, and supervised approaches. In the first category, we can find approaches based on: frequency or statistics (Hu and Liu 2004; Bafna and Toshniwal 2013; Rana and Cheah 2018; Wang et al. 2015; Luo et al. 2015); heuristics like (Singh et al. 2013) or the work of (Bancken et al. 2014) that uses a syntactic dependency path to identify entities or (Poria et al. 2014) that adopts a rule-based approach. In the semi-supervised category, we can find techniques based on lexicon (Yan et al. 2015; D'Aniello et al. 2018; Shah and Swaminarayan 2021; Klyuev and Oleshchuk 2011) or dependency trees (Yu et al. 2011) and graphs (Xu et al. 2013). Supervised techniques typically use machine learning approaches like random field, SVM (Manek et al. 2017), decision trees, neural networks, and autoencoders (Angelidis and Lapata 2018; Tomasiello 2020).

The next task is sentiment identification (or polarity identification) which requires that a sentiment score (in terms of polarity or orientation) is identified for each aspect extracted from the text (Pontiki et al. 2016). Typically, the score is in a numerical range which

indicates the intensity of positiveness or negativeness of the opinion regarding the aspect (for instance, on a five-point scale (Nakov et al. 2016), or as a decimal value in the range  $[-1; +1]$  where  $-1$  is really negative,  $+1$  is really positive). In case the objective is to understand only if the sentiment is positive or negative, the task is also called sentiment classification.

Many techniques have been proposed for sentiment identification, which can be classified into three categories (Schouten and Frasincar 2016): (1) lexicon-based approaches; (2) machine learning approaches; (3) hybrid.

### 2.3.1 Lexicon-based

The lexicon-based approaches for ABSA use lexical resources, like a dictionary, to find the scores for the sentiment words identified in the text. Typically, they also use an aggregation step to combine the scores of the single words together with mechanisms that take into account grammar rules and surrounding words to assign the final score to each aspect.

The approach proposed by Hu and Liu (2004) uses a semantic dictionary obtained from the Wordnet dictionary. The dictionary considers only adjectives and uses a binary score for the sentiment. The approach computes a sentiment score for a sentence and then simply assigns the same score to all the aspects of that sentence. The work proposed in Zhu et al. (2009), instead, splits the sentence into segments, each containing one aspect. A sentiment lexicon is used to identify the polarity of each segment and thus to each aspect. Mowlaei et al. (2020) propose an approach to adapt lexicons used for sentiment analysis to the specific problem of ABSA. In particular, they use statistical methods and a genetic algorithm to extend two lexicon generation methods for the aspect-based problem. This allows the authors to obtain good experimental results outperforming existing methods. Alexopoulos and Wallace (2015) propose an approach for creating and using domain-specific lexicons based on an ontology containing aspect-polarity relations that can be used in domains where it is difficult to find a suitable lexicon.

### 2.3.2 Machine learning

Supervised and unsupervised machine learning techniques are widely used for the identification of the sentiment score in the ABSA. Some approaches also use a lexicon as part of the training data for machine learning methods.

Traditional machine learning techniques, like Naive Bayes (Mubarak et al. 2017; Anand and Naorem 2016), Decision Trees/Random Forest (Fitri et al. 2019; Karthika et al. 2019) and SVM (Pannala et al. 2016; Wilson et al. 2005; Tripathy et al. 2016) have been widely used in the past. Esuli et al. (2020) proposes an approach for cross-lingual sentiment quantification based on Structural Correspondence Learning (SCL) which is a technique that can be applied to different kinds of classifiers to transfer knowledge through a mapping between pivot terms of two feature spaces. More recently, given the increasing trend and diffusion of deep learning approaches in all the fields of computer science, many deep learning techniques are used for sentiment identification. Convolutional Neural Network (CNN) for instance is used in Cahyadi and Khodra (2018); Maglogiannis et al. (2020); also approaches based on LSTM (Wang et al. 2016; Yu et al. 2019; Bao et al. 2019; Xing et al. 2019) and Recurrent Neural Networks (Al-Smadi et al. 2018; Nguyen and Shirai 2015; Colace et al. 2019) have been proposed. A comprehensive review of deep learning

approaches used for sentiment analysis (not only for ABSA) can be found in Yadav and Vishwakarma (2020).

### 2.3.3 Hybrid

Several works propose hybrid approaches that aim at combining the advantages of lexicon-based and machine learning (especially deep learning) approaches to ABSA. Meškelić and Frasincar (2020) propose a solution combining a lexicalized domain ontology with a neural attention model for ABSA. Asif et al. (2020) propose a hybrid approach to classify sentiment in social media posts. Specifically, the authors use a domain-specific multilingual lexicon to annotate posts and comments containing extremism sentences. Then, a machine-learning based approach is used to classify the sentiment of posts and comments published on social networks. This technique is able to reduce the effort and time needed to annotate large textual corpora to train classifiers. Mumtaz and Ahuja (2018) propose an architecture that uses a lexical resource (an English dictionary) which is given in input to an SVM. Pekar et al. (2014) combine a lexical resource with a deep linguistic processing technique. (Bao et al. 2019) integrate an attention network with a lexical resource for ABSA, while Zeng et al. (2019) combines a convolutional neural network with lexical resources.

## 3 A reference model for aspect-based sentiment analysis: KnowMIS-ABSA

In this section, we propose a new reference model, the KnowMIS-ABSA model, whose objective is to bring more clarity in the definition of the main concepts related to the ABSA. The main aspects that have driven the definition of the model are:

- preliminary works which recognize the importance of clarifying what is a sentiment, an opinion, an emotion, and the main difference between SA and OM, like the work of Munezero et al. (2014), but these do not provide a formal model for SA;
- state-of-the-art approaches and techniques which provide very interesting results but that could be better interpreted in the light of what are sentiments, opinions, emotions, and affects;
- the need for providing a more recent model of SA that takes into account the characteristics of review and the difference between sentiment, opinion, affect. Many recent approaches still use a definition of a review which has been proposed 10 years ago (Liu 2012). More recent approaches start introducing relevant concepts, as the helpfulness of the review Fang and Zhan (2015). The Web in recent years has undergone a terrific evolution. We cannot consider a review just as a traditional opinionated text: it contains more than this. It contains emoticons, likes and dislikes, comments made on an original review, number of views or web shares, etc.

The KnowMIS-ABSA model consists of the following definitions.

**Definition 1** (*Entity*) An entity  $e_k \in E$  is any abstract and/or real subject of the real or virtual world which is of interest for a user  $u_m$ ; there is at least a digital record  $rec_j \in REC$  involving the entity  $e_k \in E$  created by  $u_m \in U$ .

**Definition 2** (*User*) A user  $u_m \in U$  interacts with a website or a social platform or another kind of virtual community; his/her interactions are represented by records  $rec_j \in REC$ .

**Definition 3** (*Record*) A record  $rec_j \in REC$  is the digital representation of an interaction of a user  $u_m \in U$  with the digital platform; it consists of one or more symbols, words, sentences, texts, or other digital contents, related to one or more entities  $e_k \in E$ , created by the user in a specific time instant, and which may contain the expression of:

- emotions: social expressions of feelings and affect influenced by culture;
- sentiments: social constructs of emotions that develop over time and are enduring;
- opinions: personal interpretations of information about a topic;
- judgments: a personal opinion formed after thinking carefully and based on grounds that could be also incorrect; it usually contains an orientation

owned by an opinion holder  $h_l \in H$  (Munezero et al. 2014).

**Definition 4** (*Opinion holder*) An opinion holder  $h_l \in H$  is the subject that has a specific opinion (or sentiment, emotion, judgment, etc.) on the entity or its aspects. Such an opinion can be expressed by the holder himself (in this case, the user and the holder are the same person) or another user. To make the difference between user and opinion holder clear, consider the following review written by the user John: “My wife Mary thinks this smartphone is exceptional. I don’t think so, I hate it”. The user  $u_m$  that written the review is John. The opinion holder  $h_{l1}$  of the first sentence is Mary, the opinion holder  $h_{l2}$  of the second sentence is John himself (so  $h_{l2} \equiv u_m$ ).

**Definition 5** (*Aspect*) An aspect  $a_t^{e_k} \in A^{e_k}$  is a component, a part, a functionality, or an attribute of the entity  $e_k \in E$ ; the set  $A^{e_k}$  contains all the aspects of the entity  $e_k$ . Moreover, the aspects could be also hierarchically organized, belonging to different aspect category, as in Pontiki et al. (2016), Rosenthal et al. (2017).

Using the above definitions, it is possible to give a definition of what is a review.

**Definition 6** (*Review*) A review  $R_r \in R \subseteq REC$  is the set of one or more symbols, characters, words, sentences and/or texts, by means of which a user  $u_m$ , in a specific time instant, expresses sensations, emotions, sentiments, opinions, viewpoints, judgments, evaluations, made by himself or other holders, about one or more entities  $e_k \in E$  and optionally on their aspects  $a_t^{e_k} \in A^{e_k}, \forall e_k \in E$ , published on online platforms.

