



Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review

Komang Wahyu Trisna & Huang Jin Jie

To cite this article: Komang Wahyu Trisna & Huang Jin Jie (2022) Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review, Applied Artificial Intelligence, 36:1, 2014186, DOI: [10.1080/08839514.2021.2014186](https://doi.org/10.1080/08839514.2021.2014186)

To link to this article: <https://doi.org/10.1080/08839514.2021.2014186>



© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 07 Jan 2022.



Submit your article to this journal [↗](#)



Article views: 6212



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 6 View citing articles [↗](#)

Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review

Komang Wahyu Trisna^{a,b} and Huang Jin Jie^{a,c}

^aSchool of Computer Science and Technology, Harbin University of Science and Technology, Harbin, China; ^bDepartment of Informatics, STMIK Primakara, Bali, Indonesia; ^cKey Laboratory of Advanced Manufacturing and Intelligent Technology, Harbin University of Science and Technology, Harbin, China

ABSTRACT

The emergence of various e-commerce sites has led to an increase in review sites for various services and products. People nowadays easily get information about products and services that will be used through reviews. Here sentiment analysis plays an important role in classifying the polarity of product reviews. However, with a large number of reviews, a sentiment analysis that only gives overall polarity is not sufficient. This will make it difficult to find the reviews of certain aspects (features) of the product. Aspect-based sentiment analysis as fine-grained sentiment analysis is able to provide specific polarity for each aspect contained in a sentence. Various kinds of development methods have been carried out to provide accurate results in aspect-based sentiment analysis. This paper will discuss the various deep learning methods that have been carried out and provide the possibility of research that can be carried out from Aspect-Based Sentiment Analysis.

ARTICLE HISTORY

Received 17 December 2020
Revised 26 November 2021
Accepted 30 November 2021

Introduction

With the use of internet tools rapidly increasing, the use of e-commerce such as Amazon, Walmart, and Alibaba have recently increased as well. The growth of e-commerce, as a new shopping and marketing channel, is causing a surge in review sites for a variety of services and products (Pontiki, Galanis, and Papageorgiou 2015). Customers buy products online and freely express their ideas and thoughts. Due to this increase, huge amounts of data are generated. In business sectors, this huge amount of review data has great efforts to find out customers' sentiments and opinions, often expressed in free text, toward companies' products and services. Customer reviews are in unstructured textual form, which makes it difficult to be summarized by a computer. In addition, manual analysis of this huge amount of data is nearly impossible. Here automatic sentiment analysis comes to solve these problems. [Figure 1](#) [Figure 2](#) [Table 1](#)

CONTACT Komang Wahyu Trisna  km.ayutrisna@gmail.com  School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, China 150080

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

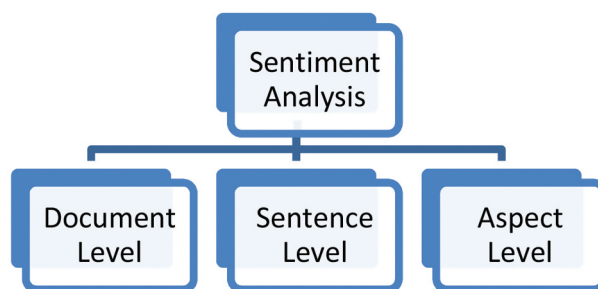


Figure 1. Level of sentiment analysis.

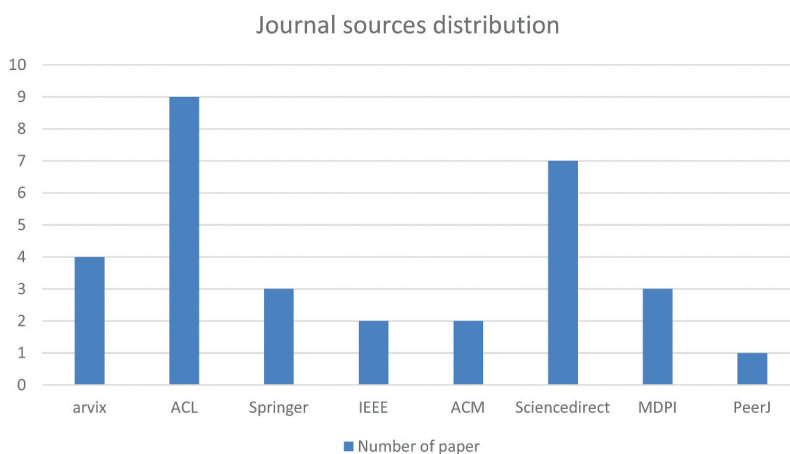


Figure 2. Journal sources distribution.

Sentiment analysis also known as opinion mining (Pang and Lee 2008), is classically classified based on the overall polarity of the input. The majority of current approaches, however, attempt to detect the overall polarity of a sentence, paragraph, or text spam and are based on the assumption that the sentiment expressed in a sentence is unified and consistent, which does not hold in reality. Sentiment analysis aims to classify text as positive or negative, sometimes even neutral according to the affective states and the subjective information of the text. In this case, sentiment analysis is called sentiment classification because it divides polarities into two or more classes. In theory, sentiment analysis is divided into three levels, namely document level, sentence level, and aspect/feature level, sometimes also called target. At the document level, sentiment classification is determined by looking at opinions from a review and determining their polarity. The system will decide whether the overall review contains positive or negative reviews (Toqir and Cheah 2016). Sentence level, the analysis is done one by one to the sentence. The last level of

sentiment analysis is aspect level sentiment analysis. Aspect-level sentiment analysis is proposed to conduct fine-grained opinion mining toward specific entities or categories of entities which are also called targets (Ren et al. 2020).

The goal of aspect-based sentiment analysis is to identify the sentiment polarity of a specific opinion target/aspect expressed in a comment or review by a reviewer. Conventional approaches mainly focus on designing a set of features such as bag-of-words, sentiment lexicon to train classification (e.g., SVM) for aspect-based sentiment analysis (Jiang et al. 2011). To make the system accurate, the aforementioned strategy requires a lot of features, and later machine learning techniques are proposed (Manek and Shenoy 2016; Syahrul and Dwi 2017). However, with the popularity of deep learning in natural language processing (NLP), researchers used that approach to solve aspect-based sentiment analysis problems (Poria, Cambria, and Gelbukh 2016; Tang, Qin, and Liu 2014).

This survey will discuss the methods used in solving aspect-based sentiment analysis problems. Through this survey, it will be possible to see the pattern of methods used in solving those problems.

Method for Collect and Review Papers

In conducting a review on aspect-based sentiment analysis, the first step is to find related journals that we will review. The search process for the journals that we will review is to search for them through Google Scholar by using several related keywords such as ABSA, aspect extraction, and deep learning in aspect extraction. From Google Scholar, we choose which papers are related to deep learning on aspect-based sentiment analysis, and then we also look at the reputation of the journals such as IEE, Springer, Science direct, and so on. Here we will show the journal sources distribution.

After finding the related paper, the next step is to review. The review process that we do is first, we sort out titles that match the topic we will review. Second, we start to look at the abstracts of the papers, and if we think they are related, we will use them to review at the next stage. Third, we look at the proposed method, and whether the method is deep learning or not, if the method used is deep learning, then we will use it. Fourth, we look at the dataset. The dataset that we will compare here will only use the most widely used dataset in the ABSA task namely, SemEval dataset and Twitter dataset. The last stage is that we look at the results of each paper, namely the accuracy and F1 values, which we then compare with several studies.

Aspect-Based Sentiment Analysis

Online seller products on the e-commerce often ask their customers to review the products that they have purchased and the related services. As e-commerce is becoming more and more popular, the number of customer reviews that

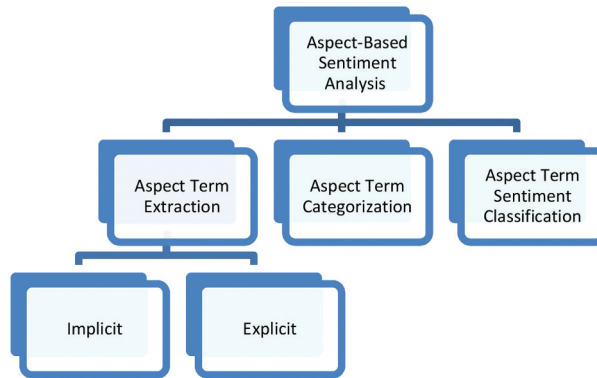


Figure 3. The sub-tasks of ABSA.

a product receives grows rapidly. For a popular product, the number of reviews can be in hundreds or even thousands. This makes it difficult for a prospective customer to read them to make the right decision on whether to purchase the product. It also makes it difficult for the product manufacturers to keep track and to manage customer opinions. These problems can be solved with a sentiment classification or sentiment analysis.