The analysis of a review aims at studying the opinions contained in it. A review consists of multiple sentences, and it may contain different, also contrasting, opinions with affective reactions or sentiments expressed by the opinion holder. To analyze such a review, the KnowMIS-ABSA defines (based on the definitions analyzed in Sect. 3) three different subjectivity forms, or sentiment dimensions (depicted in Fig. 1). The opinion holder expresses his/her opinion along these dimensions in the review. The three subjectivity forms are:

- *Sentiment*: a sentence or an opinion contained in a review may indicate a sentiment of the holder, which is a lasting, durable emotional disposition developed by the holder in a long period of time with respect to the opinion target (the entity or one of its aspects). The sentiments can only be positive or negative. It has no sense to consider a neu-

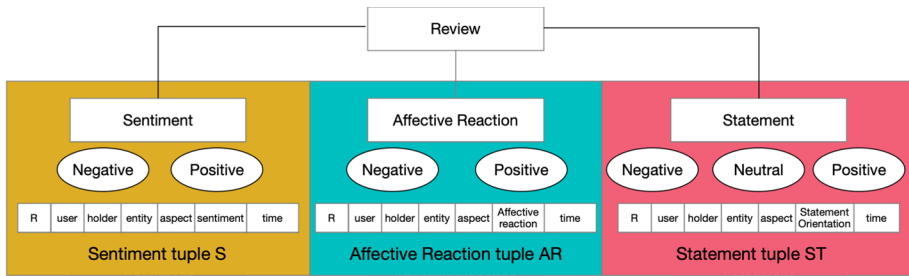


Fig. 1 The KnowMIS-ABSA model

tral sentiment (while, instead, we may have a neutral opinion or a neutral sentence expressed by the opinion holder). Therefore, the score of the sentiment is indicated by a polarity value. The time dimension of the sentiment is critical: the sentiments have a duration greater than an affective reaction or an emotion because they are more stable and are developed over a long period of time. It can be expressed in the natural language (written or spoken) with one or more words or sentences which usually indicates a durable, well-established opinion. An example containing a sentiment is: “I’ve always hated smartphones with oversized screens”

- *Affective Reaction*: a conscious declaration of an affect that describes the way the holder perceives the entity. It differs from the sentiment as it is less stable and less lasting than the sentiment. Like the sentiment, an affective reaction is positive or negative, and cannot be neutral, and therefore the score of the affective reaction is measured with a polarity value. An example containing an affective reaction is: “The screen of this smartphone is fantastic!”
- *Statement (or Affirmation)*: a sentence or an opinion without sentiment or emotion. It is a statement, an affirmation, a claim to describe a fact, to provide information, or to give a judgment. Such statements can include (implicitly) an *orientation* that can be positive, negative, or neutral. The score of a statement is given by an orientation value, which differs from the polarity value as it can be also neutral. An example containing a statement is: “The smartphone has a large screen of 6.1”.

Each of the three elements of subjectivity (Sentiment, Affective Reactions, Statement) has a different meaning, is expressed by the holder in a different form, and represents different aspects of the opinion of the holder. Therefore, they are measured and represented using different metrics. In particular, we use polarity metrics (which contain only positive or negative values) for the affective reaction and sentiment, and orientation metrics (which contain positive, neutral, and negative values) for the statements.

### 3.1 Metrics

For each opinion contained in the review, we can define a tuple that allows us to computationally represent such an opinion along with the three different levels of subjectivity (sentiment, affective reaction, or statement) and to process them. We define the following tuples, depicted in Fig. 1.

- Sentiment tuple  $S$ -tuple :  $(R_r, u_m, h_l, e_k, a_n^k, S_{kn}, t_i)$

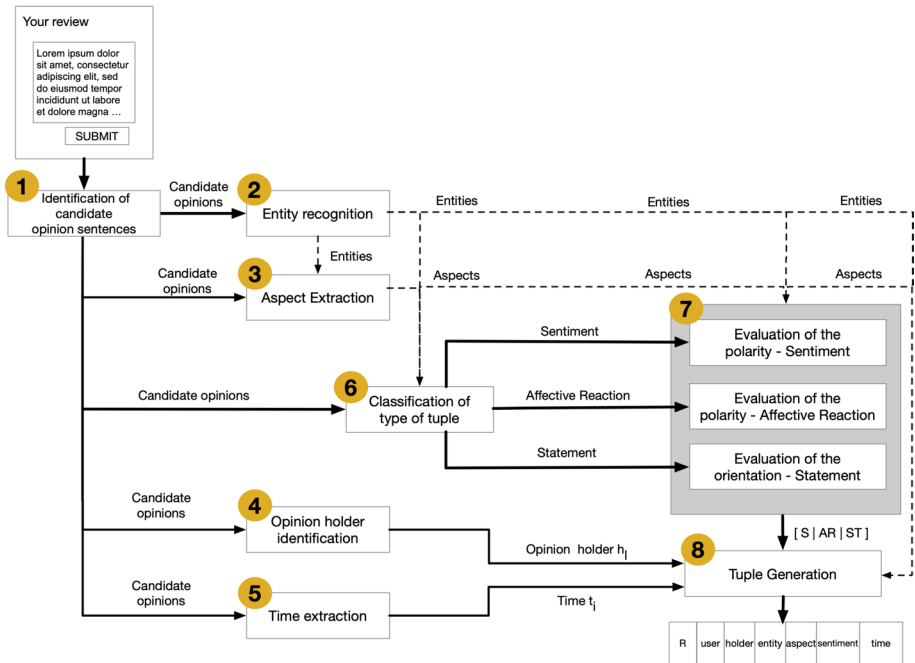


Fig. 2 Abstract process  $\rho(R_r)$  based on KnowMIS-ABSA model to generate tuples from the review

- Affective Reaction tuple  $AR-tuple : (R_r, u_m, h_l, e_k, a_n^k, AR_{kn}, t_i)$
- Statement tuple  $ST-tuple : (R_r, u_m, h_l, e_k, a_n^k, ST_{kn}, t_i)$

where  $R_r$  is the review written by the user  $u_m \in U$ ;  $h_l \in H$  is the holder of the opinion;  $e_k \in E$  is the target entity of the opinion;  $a_n^k$  is the specific aspect belonging to the entity  $e_k$ . If the opinion is related to the whole entity and not to a specific aspect,  $a_n^k$  assumes the special value *GENERAL* to indicate that the opinion refers to the entity.  $S_{kn}$  is the polarity value of the sentiment of the  $S-tuple$  for the aspect  $a_n^k$ ;  $AR_{kn}$  is the polarity value of the affective reaction of the  $AR-tuple$  for the aspect  $a_n^k$ ;  $ST_{kn}$  is the orientation value for the statement of the  $ST-tuple$  for the aspect  $a_n^k$ .  $t_i$  is the time instant in which the opinion has been expressed.

Each review can be split into a set of  $S-tuple$ ,  $AR-tuple$ , and  $ST-tuple$ , one for each opinion related to a specific aspect. The identification of the opinions together with the separation of sentiments, reactions, and judgments/statements sustains the subsequent process of analysis.

In Fig. 2, we propose an abstract process with the main steps that an approach based on the KnowMIS-ABSA model should follow. This is not a concrete implementation of a system, rather it is a generic flow that underlines the main tasks that should be done to complete an ABSA analysis. In particular, the focus of the process is the identification of the different kinds of tuples (sentiment, affective reactions, and statement) and the subsequent different processing using a different technique and metric for each of them. Therefore, the process can be concretely instantiated with any kind of computational technique in each step. However, in this section, we suggest some recent techniques for each step, while in Sect. 4 we propose the implementation of the process we realized for

the case study. Let us define the process  $\rho(R_r)$  as the process that takes a review  $R_r$  as its input and produces three sets as output:  $\mathbf{S}_r$ ,  $\mathbf{AR}_r$ ,  $\mathbf{ST}_r$  that are respectively the set of all the *S-tuple*, *R-tuple* and *ST-tuple* extracted from the review  $R_r$ . The process  $\rho(R_r)$  consists of the following steps:

1. Identification of the candidate opinion sentences: the review is analyzed to identify the sentences that may contain an opinion. This requires the use of data preprocessing techniques (tokenization, stop words removal, part-of-speech tagging, lemmatization) together with a feature extraction task to identify terms, opinion words and phrases, negations, etc. NLP techniques, text mining, and information retrieval approaches can be used for these initial tasks, as analyzed in the recent survey of Birjali et al. (2021).
2. Entity recognition: for each sentence, the entity  $e_k \in E$ , representing the target of the sentence, is identified. This step should also deal with the presence of synonyms of the entity and the indirect references to an entity (e.g., by using pronouns) in the sentence. In SemEval competitions, the importance of entity recognition, or target identification, is reported as a critical task even for ABSA (Pontiki et al. 2016; Rosenthal et al. 2017; Nakov et al. 2016). Both grammar-based (Ding et al. 2018; Basiri et al. 2020; Wiegand et al. 2016) as well as machine learning-based techniques (Gan et al. 2020; Liu et al. 2015; Li et al. 2019) can be used in this step (Liu and Zhang 2012; Khan et al. 2014).
3. Aspect extraction: the aspect targets  $a_n^k \in A^{e_k}$  are identified in this step for each sentence. As for the entity, issues like synonymy, polysemy, and indirect reference have to be addressed by the aspect identification technique. This step is not trivial as some aspects can be also implicit and should be inferred from the text (Ganganwar and Rajalakshmi 2019; Tubishat et al. 2018). Many works are based on natural language processing techniques, using linguistic feature patterns, to identify grammatical classes and syntactic relations of the text to identify aspects (Tubishat et al. 2021). Some approaches use inductive supervised learning algorithms (Rana and Cheah 2016). Other approaches perform aspect extraction and aspect sentiment classification in a unified technique (Akhtar et al. 2020). However, recent approaches are heavily leveraging on deep learning and word embeddings approach for aspect extraction (Do et al. 2019). A recent survey on aspect extraction techniques is proposed in (Nazir et al. 2020).
4. Opinion holder identification: another element of the tuple that should be identified in this step is the opinion holder  $h_l \in H$ . This is an important task for discriminating between opinions that are viewed from different perspectives (Seki et al. 2009; Wiegand et al. 2016). Many approaches have been proposed for this task, and are integrated into many ABSA techniques, ranging from NLP to learning-based approaches (Sima and Vunvulea 2013; Wiegand et al. 2016; Xu et al. 2015; Colhon et al. 2014; Kolya et al. 2012; Wiegand and Klakow 2010).
5. Time extraction: identify the time  $t_i$  in which the opinion, regarding the aspect  $a_n^k$  of the entity  $e_k$ , has been given by the opinion holder  $h_l$  (Liu and Zhang 2012; Tu et al. 2015). This task requires the identification of time expression and to identify tense in order to correctly interpret the time reference of each sentence (Medhat et al. 2014; Min and Park 2012; Preethi et al. 2015). The role of adverbs, especially adverbs of time and comparative adverbs, plays a critical role to identify the time dimension. The work of Haider et al. (2021) proposes an interesting approach to deal with adverbs in sentiment classification, addressing an open research challenge.
6. Classification of the type of tuple: Sentiment, Affective Reaction, Statement. This step is critical, although many existing approaches skip this phase. Before evaluating the

sentence, it should be identified if it contains a sentiment, an affective reaction, or just a statement. Each category should be evaluated with a different technique and using a different metric. The classification could be done with approaches similar to those used for the first step identification of candidate opinion sentences. Some approaches based on machine learning can be adopted as in Carrillo-de Albornoz et al. (2019), Xu et al. (2020) and using text classification techniques (Minaee et al. 2021; Vieira and Moura 2017).