In the case of customer reviews, it has been observed that customers always comment multiple aspects in one sentence at the same time. This kind of problem cannot be solved by traditional sentiment classification. ABSA comes to overcome this issue. Instead of predicting the overall sentiment polarity, fine-grained ABSA is proposed to better understand reviews than traditional sentiment analysis. Suppose that a restaurant customer gives a comment on the restaurant he visited “*The price is reasonable although the service is poor.*” There are two targets “*price*” and “*service*” mentioned in that sentence and they express positive and negative, respectively. The example can show that the polarity will be opposite when different aspects are considered.

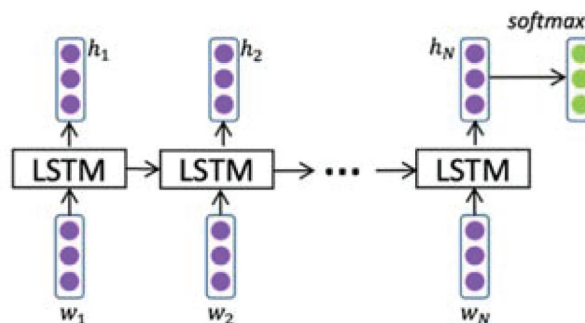


Figure 4. Architecture of standard LSTM (Wang et al. 2016).

In this survey, we divided aspect-based sentiment analysis (ABSA) into three sub-tasks: aspect term extraction, aspect term categorization, and aspect term sentiment classification. [Figure 3](#) shows the sub-tasks in ABSA.

Aspect Term Extraction (ATE)

Aspect term extraction (or aspect identification or opinion target extraction) is a subtask of ABSA that deals with identifying different aspects mentioned within a given sentence. The major task of ABSA is to extract aspects from the review text (Da'u and Salim 2019). Traditional approaches to aspect term extraction typically rely on using handcrafted features, linear, and integrated network architectures. Although these methods can achieve good performances, they are time-consuming and often very complicated. In real-life systems, a simple model with competitive results is generally more effective and preferable over complicated models.

Aspect term extraction (ATE) for opinion mining was first studied by Hu, Liu, and Street (2004). They introduce the distinction between explicit and implicit aspects. Explicit aspects are concepts that explicitly denote targets in the opinionated sentence. For instance, in the example sentence “*I love the touchscreen of my phone but the battery life is so short.*” *Touchscreen* and *battery life* are explicit aspects as they are explicitly mentioned in the sentence. On the other hand, an aspect can also be expressed indirectly through an implicit aspect (Chen and Chen 2016), e.g., in the sentence “*This camera is sleek and very affordable,*” which implicitly provides a positive opinion about the aspect “*appearance*” and “*price*” of the entity camera. As mentioned in the first example, it contains two aspects, namely “*touchscreen*” and “*battery life.*” In this case, applying a sentence level polarity detection technique would mistakenly result in a polarity value close to neutral since the two opinions expressed by the user are opposite. Hence, aspect is necessary to first deconstruct sentence into product features and then assign a separate polarity value to each of these features. Given another example sentence, “*The screen of my phone is really nice and its resolution is superb*” for a phone review contains positive polarity, i.e., the author likes the phone. However, more specifically, the positive opinion is about its screen and resolution. These concepts are thus called opinion targets, or aspects, of this opinion. The task of identifying the aspects in the given opinionated text is called aspect extraction (Poria, Cambria, and Gelbukh 2016).

Rule-based methods have become a popular method in early research ATE was proposed by Poria et al. (2014), they aim to solve the problem of ATE from product reviews by proposing a novel rule-based approach that exploits common-sense knowledge and sentence dependency trees to detect both explicit and implicit aspects (Liu et al. 2016) proposed double propagation as the base and improved its results dramatically through aspect recommendation, semantic similarity-based, and aspect association-based. Since rule-based has grown in

popularity, some studies have improved how automated systems identify the appropriate rule for ATE (Liu and Gao 2016; Rana and Cheah 2017). Rule-based methods usually do not group extracted aspect terms into categories.

In the supervised approaches, machine-learning systems are trained on manually annotated data to extract targets in the reviews. The most common techniques employed in supervised approaches are decision trees, Support Vector Machines (SVM) (Manek and Shenoy 2016), K-nearest neighbor (Shah et al. 2020), Naïve Bayesian classifiers (Kaur 2021), and some in neural networks (Kessler 2009; Pontiki et al. 2016). However, unsupervised methods are adopted to avoid reliance on labeled data needed for supervised learning (He et al. 2017), automatically extract product features using syntactic and contextual patterns without the need of annotated data (Liu et al. 2015; Liu and Gao et al. 2016).

Conditional Random Field (CRF), one of the supervised-based methods, is a promising method used in the named entity recognition (NER) problems. Due to this reason, most of researchers use CRF in aspect extraction tasks since the NER and aspect extraction have a similar problem. Shu et al. (2017b) used the lifelong learning method combined with CRF (L-CRF) to leverage the knowledge gained from extraction results of previous domains, which are unlabeled data, to improve its extraction.

The methods mentioned above have their own limitations. CRF is a linear model, so it needs a large number of features to work well, linguistic patterns need to be crafted by hand, and they crucially depend on the grammatical accuracy of the sentence. Recently, methods based on the deep learning model make promising results in any sentiment analysis task including aspect extraction. To overcome the above limitations, Poria, Cambria, and Gelbukh (2016) proposed a deep convolutional neural network (CNN), a non-linear supervised classifier that can more easily fit the data. They also introduced specific linguistic patterns and combined a linguistic pattern approach with a deep learning approach for the ATE task.

CNN is widely used in research in the field of image processing. Since CNN does not require complex computations, Kim 2014 proposed CNN for sentence classification to get a promising result. Due to that result, CNN has become more popular in text classification tasks especially ATE tasks. In ATE task, each aspect has a different domain depending on the context of the sentence. With the modification of the embedding layer, Xu et al. (2018) used a double embedding layer to give better performance for the CNN layer (DE-CNN). They used general embedding and domain-specific embedding to make a better performance.

Using the same double embedding method, Shu et al. (2017) modified the standard CNN in (Xu et al. 2018). They called controlled CNN (Ctrl), which has the idea by asynchronously updating control modules and CNN layers, it can boost the performance of a single task. Da' u and Salim (2019) with leveraging different embedding layers used multichannel convolutional neural

network (MCNN). Similar to DE-CNN, they also use two-word embedding, general embedding, and domain embedding, but the difference is that in this approach, they use an additional embedding layer, namely Part of Speech (POS) tagging and the convolutional layer used to extract local features from the embedding layer. Jabreel, Hassan, and Moreno (2018) used a bidirectional gated recurrent neural network to extract the target of the tweets.

Aspect Term Categorization (ATC)

The second sub-task of ABSA is to cluster synonymous aspect terms into categories where each category represents a single aspect, which we call an aspect category (Mukherjee and Liu 2012). For the given example, in the sentence “*I have to say they have one of the fastest delivery time in the city,*” the aspect term is “*delivery time.*” For example, we can cluster aspect terms with similar meaning into categories where each category represents a single aspect (e.g., cluster “*delivery time,*” “*waiter,*” and “*staff*” into one aspect **service**). Ganu, Elhadad, and Marian (2009) adopted a category-specific sentiment classification in restaurant reviews. They first identified six basic categories for a restaurant. Then established that the textual entity of the review is a better indicator than the other meta-information such as star ratings.

Mukherjee and Liu (2012) solved the problem in a different setting where the user provides multiple headwords for multiple aspect categories and the model extracts and groups the aspect terms into categories simultaneously by proposing two new statistical models Seeded Aspect Sentiment model (SAS) and the Maximum-Entropy Seeded Aspect Sentiment model (ME-SAS). Aspect category classification (ACC) and aspect term extraction (ATE) are often treated independently, even though they are closely related. Intuitively, the learned knowledge of one task should inform the other learning task. Xue et al. (2017) proposed a multi-task learning model based on neural networks (MTNA) to solve both tasks. ACC as a supervised classification task where the sentence should be labeled according to a subset of predefined aspect labels and ATE as a sequential labeling task where the word tokens related to the given aspects should be tagged according to a predefined tagging scheme, such as IOB (Inside, Outside, Beginning). They combine BiLSTM for ATE and CNN for ACC together in a multi-task framework. Senarath, Jihan, and Ranathunga (2019) proposed mixture classifiers for aspect extraction, combining the proposed improved CNN with an SVM that uses state-of-the-art manually engineered features.