7. Evaluation of the polarity/orientation for each tuple: for each kind of sentence (sentiment, affective reaction, statement) a technique is used to measure the polarity or the orientation. Many different techniques can be used for this task, as discussed in Sect. 3. Recent surveys of sentiment and orientation classification techniques are in Yue et al. (2019), Birjali et al. (2021).
8. Construction of the tuples: the tuples for each sentence are then defined as above described.

## 4 A qualitative case study

The objective of this case study is to highlight the existing differences between affective reactions, sentiments, and statements. The case study is based on the annotated dataset created by Hu and Liu (2004). The dataset contains real reviews on five products extracted from Amazon.com. We evaluate the performances of three existing sentiment analysis tools regarding their capability to evaluate the polarity or orientation of each sentence contained in the dataset.

The aim of the comparison is not to identify the tool which gives the best accuracy in the sentiment analysis task. Rather, to fulfill the objective of this case study, we will compare two cases: (1) the case where each subjectivity sentence is considered as containing a polarity/sentiment (as in traditional approaches of ABSA); (2) the case in which a sentence may contain one of the following: an affective reaction, a sentiment, or a statement (which could be both neutral or containing an orientation) as per the KnowMIS-ABSA model. In the second case, we instantiate some of the steps of the process of Fig. 2, by classifying the sentences as sentiments, affective reactions and statements. Comparing the results, we will see the differences between a traditional approach and the proposed model. What emerges is that the state-of-the-art tools have good performances in the analysis of sentiments and affective reactions, but worse performances when analyzing statements. This is because such tools have been constructed to find sentiments, not the orientation that could be hidden in factual sentences like statements. A finer analysis, by creating ad-hoc tools for such kinds of sentences, could greatly improve the performance of these tools. The case study, therefore, will demonstrate that a difference between these concepts exists and that they should be treated differently.

### 4.1 Data

The dataset “Customer Review Dataset<sup>1</sup>” (Hu and Liu 2004) contains the reviews posted on Amazon.com regarding five electronic devices: Apex AD2600 Progressive-scan DVD

<sup>1</sup> <http://www.cs.uic.edu/~liub/FBS/CustomerReviewData.zip>.



**Table 2** Examples of sentences and annotation provided to the annotators

	Sentence	Class	Polarity/ orientation
1	In summary, I love this camera and would whole-heartedly recommend this to anyone wanting a high quality point and shoot camera in the medium price range.	Sentiment	Positive
2	I bought this camera because it fit my budget and the pre-production and production model reviews were positive .	Statement	Positive
3	Lens cap is a really annoying .	Affective Reaction	Negative
4	The battery life is very good, I got about 90 min with the lcd turned on all the time , the first time around, and I have been using it with the LCD off every now and then, and have yet needed to recharge it .	Affective Reaction	Positive

The annotation procedure was the following. We provided each annotator with the instructions shown in Table 1. We made available to the annotators some examples to let them familiarize themselves with the procedure (reported in 2). Each annotator performed the task independently to avoid influences. For each sentence the annotator should indicate:

- the type of sentence (sentiment, affective reaction, statement);
- in the case of sentiment or affective reaction, the polarity (negative or positive) of the overall sentence; in the case of statement, the orientation (negative, neutral, or positive).
- for each aspect contained in the sentence, the polarity or orientation as above.

The annotator could add a comment to explain the reasons for the made choices and for reporting eventual doubts. Once the annotation process was completed, we proceed with the aggregation and consolidation of the annotations. We discarded the ones wherein the annotator indicates a strong doubt in the comments. In cases where a majority emerged (three out of the five annotators agreed on the classification and the polarity/orientation), we accept the annotation. Otherwise, we adopted a consensus reaching process [proposed by Herrera et al. (1996), Dong et al. (2018) for group decision making] in which the annotators together are asked to discuss among them about the specific sentence and each one explains the reasons behind the previous choices. After this, several rounds of voting are performed, trying to reach a consensus on the final choice. In this case, however, to give more flexibility to the voting process, the annotators express their preferences using utility values. This allows them to express more than one preference but with a different weight (e.g., a sentence could be considered as positive with a utility value of 0.7 and neutral with a value 0.3). Further details on this process are in Herrera-Viedma et al. (2014), Herrera et al. (1996), D’Aniello et al. (2021). If a consensus value of at least 0.8 is reached, the process is stopped and the alternative with the maximum consensus is chosen as the label of that sentence (both for the class and for the polarity/orientation). If the consensus is not reached, another round of discussion and voting is performed. The process ends when the consensus threshold has reached or after three rounds (in this case, we select in any case the choices on which there is the maximum consensus).

The classifications provided by the experts are reported in Table 3. Of the 403 sentences, the experts have identified 16 sentences containing an expression of sentiment; (2) 193 sentences with affective reactions; (3) and 194 sentences containing statements. In this table, each sentence is indicated by three numbers  $x.y.z$  where  $x$  refers to the number (or position) of the review in the Hu-Liu dataset (considering the order by which the reviews appear in the dataset);  $y$  refers to the number (position) of the sentence in each review (considering the order and the splitting performed by Hu and Liu in their dataset);  $z$  refers to the aspect (in the case where more than one aspect is cited in the sentence).

## 4.2 Methods

Three tools for sentiment analysis are used for the case study:

1. Google Cloud Natural Language API (<https://cloud.google.com/natural-language>): the API uses machine learning techniques to support different tasks of natural language analysis, among which entity recognition, content classification, and sentiment analysis. In particular, the library supports both sentiment analysis and aspect-based sentiment analysis, at a document or sentence level. The polarity of the sentiment is given in the range  $[-1; +1]$ .
2. Python Natural Language Toolkit (NLTK <https://www.nltk.org/>): the NLTK is the platform for building Python programs working with natural language. It provides interfaces to over 50 corpora and lexical resources as well as a suite of text processing libraries for classification, parsing, semantic reasoning. In our case study, we used an already available implementation of a sentiment analysis demo tool implemented with Python NLTK (available at <https://text-processing.com/demo/sentiment/>). This demo classifies a text as positive, negative, or neutral. The tool uses a hierarchical classification: first, it decides if the text is neutral or not (by giving also a value between 0 and 1 to the level of neutrality); if it is not neutral, a sentiment polarity is determined (with two values between 0 and 1 for the positive and negative sentiment).
3. ParallelDots Text Analysis API (<https://www.paralldots.com/text-analysis-apis>): ParallelDots offers a text classification and a NLP API that uses machine learning models (trained on more than a billion documents) offering different tasks like sentiment analysis, emotion recognition, entity recognition, semantic analysis. Regarding the sentiment score, it is given as three percentages respectively for positive, neutral, and negative sentiment.

We compare the results of the three tools applied to the sentences of the Hu-Liu dataset. Specifically, each tool is used to perform an analysis of the sentiment at a document level (i.e., the whole review) and at aspect-level. Indeed, in our case, each sentence has only one aspect, and thus even if we analyze a sentence, we are analyzing the sentiment towards a specific aspect. The results of these tools are compared with the annotations of the dataset to obtain a confusion matrix for each tool.

To obtain these confusion matrices for each tool, we need to compare the results with the annotated dataset. Considering that each tool has a different score mechanism for the sentiment, we need to uniform these scores. Moreover, considering that the annotations of the dataset consist only in the discrete values positive or negative, we need to transform the score of each tool to these two discrete values. We perform the following transformations for the score of each tool:

**Table 3** Dataset KnowMIS: Classification of the sentences of the Hu-Liu dataset (Hu and Liu 2004) performed by five experts according to the KnowMIS-ABSA model