Akhtar, Garg, and Ekbal (2020) proposed two strategies for joint learning of the two tasks (aspect term extraction and aspect sentiment classification). The first approach is based on an end-to-end framework, where the two tasks are solved in a sequence. BiLSTM-CNN is used in this task, the BiLSTM is responsible for learning the sequential pattern of tokens in a sentence. In addition, the self-attention module aims to assist the system in learning the importance of other tokens in the sentence, for tagging the current token as inside (or outside) of an aspect term, and using the softmax function for BIO classification. The CNN layer is used for capturing the local context of each token. In contrast, the second approach combines the two tasks into a single task and solves them as one task in architecture. They classified each token into one of the nine classes, i.e., B-Positive, I-Positive, B-Negative, I-Negative, B-Neutral, I-Neutral, B-Conflict, I-Conflict, and O.

Aspect Term Sentiment Classification

The last sub-task in ABSA is aspect term sentiment classification. After extracting the aspect term in a review sentence and classifying it, now the proposed of ABSA is classified to the sentiment polarity. Tang et al. (2016) proposed a target-dependent sentiment classification using Long Short-Term Memory. The term target here is the same meaning as the term aspect that we use in this survey. Target-dependent sentiment classification is typically regarded as a kind of text classification problem in the literature. Therefore, a standard text classification approach such as a feature-based Support Vector Machine (Jiang et al. 2011; Pang et al. 2002) can be naturally employed to build a sentiment classifier. For example, Jiang et al. (2011) manually designed target-independent features and target-dependent features with expert knowledge, syntactic parser, and external resources. Despite the effectiveness of feature engineering, it is labor-intensive and unable to discover the discriminative or explanatory factor of data. To handle this problem, Tan et al. (2014) proposed a method to transfer a dependency tree of a sentence into a target-specific recursive structure and use an Adaptive Recursive Neural Network to learn higher-level representation. Otherwise, Vo and Zhang (2015) use rich features including sentiment-specific word embedding and sentiment lexicons.

Sentiment polarity in ABSA not only depends on the aspect on the sentence review but also depends on the content. In Wang et al. (2016), they found that the sentiment polarity of a sentence is highly dependent on both content and aspect. For example, “*Staff are not that friendly, but the taste covers all.*” Will be *positive* if the aspect is *food* but *negative* when considering the aspect *service*. Polarity could be opposite when different aspects are considered.

Method for Aspect-Based Sentiment Classification (ABSA)

Traditional approaches to solve ABSA problems are to manually design a set of features. These conventional approaches mainly focus on designing a set of features such as bag-of-words, sentiment lexicon to train classification (e.g., SVM) (Jiang et al. 2011), rule-based methods (Ding et al. 2008), and statistic-based methods (Jiang et al. 2011; Zhao et al. 2010). However, feature engineering is labor-intensive and almost researches its performance bottleneck (Ma et al. 2017) because it needs a lot of labeled data. With the development of deep learning techniques, some researchers have designed effective neural networks to automatically generate useful low-dimensional representations from targets and their contexts and obtain a promising result on the ABSA task. In this part, we will discuss the methods used in ABSA.

Method Based on Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) are a type of Neural Network where the output from the previous step is fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases when it is required to predict the next word of a sentence, the previous words are required, and hence there is a need to remember the previous words. Thus, RNN came into existence, which solved this issue with the help of a hidden layer. Here we will describe RNN based that researcher use in ABSA tasks.

Long Short-Term Memory (LSTM)

Standard RNN has gradient vanishing or exploding problems, where gradients may grow or decay exponentially over long sequences. In order to overcome the issues, Long Short-term Memory Network (LSTM) was developed and achieved superior performance (Hochreiter and Schmidhuber 1997). Compared with traditional feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images) but also entire sequences of data (such as text, speech, and video). In LSTM architecture, there are three gates (input gate, forget gate, and output gate) and a cell memory state. The input gate is responsible for the addition information to the cell state. A forget gate is responsible for removing information from the cell state. The information that is no longer required for the LSTM to understand things or the information that is of less importance is removed via the multiplication of a filter. This is required for optimizing the performance of the LSTM network. Output gate has the job of selecting useful