Sentiment	POS	1.7.2, 3.4.1, 8.11, 11.6, 12.5, 16.1, 17.1, 20.2.2, 22.1, 22.12, 23.2, 24.17, 24.18, 30.1, 31.4
	NEG	4.8
Affective Reaction	POS	1.1, 1.2, 1.3–1.7.1, 1.8, 1.10, 2.3–2.7, 3.4.2, 3.6, 4.6, 5.15, 6.6.1, 8.2, 8.12, 9.4–9.6, 9.10, 9.15, 11.3–11.5, 11.7–11.9, 11.13, 11.14, 11.17, 12.2, 12.4, 13.2, 13.3, 13.6–13.8, 14.1, 14.2, 14.4, 14.5, 14.8, 15.1, 15.4, 15.5, 16.2, 16.6, 16.8, 16.11, 18.2–18.5, 18.8, 19.1, 19.3, 19.4, 19.5, 20.2.1, 20.18, 20.20, 20.21, 21.1, 21.2, 22.2–22.4, 23.1, 23.3, 24.4, 24.5, 24.7, 25.1–25.5, 26.3, 26.5, 26.6, 27.3, 27.4, 27.7, 27.15, 27.17, 28.1–28.6, 29.2, 30.4, 30.5, 30.9–30.11, 31.2, 31.3, 31.5, 32.3–32.5, 32.8–32.12, 33.2–33.5, 34.1–34.3
	NEG	4.7, 5.3, 5.8, 5.9, 5.11, 5.14, 6.1, 6.3, 6.6.2, 7.1–7.3, 7.5, 7.6, 8.14, 8.15, 8.16, 10.5, 11.5, 13.5, 18.6, 20.6, 20.8, 20.10, 24.8, 24.10, 24.13, 27.5, 31.6, 31.7, 32.14, 32.16,
Statement	POS	2.2, 3.3, 3.5, 4.9, 4.12, 5.1, 5.16, 8.9, 8.10, 8.22, 8.23, 9.2, 9.3, 9.12–9.14, 10.1, 10.11, 11.2, 11.11, 12.3, 13.4, 14.6, 15.3, 16.3, 16.7, 16.10, 16.12, 16.13, 19.2, 20.1, 20.4, 20.16, 20.19, 21.3, 21.4, 24.3, 24.16, 24.19, 24.20, 24.23–24.25, 26.2, 26.4, 26.8, 27.2, 27.13, 27.14, 27.18, 28.8, 30.2, 30.3, 30.8, 31.1, 32.1, 32.2, 32.6, 32.7, 32.15, 32.17, 32.18
	NEU	1.9, 2.1, 3.1, 3.2, 4.1–4.5, 5.13, 7.7, 8.1, 8.3–8.8, 8.18, 8.20, 8.21, 9.1, 9.7, 9.11, 10.2–10.4, 10.6, 10.7, 10.9, 11.1, 11.15, 11.16, 12.1, 13.1, 15.2, 16.4, 16.5, 16.9, 18.1, 19.2, 20.13, 20.15, 20.17, 22.6, 22.11, 24.1, 24.2, 24.11, 24.14, 24.21, 24.22, 25.7, 26.1, 27.1, 27.6, 27.8, 27.12, 28.7, 29.1, 30.6, 30.7, 31.9, 32.19, 32.23, 33.1
	NEG	2.8, 4.10, 4.11, 4.13, 5.2, 5.4, 5.5–5.7, 5.10, 5.12, 6.2, 6.4, 6.5, 7.4, 8.13, 8.17, 8.19, 10.8, 10.11, 11.10, 11.12, 14.3, 14.7, 18.7, 18.9, 20.3, 20.5, 20.7, 20.9, 20.11, 20.12, 20.14, 22.5, 22.7–22.10, 24.6, 24.9, 24.12, 24.15, 24.26, 25.6, 26.7, 27.9–27.11, 27.16, 31.8, 32.13, 32.20–32.22

Each number in the cells indicates a sentence of the dataset Hu-Liu according to this format: x.y.z where x is the number of the review, y is the number of the sentence in that review, z indicates the number of the aspect when a sentence contains more than one aspect and the sentence is classified in different classes for each aspect. If the sentences are classified in the same class, we omit z. The notation x.y1-x.y2 indicates the range of sentences for review x from y1 to y2. Example: 1.7.2 is the seventh sentence in the first review classified according to the second aspect identified by Hu and Liu in their dataset

- Google Cloud NL API: [0.25; 1]: Positive; [- 0.25; 0.25): Neutral; [- 1; - 0.25): Negative.
- Python NLTK: when no polarity is identified by the tool, we consider the value Neutral; in the case the positive polarity has a score greater than the negative one, we consider the Positive value, otherwise the Negative value.
- ParralelDots API: since this tool always gives a score on the three values (positive, neutral, and negative), as a result of the analysis we take the categorical value with the maximum score.

Using this setting, we execute two experiments. First, we compare the tools on the Hu-Liu dataset in which the sentences belong only to two categories: (1) sentences with sentiments and (2) sentences with claims (neutral). Second, we execute the process of Fig. 2 on the annotations of the KnowMIS dataset. For this experiment, the process has been executed in a semi-automatic way, as depicted in Fig. 3. Indeed, considering that we already have the sentences (candidate opinions) with the related entity and aspects in the two datasets, the step 1–3 are not needed as they have been executed by humans. Furthermore, as our aim is to evaluate only the polarity or the orientation of the sentences, we do not need steps 4, 5 and 8 (opinion holder, time extraction and tuple generation). Indeed, we need step 6 (classification of the type of tuples) performed by the experts during the annotation of the KnowMIS dataset (see Sect. 4.1) and step 7 in which the three tools have been used to evaluate the polarity or the orientation. It is clear that in a complete ABSA system, all these steps will be performed automatically (as explained in Sect. 3.1). Specifically, we underline that the Google NL tool is also able to extract both entity and aspects from sentences with good accuracy and to give a polarity for both.

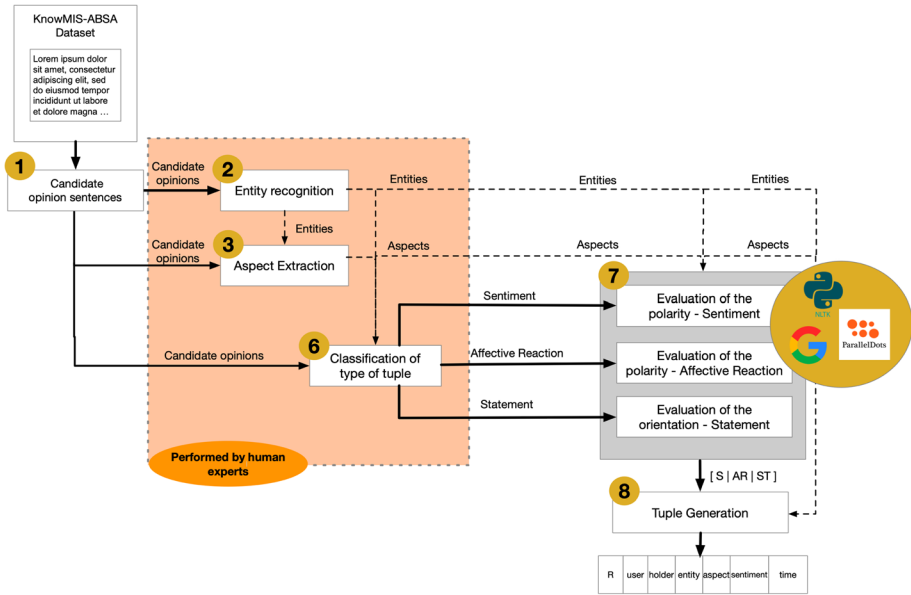
The comparison between the two experiments (the one based on the Hu-Liu dataset and that based on the KnowMIS-ABSA dataset with the above-described process) allows us to qualitatively evaluate the advantages of the proposed model.

### 4.3 Results and discussion

The first set of experiments concerns the evaluation of the three tools to analyze the sentiment of all the sentences of the Hu-Liu dataset at a sentence-level. In this first case, we do not distinguish between affirmations, affective reactions, and sentiment. Table 4 contains the confusion matrices for the three tools. On the rows, there is the classification of the sentences in the three classes (Pos: Positive, Neu: Neutral, Neg: Negative) as annotated in the dataset (i.e., this represents the ground truth). On the columns, for each tool, there is the number of sentences classified as positive, neutral, and negative. Therefore, on the diagonal we have the sentences correctly classified; out of the diagonal, there are the sentences incorrectly classified. The accuracy of each tool computed as the ratio between the sentences correctly classified on the total number of sentences is depicted in Fig. 4a). Figure 4b) shows the F-score for each tool and each class. The f-score for each class is computed as follows:

$$F - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (1)$$

where the precision and recall for a given class is given by:



**Fig. 3** Instantiation of the KnowMIS-ABSA process to the case study. The evaluation of sentiments, affective reactions and statements are performed comparing three tools: Google Cloud NL API, Python NLTK, Parallel Dots

**Table 4** Confusion matrices for each tool applied to all the sentences of the Hu-Liu dataset

All Sentences (Hu-Liu dataset)

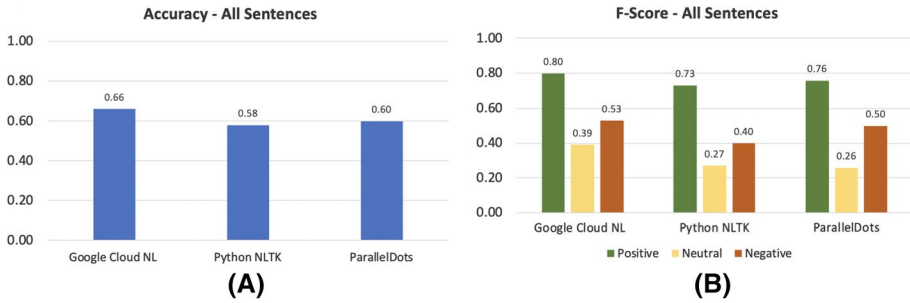
	Google NL API			Python NLTK			ParallelDots		
	Pos	Neu	Neg	Pos	Neu	Neg	Pos	Neu	Neg
Expert									
Pos	192	49	5	178	27	41	176	50	20
Neu	21	35	11	27	17	23	27	23	17
Neg	22	30	38	38	13	39	14	34	42

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

where TP is the number of true positives, FP the number of false positives and FN the number of false negatives.

It can be observed, both from the accuracy and F-Score, that Google NL API obtains the best performance. We can also observe that the performance regarding the neutral class is quite poor. According to us, this could be also related to the fact that the dataset does not distinguish between sentences that may be neutral (like sentences containing statements) and sentences that cannot be neutral for definition (like sentences containing



**Fig. 4** Accuracy and F-Score of the three tools applied to all the sentences of the Hu-Liu dataset

sentiments and affective reactions). This suggests that we cannot process all the sentences in the same way. Moreover, if we observe the confusion matrices and in particular the row containing neutral sentences, we see that all the tools have often classified these as positive or negative. This indeed is in line with the main idea of the KnowMIS-ABSA model: a sentence that contains just a statement or an affirmation can also have an orientation (towards positive or negative) and thus it is not true that they are always neutral. And this, in part, seems to be revealed by the tools.

The second set of experiments regards the analysis of the whole reviews (review-level sentiment analysis) of the Hu-Liu dataset. Table 5 contains the confusion matrices. Figure 5a, b show the accuracy and F-score of the tools applied to the reviews. Again, Google NL API seems to over-perform the other two, but what is interesting to our analysis is the fact that, in the considered dataset, there is only one review that has been considered neutral. When compared with the results of the sentence-level analysis, we can observe how in an opinionated document can coexist sentences that contain an opinion or a form of subjectivity with other sentences that instead are just neutral and do not contain subjectivity. We also have a confirmation that modern tools are better at analyzing the sentiment of an entire review (or a longer text) rather than single sentences (or shorter text).

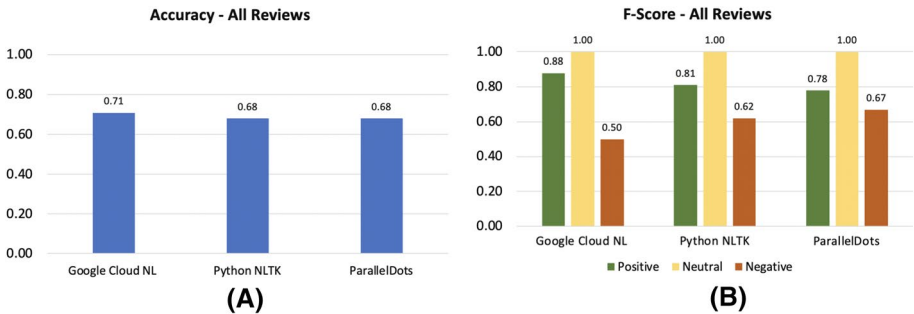
In the last set of experiments, the KnowMIS-ABSA model is applied to classify the sentences in the three classes (sentiment, affective reaction, and statement) before applying the three tools to the analysis of the sentences of the KnowMIS dataset (Table 3).

Table 6 shows the confusion matrices of the three tools applied to only the sentences containing sentiments. Figure 6a, b show the accuracy and F-scores. In this case, the ParallelDots API seems to perform slightly better than the Google Cloud NL API. An interesting result is that all three tools have improved their performance with respect to the case where we consider all the sentences. This is due to the fact that we selected only the sentences that contain sentiment. Usually, such sentences contain sentiment words (love, hate, etc.) to express the sentiment. Therefore, it is easier for all the tools to identify the polarity, considering also that the polarity can not be neutral (so they surely have a polarity, instead of statements that can be also neutral).

Table 7 shows the confusion matrices for the sentences with affective reactions, and Fig. 7a, b the accuracy and F-scores. We observe that Google NL API has slightly better performances, but what is interesting is that again all the tools improve their performances when compared with the case wherein all the sentences are analyzed. Indeed, many affective reactions are quite easy to detect, especially for the presence of affective words and adjectives that indicate a positive or negative polarity.

**Table 5** Confusion matrices for each tool applied to all the reviews of the Hu-Liu dataset (review-level sentiment analysis)

All Reviews (Hu-Liu dataset)									
	Google NL API			Python NLTK			ParallelDots		
	Pos	Neu	Neg	Pos	Neu	Neg	Pos	Neu	Neg
Expert									
Pos	22	5	0	19	6	2	19	6	2
Neu	1	0	0	0	0	1	1	0	0
Neg	0	4	2	1	1	4	2	0	4



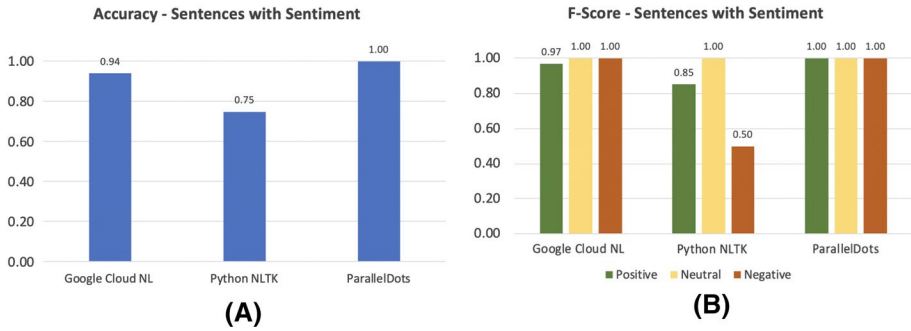
**Fig. 5** Accuracy and F-Score of the three tools applied to all the reviews of the Hu-Liu dataset

**Table 6** Confusion matrices for each tool applied only to the sentences with sentiment of the KnowMIS dataset

Sentences with Sentiment (KnowMIS dataset)									
	Google NL API			Python NLTK			ParallelDots		
	Pos	Neu	Neg	Pos	Neu	Neg	Pos	Neu	Neg
Expert									
Pos	14	1	0	11	2	2	15	0	0
Neu	0	0	0	0	0	0	0	0	0
Neg	0	0	1	0	0	1	0	0	1

Lastly, Table 8 contains the confusion matrices related to the sentences with statements, with Fig. 8a, b showing accuracy and F-scores. What emerges is that all the tools do not perform very well on this set of sentences, with only Google NL API having an accuracy greater than 50%. This shows that the analysis of sentences containing a statement (which usually are factual sentences but that may contain an orientation) is often more complex to be analyzed with respect to those sentences that contain sentiments or reactions. But what is more interesting for our analysis is that all the three tools find that some kind of orientation (positive or negative) could be present in those sentences which usually are just considered as objective claims. Indeed, the analysis of such statements may hide some hidden judgment by the opinion holders that would be worth investigating.

The results obtained by this qualitative evaluation can be summarized as follows. Typically, the sentiment analysis at a review-level is more efficient than analysis at sentence-level. The class which seems to be more problematic for the three tools is the neutral one: it is more difficult to understand if an opinion of the user is neutral rather than positive

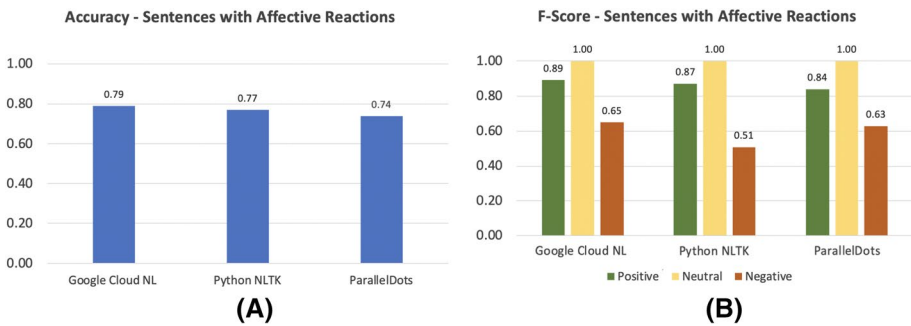


**Fig. 6** Accuracy and F-Score of the three tools applied to all the sentences with sentiment of the KnowMIS dataset

**Table 7** Confusion matrices for each tool applied to the sentences with affective reactions of the KnowMIS dataset

Sentences with Affective Reactions (KnowMIS dataset)

	Google NL API			Python NLTK			ParallelDots		
	Pos	Neu	Neg	Pos	Neu	Neg	Pos	Neu	Neg
Expert									
Pos	136	24	2	131	10	21	122	25	15
Neu	0	0	0	0	0	0	0	0	0
Neg	6	9	16	8	5	18	5	5	21

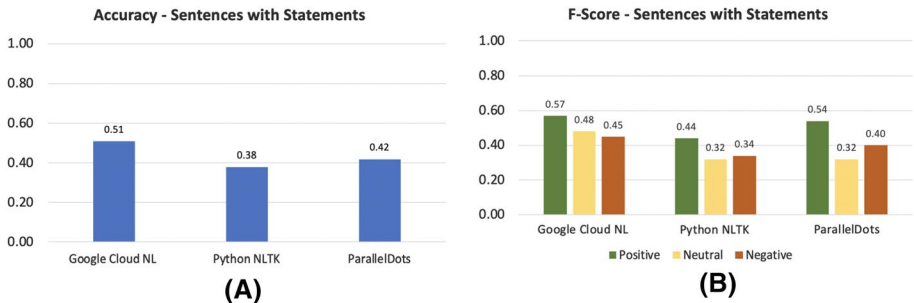


**Fig. 7** Accuracy and F-Score of the three tools applied to all the sentences with affective reactions of the KnowMIS dataset

or negative. This can also be explained by considering that in the style of a review, even a sentence that could seem neutral indeed contains at least an orientation of the opinion holder towards the product. And this is one of the main characteristics of the KnowMIS-ABSA model: a statement can have an orientation even when it does not contain a sentiment or an affective reaction. Indeed, if we observe the results obtained only for the sentences with statement, that has been annotated in the original dataset by Hu and Liu as “claims” that do not contain subjectivity, only the 39% of such sentence (75 out of 194) have been effectively classified as neutral. In the remaining 61%, the tools have identified

**Table 8** Confusion matrices for each tool applied to the sentences with statements of the KnowMIS dataset

Sentences with Statements (KnowMIS dataset)									
	Google NL API			Python NLTK			ParallelDots		
	Pos	Neu	Neg	Pos	Neu	Neg	Pos	Neu	Neg
Expert									
Pos	42	24	3	36	15	18	39	25	5
Neu	21	35	11	27	17	23	27	23	17
Neg	16	21	21	30	8	20	9	29	20



**Fig. 8** Accuracy and F-Score of the three tools applied to all the sentences with statements of the KnowMIS dataset

that an orientation (positive or negative) is present in these sentences, although there is not an affective or emotional content. Therefore, differently from what the classical approaches like the one proposed by Liu (2012) foresee, according to which a statement should be necessarily neutral, the obtained results seem to confirm that the statements may contain a positive or negative orientation. Eventually, an analysis according to the KnowMIS-ABSA model allows the identification of other kinds of subjectivity, besides the traditional concept of sentiment, which can be useful, for instance, in the context of marketing analysis.

Another relevant result is that performing a separate analysis of the sentences containing sentiments from the ones containing affective reactions and from the statements seems to slightly improve the results. Therefore, this suggests that the research should focus on improving performances regarding statements and factual sentences to improve the results of ABSA solutions, using different tools and techniques for each of the three categories.