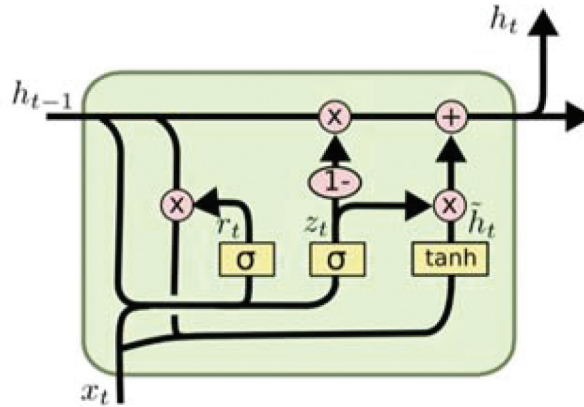


Figure 5. Gated Recurrent Unit (GUR) Source: <https://colah.github.io/posts/2015-08-Understanding-LSTMs>.

information from the current cell state and showing it out as output is done via the output gate. Figure 4 illustrates the architecture of standard LSTM. Figure 5 Figure 6 Figure 7

More formally, each cell in LSTM can be computed as follows:

$$X = \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} \quad (1)$$

$$f_t = \sigma(W_f \cdot X + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot X + b_i) \quad (3)$$

$$o_t = \sigma(W_o \cdot X + b_o) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c \cdot X + b_c) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Where $W_i, W_f, W_o \in \mathbb{R}^{d \times 2d}$ are the weighted matrices and $b_i, b_f, b_o \in \mathbb{R}^d$ are biases of LSTM to be learned during training, parameterizing the transformation of the input, forget, and output gates, respectively. σ is the sigmoid function and \odot stands for element-wise multiplication. x_t includes the inputs of the LSTM cell unit, representing the word embedding vectors w_t in Figure 4. The vector of the hidden layer is h_t .

Since LSTM can capture sequence model, most researchers use LSTM-based method in text classification and sentiment classification tasks, especially in ABSA. Target-dependent sentiment classification (TD-LSTM) was first proposed by Tang et al. (2016). The proposed method is an extension of

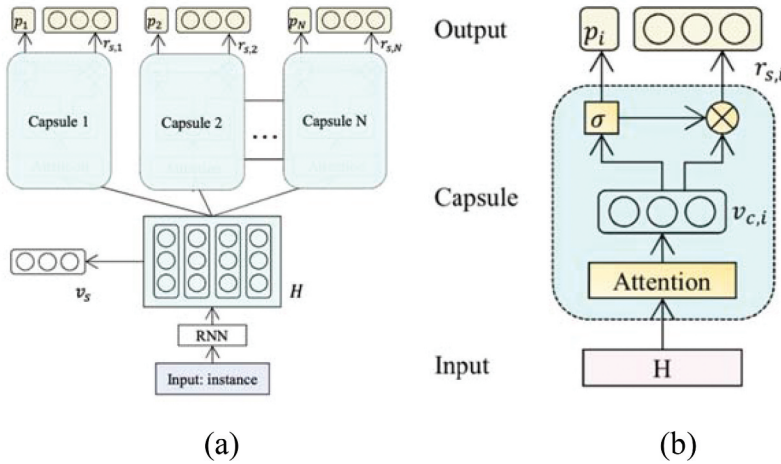


Figure 6. a. Capsule structure on sentiment analysis; b. architecture of a single capsule (Wang, Sun, and Han 2018).

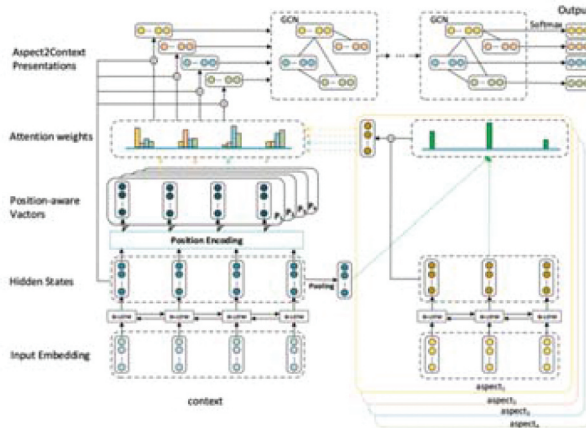


Figure 7. An example to illustrate the usefulness of the sentiment dependencies between multiple aspects (Zhao, Hou, and Wu 2019).

the LSTM model, which the standard LSTM model cannot capture any target information so that it predicts the same result for different targets in a sentence. The general proposed model is to use two LSTM neural networks, left and right LSTM, to model the preceding, and following contexts, respectively where the target is placed in the middle of the sentence.