### 5 Conclusion

The analysis of the literature regarding the latest advancements in the field of SA and ABSA made evident the terrific progress that has been made in the automatic understanding and evaluation of human opinions. Such an analysis has also highlighted the great confusion lying in the field regarding the basic concepts that underline the ABSA, like sentiment, opinion, affect, emotion. Moved by these considerations, in this work, we proposed an overview and a new reference model for ABSA solutions. The model has been used in a case study related to the analysis of product reviews, comparing it with a traditional

approach to ABSA. This qualitative evaluation highlighted that the differences between the different dimensions (sentiment, affective reaction, statement) and metrics (polarity or orientation) defined in the model allow us to better analyze a set of reviews, even when using the same computational techniques and software tools.

From the marketing perspective, the KnowMIS-ABSA model could be effective in the interpretation of the results of any ABSA technique used to analyze end-users reviews. In particular, the capability to distinguish between sentiments and affective reactions could drive the identification of marketing strategies which could be more effective to target a specific group of users. Furthermore, the analysis of statements, that could contain an orientation also when there is no emotion or sentiment in them, can represent another source of important information for the marketing.

In future works, thanks to a clearer definition of the basic concepts of ABSA proposed by the KnowMIS-ABSA model, new approaches and techniques can be defined upon it. This work, indeed, gave rise to further reflections on additional elements of ABSA solutions, in particular on the new characteristics of reviews in modern websites. Therefore, in future work, we will extend the KnowMIS-ABSA model to strengthen and extend it with considerations on reviews and other elements of ABSA. One of the elements to consider is the analysis of comments and opinions related to the reviews published by other users. This requires the evaluation of the importance of the review itself, as well as of the comments, and to understand if the comment has a positive, neutral, or negative opinion with respect to the original review. An approach similar to the one proposed in Aljuaid et al. (2021), wherein a technique to evaluate the importance and polarity of citations in scientific documents is proposed, will be investigated. Lastly, having demonstrated in the current work that implementation of the model is feasible and provides some advantages, we will develop a completely automatic solution for ABSA based on the KnowMIS-ABSA model, by selecting and implementing adequate computational techniques.

**Acknowledgements** The authors would like to thank Giovanni Caiazzo, Antonio Citro, Stefania De Vivo, Filippo Genoese, Alessandro Leo, Francesca Marzano for their collaboration and their contribution to bibliographical researches, to the experiments and data acquisition.

**Author contributions** The contributions of the authors are as follows: conceptualization: G.D., M.G., I.L.R.; model definition: G.D., M.G., I.L.R.; methodology: G.D., M.G.; validation: G.D., M.G.; investigation: G.D., M.G.; writing-original draft: G.D., M.G.; writing- review and editing: G.D., M.G.. All authors have read and agreed to the published version of the manuscript.

**Availability of data and material** The dataset used in the case study is publicly available (Hu and Liu 2004); the additional data used in the experiments containing the classifications performed by the experts are all included in the article. The tools used for the experiments are freely available online.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akhtar MS, Garg T, Ekbal A (2020) Multi-task learning for aspect term extraction and aspect sentiment classification. *Neurocomputing* 398:247–256. <https://doi.org/10.1016/j.neucom.2020.02.093>
- 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. <https://doi.org/10.1016/j.jocs.2017.11.006>
- Alexopoulos P, Wallace M (2015) Creating domain-specific semantic lexicons for aspect-based sentiment analysis. In: 2015 10th international workshop on semantic and social media adaptation and personalization (SMAP), pp 1–6. <https://doi.org/10.1109/SMAP.2015.7370083>
- Aljuaid H, Iftikhar R, Ahmad S, Asif M, Tanvir Afzal M (2021) Important citation identification using sentiment analysis of in-text citations. *Telemat Inform.* <https://doi.org/10.1016/j.tele.2020.101492>
- Aly ES, van der Haar DT (2020) Slang-based text sentiment analysis in Instagram. In: Yang XS, Sherratt S, Dey N, Joshi A (eds) Fourth international congress on information and communication technology. Springer, Singapore, pp 321–329
- Anand D, Naorem D (2016) Semi-supervised aspect based sentiment analysis for movies using review filtering. *Procedia Comput Sci* 84:86–93. <https://doi.org/10.1016/j.procs.2016.04.070> proceeding of the Seventh International Conference on Intelligent Human Computer Interaction (IHCI 2015)
- Angelidis S, Lapata M (2018) Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised. In: Proceedings of the 2018 conference on empirical methods in natural language processing, Association for Computational Linguistics, Brussels, Belgium, pp 3675–3686. <https://doi.org/10.18653/v1/D18-1403>
- Appel O, Chiclana F, Carter J, Fujita H (2016) A hybrid approach to the sentiment analysis problem at the sentence level. *Knowl Based Syst* 108:110–124. <https://doi.org/10.1016/j.knosys.2016.05.040>
- Asif M, Ishtiaq A, Ahmad H, Aljuaid H, Shah J (2020) Sentiment analysis of extremism in social media from textual information. *Telemat Inform.* <https://doi.org/10.1016/j.tele.2020.101345>
- Bafna K, Toshniwal D (2013) Feature based summarization of customers' reviews of online products. *Procedia Comput Sci* 22:142–151. <https://doi.org/10.1016/j.procs.2013.09.090>, 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems—KES2013
- Bancken W, Alfarone D, Davis J (2014) Automatically detecting and rating product aspects from textual customer reviews. In: Proceedings of the 1st international conference on interactions between data mining and natural language processing—volume 1202, CEUR-WS.org, Aachen, DEU, DMNLP'14, pp 1–16
- Bao L, Lambert P, Badia T (2019) Attention and lexicon regularized LSTM for aspect-based sentiment analysis. In: Proceedings of the 57th annual meeting of the association for computational linguistics: student research workshop, Association for Computational Linguistics, Florence, Italy, pp 253–259. <https://doi.org/10.18653/v1/P19-2035>
- Basiri ME, Abdar M, Kabiri A, Nemati S, Zhou X, Allahbakhshi F, Yen NY (2020) Improving sentiment polarity detection through target identification. *IEEE Trans Comput Soc Syst* 7(1):113–128. <https://doi.org/10.1109/TCSS.2019.2951326>
- Behdenna S, Barigou F, Belalem G (2018) Document level sentiment analysis: a survey. *EAI Endorsed Trans Context Aware Syst Appl.* <https://doi.org/10.4108/eai.14-3-2018.154339>
- Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends. *Knowl Based Syst.* <https://doi.org/10.1016/j.knosys.2021.107134>
- Cahyadi A, Khodra ML (2018) Aspect-based sentiment analysis using convolutional neural network and bidirectional long short-term memory. In: 2018 5th international conference on advanced informatics: concept theory and applications (ICAICTA), pp 124–129. <https://doi.org/10.1109/ICAICTA.2018.8541300>
- Cambria E (2013) An introduction to concept-level sentiment analysis. In: Castro F, Gelbukh A, González M (eds) *Advances in soft computing and its applications*. Springer, Heidelberg, pp 478–483
- Cambria E, Schuller B, Liu B, Wang H, Havasi C (2013) Knowledge-based approaches to concept-level sentiment analysis. *IEEE Intell Syst* 28(2):12–14. <https://doi.org/10.1109/MIS.2013.45>
- 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, association for computing machinery, New York, NY, USA, pp 105–114
- Casillo M, Clarizia F, Colace F, Lombardi M, Pascale F, Santaniello D (2019) An approach for recommending contextualized services in e-tourism. *Information (Switzerland)*. <https://doi.org/10.3390/INFO10050180>

- Casillo M, Clarizia F, D'Aniello G, De Santo M, Lombardi M, Santaniello D (2020) Chat-bot: a cultural heritage aware teller-bot for supporting touristic experiences. *Pattern Recogn Lett* 131:234–243. <https://doi.org/10.1016/j.patrec.2020.01.003>
- Carrillo-de Albornoz J, Aker A, Kurtic E, Plaza L (2019) Beyond opinion classification: extracting facts, opinions and experiences from health forums. *PLoS ONE* 14(1):1–20. <https://doi.org/10.1371/journal.pone.0209961>
- Chang M, D'Aniello G, Gaeta M, Orciuoli F, Sampson D, Simonelli C (2020) Building ontology-driven tutoring models for intelligent tutoring systems using data mining. *IEEE Access* 8:48151–48162. <https://doi.org/10.1109/ACCESS.2020.2979281>
- Colace F, de Santo M, Greco L (2014) Safe: a sentiment analysis framework for e-learning. *Int J Emerg Technol Learn* 9(6):37–41. <https://doi.org/10.3991/ijet.v9i6.4110>
- Colace F, Loia V, Tomasiello S (2019) Revising recurrent neural networks from a granular perspective. *Appl Soft Comput* 82:105535. <https://doi.org/10.1016/j.asoc.2019.105535>
- Colhon M, Bădică C, Şendre A (2014) Relating the opinion holder and the review accuracy in sentiment analysis of tourist reviews. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol 8793, pp 246–257. [https://doi.org/10.1007/978-3-319-12096-6\\_22](https://doi.org/10.1007/978-3-319-12096-6_22). cited By 14
- D'Aniello G, Gaeta M, Loia F, Reformat M, Toti D (2018) An environment for collective perception based on fuzzy and semantic approaches. *J Artif Intell Soft Comput Res* 8(3):191–210. <https://doi.org/10.1515/jaiscr-2018-0013>
- D'Aniello G, Gaeta M, Orciuoli F, Sansonetti G, Sorgente F (2020) Knowledge-based smart city service system. *Electronics (Switzerland)* 9(6):1–22. <https://doi.org/10.3390/electronics9060965>
- D'Aniello G, Gaeta M, Lepore M, Perone M (2021) Knowledge-driven fuzzy consensus model for team formation. *Expert Syst Appl* 184:115522. <https://doi.org/10.1016/j.eswa.2021.115522>
- Ding J, Sun H, Wang X, Liu X (2018) Entity-level sentiment analysis of issue comments. In: *Proceedings of the 3rd international workshop on emotion awareness in software engineering*, Association for Computing Machinery, New York, NY, USA, SEmotion '18, pp 7–13. <https://doi.org/10.1145/3194932.3194935>
- Do HH, Prasad P, Maag A, Alsadoon A (2019) Deep learning for aspect-based sentiment analysis: a comparative review. *Expert Syst Appl* 118:272–299. <https://doi.org/10.1016/j.eswa.2018.10.003>
- Dong Y, Zha Q, Zhang H, Kou G, Fujita H, Chiclana F, Herrera-Viedma E (2018) Consensus reaching in social network group decision making: research paradigms and challenges. *Knowl Based Syst* 162:3–13. <https://doi.org/10.1016/j.knosys.2018.06.036>
- Esuli A, Moreo A, Sebastiani F, Cambria E (2020) Cross-lingual sentiment quantification. *IEEE Intell Syst* 35(3):106–114. <https://doi.org/10.1109/MIS.2020.2979203>
- Fang X, Zhan J (2015) Sentiment analysis using product review data. *J Big Data* 2(1):5. <https://doi.org/10.1186/s40537-015-0015-2>
- Farhadloo M, Rolland E (2016) *Fundamentals of sentiment analysis and its applications*. Springer, Cham, pp 1–24
- Fitri VA, Andreswari R, Hasibuan MA (2019) Sentiment analysis of social media twitter with case of ANTI-LGBT campaign in Indonesia using naïve Bayes, decision tree, and random forest algorithm. *Procedia Comput Sci* 161:765–772. <https://doi.org/10.1016/j.procs.2019.11.181>, the Fifth Information Systems International Conference, 23–24 July 2019, Surabaya, Indonesia
- Gan C, Wang L, Zhang Z (2020) Multi-entity sentiment analysis using self-attention based hierarchical dilated convolutional neural network. *Futur Gener Comput Syst* 112:116–125. <https://doi.org/10.1016/j.future.2020.05.022>
- Ganganwar V, Rajalakshmi R (2019) Implicit aspect extraction for sentiment analysis: a survey of recent approaches. *Procedia Comput Sci* 165:485–491. <https://doi.org/10.1016/j.procs.2020.01.010>, 2nd International Conference on Recent Trends in Advanced Computing ICRATAC - DISRUP - TIV INNOVATION, 2019 November 11–12, 2019
- Gurini D, Gasparetti F, Micarelli A, Sansonetti G (2013) A sentiment-based approach to twitter user recommendation. In: *CEUR Workshop proceedings, CEUR-WS*, vol 1066, conference of 5th ACM RecSys workshop on recommender systems and the social web, RSWeb 2013- Co-located with the 7th ACM conference on recommender systems, RecSys 2013; Conference Date: 13 October 2013; Conference Code:111232
- Haider S, Tanvir Afzal M, Asif M, Maurer H, Ahmad A, Abuarqoub A (2021) Impact analysis of adverbs for sentiment classification on twitter product reviews. *Concurr Comput Pract Exp*. <https://doi.org/10.1002/cpe.4956>
- Hemmatian F, Sohrabi MK (2019) A survey on classification techniques for opinion mining and sentiment analysis. *Artif Intell Rev* 52(3):1495–1545. <https://doi.org/10.1007/s10462-017-9599-6>

- Herrera F, Herrera-Viedma E, Verdegay J (1996) A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets Syst* 78(1):73–87. [https://doi.org/10.1016/0165-0114\(95\)00107-7](https://doi.org/10.1016/0165-0114(95)00107-7)
- Herrera-Viedma E, Cabrerizo FJ, Kacprzyk J, Pedrycz W (2014) A review of soft consensus models in a fuzzy environment. *Inf Fusion* 17:4–13. <https://doi.org/10.1016/j.inffus.2013.04.002>
- Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIG-KDD international conference on knowledge discovery and data mining, Association for Computing Machinery, New York, NY, USA, KDD '04, pp 168–177. <https://doi.org/10.1145/1014052.1014073>
- Jiaxiang H (2020) Building domain specific lexicon based on TikTok comment dataset. *ArXiv abs/2012.08773*
- Karthika P, Murugeswari R, Manoranjithem R (2019) Sentiment analysis of social media network using random forest algorithm. In: 2019 IEEE international conference on intelligent techniques in control, optimization and signal processing (INCOS), pp 1–5. <https://doi.org/10.1109/INCOS45849.2019.8951367>
- Khan K, Baharudin B, Khan A, Ullah A (2014) Mining opinion components from unstructured reviews: a review. *J King Saud Univ Comput Inf Sci* 26(3):258–275. <https://doi.org/10.1016/j.jksuci.2014.03.009>
- Kim SM, Hovy E (2004) Determining the sentiment of opinions. In: Proceedings of the 20th international conference on computational linguistics, Association for Computational Linguistics, USA, COLING '04, p 1367-es. <https://doi.org/10.3115/1220355.1220555>
- Klyuev V, Oleshchuk V (2011) Semantic retrieval: an approach to representing, searching and summarising text documents. *Int J Inf Technol Commun Converg* 1(2):221–234. <https://doi.org/10.1504/IJITCC.2011.039287>
- Kolya A, Das D, Ekbal A, Bandyopadhyay S (2012) Roles of event actors and sentiment holders in identifying event-sentiment association. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7181 LNCS(PART 1), pp 513–525. [https://doi.org/10.1007/978-3-642-28604-9\\_421](https://doi.org/10.1007/978-3-642-28604-9_421)
- Li X, Bing L, Li P, Lam W (2019) A unified model for opinion target extraction and target sentiment prediction. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, no. 01, pp 6714–6721. <https://doi.org/10.1609/aaai.v33i01.33016714>
- 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. Springer, Boston, pp 415–463. [https://doi.org/10.1007/978-1-4614-3223-4\\_13](https://doi.org/10.1007/978-1-4614-3223-4_13)
- Liu P, Joty S, Meng H (2015) Fine-grained opinion mining with recurrent neural networks and word embeddings. In: Proceedings of the 2015 conference on empirical methods in natural language processing, Association for Computational Linguistics, Lisbon, Portugal, pp 1433–1443. <https://doi.org/10.18653/v1/D15-1168>
- Loia V, Senatore S (2014) A fuzzy-oriented sentic analysis to capture the human emotion in web-based content. *Knowl Based Syst* 58:75–85. <https://doi.org/10.1016/j.knosys.2013.09.024>
- Luo W, Zhuang F, Zhao W, He Q, Shi Z (2015) QPLSA: utilizing quad-tuples for aspect identification and rating. *Inf Process Manag* 51(1):25–41. <https://doi.org/10.1016/j.ipm.2014.08.004>
- Maglogiannis I, Iliadis L, Pimenidis E, Alexandridis G, Michalakakis K, Aliprantis J, Polydoros P, Tsantilas P, Caridakis G (2020) A deep learning approach to aspect-based sentiment prediction. In: Artificial intelligence applications and innovations 16th IFIP WG 125 international conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020. Proceedings, Part I 583:397–408
- Manek AS, Shenoy PD, Mohan MC, Venugopal K (2017) Aspect term extraction for sentiment analysis in large movie reviews using GINI index feature selection method and SVM classifier. *World Wide Web* 20(2):135–154
- Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J* 5(4):1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Meškelė D, Frasinca F (2020) Aldonar: a hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model. *Inf Process Manag* 57(3):102211. <https://doi.org/10.1016/j.ipm.2020.102211>
- Min HJ, Park JC (2012) Identifying helpful reviews based on customer's mentions about experiences. *Expert Syst Appl* 39(15):11830–11838. <https://doi.org/10.1016/j.eswa.2012.01.116>
- Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlu M, Gao J (2021) Deep learning—based text classification: a comprehensive review. *ACM Comput Surv*. <https://doi.org/10.1145/3439726>
- Mowlaei ME, Saniee Abadeh M, Keshavarz H (2020) Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Syst Appl* 148:113234. <https://doi.org/10.1016/j.eswa.2020.113234>