Wang et al. (2016) proposed a new idea with adding aspect embedding in LSTM (ATAE-LSTM). Apart from adding an aspect embedding, they also use the attention mechanism. This is the first time to propose aspect embedding in ABSA task. With the use of attention mechanism in

Table 1. Comparison proposed method in aspect extraction task with different domains.

No.	Proposed method	Dataset	Word embedding	Base method	Results (F1)		
					Laptop	Restaurant	Twitter
1	CTRL (2019)	SemEval 2014 (Laptop) SemEval 2015 (Restaurant)	GloVe + Domain Embedding	Controlled CNN	82.73	75.64	-
2	MCNN (2019)	SemEval 2014 (Restaurant)	Word2Vec + Amazon +Yelp	CNN + POS Tag	80.63	86.89	-
		SemEval 2015 (Restaurant) Semeval 2016 (Restaurant)			-	72.65	-
3	DE-CNN (2018)	SemEval 2014 (Laptop)	Glove + Amazon + Yelp	Double Embedding + CNN	81.59	-	-
		SemEval 2016 (Restaurant)			-	74.37	-
4	TI-biGRU (2017)	Twitter	GloVe	biGRU	-	-	89.25

ATAE-LSTM, Ma et al. (2017) considered to separate modeling of targets, especially with the aid of contexts, they propose the interactive attention networks (IANs), which is based on the LSTM model and attention mechanism to interactively learn attention in the contexts and targets, and generate representation of the targets and contexts separately.

Previous LSTM-based methods mainly focused on modeling text separately, while they modeled aspects and texts simultaneously using LSTMs. Furthermore, the target representation and text representation generated from LSTMs interact with each other by an attention-over-attention (AOA) module than inspired by the use of AOA in question-answering (Cui et al. 2017). Meanwhile, He et al. (2018) proposed a modification of attention modeling based on LSTM.

Joint attention LSTM network (JAT-LSTM) is proposed to combine the aspect attention and sentiment attention to construct a joint attention LSTM network (Cai and Li 2018). The model concatenates the aspect terms embedding and sentiment terms embedding with sentence embedding as the input of the LSTM network to make the input information of the LSTM network richer.

In Table 2, we present the results of the accuracy and F1 studies of several proposed ABSA models that use LSTM as the base model. We can see that JAT-LSTM has the highest accuracy among other models with an LSTM base in the SemEval 2014 dataset. But for SemEval 2015, 2016, and Twitter datasets, here we can combine because only one approach uses these datasets. We will combine later with another method. Laptop and Rest stand for the SemEval 2014 domain, Twitter stands for the Twitter domain, Rest 15 stands for the

SemEval 2015 Restaurant domain, and Rest 16 stands for the SemEval 2016 Restaurant domain. Value in bold indicate the model that gets the highest score.

Gated Recurrent Unit (GRU)

Gated Recurrent Units (GRUs) is a variant of LSTM, introduced in (Cho et al. 2014). GRUs were designed to have more persistent memory, making them very useful to capture long-term dependencies between the elements of a sequence. It combines the forget gate and input gate into a single update gate. It also merges the cell state and hidden state, among other changes. The resulting model is simpler than standard LSTM models and has become a popular model in many tasks. GRU has *reset* (r_t) and *update* (z_t) gates. The former has the ability to completely reduce the past hidden state h_{t-1} if it considers that it is irrelevant to the computation of the new state, whereas the latter is responsible for determining how much of h_{t-1} should be carried forward to the next state h_t . The output h_t of a GRU depends on the input x_t and the previous state h_{t-1} , and it is computed as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}; x_t] + b_r) \quad (7)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}; x_t] + b_z) \quad (8)$$

$$\tilde{h}_t = \tanh(W_h \cdot [(r_t \odot h_{t-1}); x_t] + b_h) \quad (9)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (10)$$

r_t and z_t denote the *reset* and *update* gates, respectively, \tilde{h}_t is the candidate output state and h_t is the actual output state at time t . The symbol \odot stands for element-wise multiplications, σ is a sigmoid function and stands for the vector concatenation operation. $W_r, W_z, W_h \in \mathbb{R}^{d_h \times (d+d_h)}$ and $b_r, b_z, b_h \in \mathbb{R}^{d_h}$ are the parameters of the reset and update gates, respectively, where d_h is the dimension of the hidden state.