- Mubarok MS, Adiwijaya, Aldhi MD (2017) Aspect-based sentiment analysis to review products using naïve Bayes. In: AIP conference proceedings, vol 1867, no. 1, pp 020060. <https://doi.org/10.1063/1.4994463>
- Mumtaz D, Ahuja B (2018) A lexical and machine learning-based hybrid system for sentiment analysis. Springer, Singapore, pp 165–175. [https://doi.org/10.1007/978-981-10-4555-4\\_11](https://doi.org/10.1007/978-981-10-4555-4_11)
- Munzero M, Montero C, Sutinen E, Pajunen J (2014) Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Trans Affect Comput* 5(02):101–111. <https://doi.org/10.1109/TAFFC.2014.2317187>
- Nakov P, Ritter A, Rosenthal S, Sebastiani F, Stoyanov V (2016) SemEval-2016 task 4: Sentiment analysis in Twitter. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), Association for Computational Linguistics, San Diego, California, pp 1–18, 10.18653/v1/S16-1001
- Nazir A, Rao Y, Wu L, Sun L (2020) Issues and challenges of aspect-based sentiment analysis: a comprehensive survey. *IEEE Trans Affect Comput*. <https://doi.org/10.1109/TAFFC.2020.2970399>
- Neviarouskaya A, Prendinger H, Ishizuka M (2007) Textual affect sensing for sociable and expressive online communication. In: Paiva ACR, Prada R, Picard RW (eds) *Affective computing and intelligent interaction*. Springer, Berlin, pp 218–229
- Nguyen TH, Shirai K (2015) PhraseRNN: Phrase recursive neural network for aspect-based sentiment analysis. In: Proceedings of the 2015 conference on empirical methods in natural language processing, Association for Computational Linguistics, Lisbon, Portugal, pp 2509–2514. <https://doi.org/10.18653/v1/D15-1298>
- Novielli N, Calefato F, Lanubile F (2020) Love, joy, anger, sadness, fear, and surprise: Se needs special kinds of ai: a case study on text mining and se. *IEEE Softw* 37(3):86–91. <https://doi.org/10.1109/MS.2020.2968557>
- Pannala NU, Nawarathna CP, Jayakody JTK, Rupasinghe L, Krishnadeva K (2016) Supervised learning based approach to aspect based sentiment analysis. In: 2016 IEEE international conference on computer and information technology (CIT), pp 662–666. <https://doi.org/10.1109/CIT.2016.107>
- Pekar V, Afzal N, Bohnet B (2014) Ubham: lexical resources and dependency parsing for aspect-based sentiment analysis. In: *SemEval@COLING*
- Pontiki M, Galanis D, Papageorgiou H, Androutsopoulos I, Manandhar S, Al-Smadi M, Al-Ayyoub M, Zhao Y, Qin B, De Clercq O, Hoste V, Apidianaki M, Tannier X, Loukachevitch N, Kotelnikov E, Bel N, Jiménez-Zafra SM, Eryiğit G (2016) SemEval-2016 task 5: aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), Association for Computational Linguistics, San Diego, CA, pp 19–30. <https://doi.org/10.18653/v1/S16-1002>
- Poria S, Cambria E, Ku LW, Gui C, Gelbukh A (2014) A rule-based approach to aspect extraction from product reviews. In: *SocialNLP@COLING*
- Preethi P, Uma V, kumar A (2015) Temporal sentiment analysis and causal rules extraction from tweets for event prediction. *Procedia Computer Science* 48:84–89. <https://doi.org/10.1016/j.procs.2015.04.154>, international Conference on Computer, Communication and Convergence (ICCC 2015)
- Rana TA, Cheah YN (2016) Aspect extraction in sentiment analysis: comparative analysis and survey. *Artif Intell Rev* 46(4):459–483. <https://doi.org/10.1007/s10462-016-9472-z>
- Rana TA, Cheah YN (2018) Improving aspect extraction using aspect frequency and semantic similarity-based approach for aspect-based sentiment analysis. In: Meesad P, Sodsee S, Unger H (eds) *Recent advances in information and communication technology 2017*. Springer, Cham, pp 317–326
- Rao G, Huang W, Feng Z, Cong Q (2018) LSTM with sentence representations for document-level sentiment classification. *Neurocomputing* 308:49–57. <https://doi.org/10.1016/j.neucom.2018.04.045>
- Rosenthal S, Farra N, Nakov P (2017) SemEval-2017 task 4: Sentiment analysis in Twitter. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017), Association for Computational Linguistics, Vancouver, Canada, pp 502–518. <https://doi.org/10.18653/v1/S17-2088>
- Saad S, Saberi B (2017) Sentiment analysis or opinion mining: a review. *Int J Adv Sci Eng Inf Technol* 7(5):1660–1666. <https://doi.org/10.18517/ijaseit.7.5.2137>
- Schouten K, Frasinca F (2016) Survey on aspect-level sentiment analysis. *IEEE Trans Knowl Data Eng* 28(3):813–830. <https://doi.org/10.1109/TKDE.2015.2485209>
- Seki Y, Kando N, Aono M (2009) Multilingual opinion holder identification using author and authority viewpoints. *Inf Process Manag* 45(2):189–199. <https://doi.org/10.1016/j.ipm.2008.11.004>
- Shah P, Swaminarayan P (2021) Lexicon-based sentiment analysis on movie review in the Gujarati language. *Int J Inf Technol Commun Converg* 4(1):63–72. <https://doi.org/10.1504/IJITCC.2021.119111>
- Shirsat VS, Jagdale RS, Deshmukh SN (2017) Document level sentiment analysis from news articles. In: 2017 international conference on computing, communication, control and automation (ICCUBE), pp 1–4. <https://doi.org/10.1109/ICCUBE.A.2017.8463638>

- Sima I, Vunvulea M (2013) A rule-based, domain independent approach for opinion and holder identification. In: Proceedings—2013 IEEE 9th international conference on intelligent computer communication and processing, ICCP 2013, pp 55–62. <https://doi.org/10.1109/ICCP.2013.6646081>
- Singh VK, Piriyani R, Uddin A, Wailla P (2013) Sentiment analysis of movie reviews: a new feature-based heuristic for aspect-level sentiment classification. In: 2013 international multi-conference on automation, computing, communication, control and compressed Sensing (iMac4s), pp 712–717. <https://doi.org/10.1109/iMac4s.2013.6526500>
- Sokolova M, Bobicev V (2011) Sentiments and opinions in health-related web messages. In: Proceedings of the international conference recent advances in natural language processing, 2011, pp 132–139
- Tomasiello S (2020) Least-squares fuzzy transforms and autoencoders: some remarks and application. IEEE Trans Fuzzy Syst. <https://doi.org/10.1109/TFUZZ.2020.3007442>
- Tripathy A, Agrawal A, Rath SK (2016) Classification of sentiment reviews using n-gram machine learning approach. Expert Syst Appl 57:117–126. <https://doi.org/10.1016/j.eswa.2016.03.028>
- Tu W, Cheung D, Mamoulis N (2015) Time-sensitive opinion mining for prediction. In: Proceedings of the AAAI conference on artificial intelligence, vol 29
- 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. <https://doi.org/10.1016/j.ipm.2018.03.008>
- Tubishat M, Idris N, Abushariah M (2021) Explicit aspects extraction in sentiment analysis using optimal rules combination. Futur Gener Comput Syst 114:448–480. <https://doi.org/10.1016/j.future.2020.08.019>
- Vieira JPA, Moura RS (2017) An analysis of convolutional neural networks for sentence classification. In: 2017 XLIII Latin American Computer Conference (CLEI), pp 1–5. <https://doi.org/10.1109/CLEI.2017.8226381>
- Wang G, Zhang Z, Sun J, Yang S, Larson CA (2015) POS-RS: a random subspace method for sentiment classification based on part-of-speech analysis. Info Process Manag 51(4):458–479. <https://doi.org/10.1016/j.ipm.2014.09.004>
- Wang H, Castanon JA (2015) Sentiment expression via emoticons on social media. In: 2015 IEEE international conference on big data (Big Data), pp 2404–2408. <https://doi.org/10.1109/BigData.2015.7364034>
- Wang Y, Huang M, Zhu X, Zhao L (2016) Attention-based LSTM for aspect-level sentiment classification. In: Proceedings of the 2016 conference on empirical methods in natural language processing. Association for Computational Linguistics, Austin, Texas, pp 606–615. <https://doi.org/10.18653/v1/D16-1058>
- Wiegand M, Klakow D (2010) Convolution kernels for opinion holder extraction. In: NAACL HLT 2010—human language technologies: the 2010 annual conference of the North American Chapter of the Association for Computational Linguistics, proceedings of the main conference, pp 795–803
- Wiegand M, Bocionek C, Ruppenhofer J (2016) Opinion holder and target extraction on opinion compounds—a linguistic approach. In: 2016 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies, NAACL HLT 2016—proceedings of the conference, pp 800–810. <https://doi.org/10.18653/v1/n16-1094>
- Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on human language technology and empirical methods in natural language processing, Association for Computational Linguistics, USA, HLT '05, pp 347–354. <https://doi.org/10.3115/1220575.1220619>
- Xing B, Liao L, Song D, Wang J, Zhang F, Wang Z, Huang H (2019) Earlier attention? Aspect-aware LSTM for aspect-based sentiment analysis. In: Proceedings of the twenty-eighth international joint conference on artificial intelligence, pp 5313–5319
- Xu L, Liu K, Lai S, Chen Y, Zhao J (2013) Walk and learn: A two-stage approach for opinion words and opinion targets co-extraction. In: Proceedings of the 22nd international conference on world wide web, Association for Computing Machinery, New York, NY, USA, WWW '13 Companion, pp 95–96. <https://doi.org/10.1145/2487788.2487831>
- Xu L, Bing L, Lu W, Huang F (2020) Aspect sentiment classification with aspect-specific opinion spans. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), Association for Computational Linguistics, Online, pp 3561–3567. <https://doi.org/10.18653/v1/2020.emnlp-main.288>
- Xu R, Gui L, Xu J, Lu Q, Wong KF (2015) Cross lingual opinion holder extraction based on multi-kernel SVMs and transfer learning. World Wide Web 18(2):299–316. <https://doi.org/10.1007/s11280-013-0246-0>

- Yadav A, Vishwakarma DK (2020) Sentiment analysis using deep learning architectures: a review. *Artif Intell Rev* 53(6):4335–4385
- Yadollahi A, Shahraki AG, Zaiane OR (2017) Current state of text sentiment analysis from opinion to emotion mining. <https://doi.org/10.1145/3057270>
- Yan Z, Xing M, Zhang D, Ma B (2015) EXPRS: an extended pagerank method for product feature extraction from online consumer reviews. *Inf Manag* 52(7):850–858. <https://doi.org/10.1016/j.im.2015.02.002>, novel applications of social media analytics
- Yu J, Zha ZJ, Wang M, Chua TS (2011) Aspect ranking: Identifying important product aspects from online consumer reviews. In: Proceedings of the 49th annual meeting of the association for computational linguistics: Human Language Technologies, Association for Computational Linguistics, Portland, Oregon, USA, pp 1496–1505
- Yu J, Jiang J, Xia R (2019) Global inference for aspect and opinion terms co-extraction based on multi-task neural networks. *IEEE/ACM Trans Audio Speech Lang Process* 27(1):168–177. <https://doi.org/10.1109/TASLP.2018.2875170>
- 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. <https://doi.org/10.1007/s10115-018-1236-4>
- Zeng D, Dai Y, Li F, Wang J, Sangaiah AK (2019) Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism. *J Int Fuzzy Syst* 36(5):3971–3980. <https://doi.org/10.3233/JIFS-169958>
- Zhang L, Liu B (2014) Aspect and entity extraction for opinion mining. Springer, Berlin, pp 1–40. [https://doi.org/10.1007/978-3-642-40837-3\\_1](https://doi.org/10.1007/978-3-642-40837-3_1)
- Zhu J, Wang H, Tsou BK, Zhu M (2009) Multi-aspect opinion polling from textual reviews. In: Proceedings of the 18th ACM conference on information and knowledge management, Association for Computing Machinery, New York, NY, USA, CIKM '09, pp 1799–1802. <https://doi.org/10.1145/1645953.1646233>

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