Since GRU has a simpler model than LSTM, many researchers have started developing ABSA tasks using the GRU model. Most of the ABSA tasks assume that the target or aspect is known before determining the sentiment polarity. Meanwhile, Jabreel, Hassan, and Moreno (2017) proposed a model for target-dependent sentiment analysis of tweet that has the ability to identify and extract the target of the tweets, representing the relatedness between the targets and its contexts and identifying the polarities of the tweets toward the targets. They used the bi-GRU model (TD-BiGRU) to extract the target and determine the sentiment polarities.

Gu et al. (2018), in their work, proved that when an aspect term occurs in a sentence, its neighboring words should be given more attention than other words with long distance. They proposed a position-aware bidirectional attention network (PBAN) based on bidirectional GRU. It not only concentrates on the position information of the aspect terms, but also mutually models the relation between aspect term and sentence by employing bidirectional attention mechanism. Previous works of ABSA have proved that the interaction between aspects and the contexts is important. Otherwise, most of those works ignore the position information of the aspect when encoding the sentence. HAPN (Hierarchical Attention-based Position-aware Network) (Li, Liu, and Zhou 2018) is proposed to solve this problem. They introduce the position embeddings when modeling the sentence and further generate the position-aware representations. In addition, they propose a hierarchical attention-based fusion mechanism to fuse the clues of aspects and the contexts. The results demonstrate that the proposed approach is effective for aspect-level sentiment classification, and it outperforms state-of-the-art approaches with remarkable gains.

In Table 3, we present the results of the accuracy and F1 studies of several proposed ABSA models that use GRU as the base model. We can see that HAPN has the highest accuracy among other models with GRU base in the SemEval 2014 dataset but if we compare with the LSTM base model, still LSTM in JAT-LSTM has the highest performance. For the Twitter dataset, TD-biGRU gets the highest accuracy performance.

Method Based on Convolutional Neural Network (CNN)

Convolutional Neural Networks (ConvNets or CNNs) are a category of Neural Networks that have proven very effective in areas such as image recognition and classification. CNNs have been successful in identifying faces, objects, and traffic signs apart from powering vision in robots and self-driving cars. In the past years CNN shows breakthrough results in some NLP tasks, one particular task is sentence classification (Kim 2014), i.e., classifying short phrases into a set of pre-defined categories.

Most Convolutional Neural Network (CNN) methods are used in text classification tasks and can achieve state-of-the-art performance on many standard sentiment classification datasets (Le, Cerisara, and Denis 2017). The CNN model consists of an embedding layer, a one-dimension convolutional layer, and a max-pooling layer. The embedding layer is usually initialized with pre-trained embeddings such as Glove (Pennington, Socher, and Manning 2014).

Since the ABSA is a part of text classification, some researchers used a method based on CNN. CNN cannot stand alone to solve the ABSA problem, some studies combine the CNN with a sequence model such as

LSTM. LSTM employed in ABSA have weaknesses, such as lacking position invariance and lacking sensitivity to local key patterns. LSTM combines with attention mechanism widely used in ABSA task, but it is time inefficient because they processed the given sequence in a circular manner and need more training time. Meanwhile, the CNN model can address these limitations because the convolution layer can easily perform parallel operations during training without waiting for the results of the previous step, but it is weak at capturing long-distance dependency and modeling sequence information. Many researchers try to tackle this limitation such as Xue and Li (2018) instead of using attention mechanism, they proposed a model based on convolutional neural network and gating mechanisms with aspect embedding (GCAE), which is more accurate and efficient.

Sparse attention-based separable dilated convolutional neural network (SA-SDCNN) (Gan et al. 2019) is a proposed method that is composed of multi-channel embedding layer, separable dilated convolutional module, sparse attention layer, and output layer. The multichannel embedding using three different embedding layers, namely word2vec, glove, and SSWE for sentiment embedding.

Ren et al. (2020) proposed the distillation network (DNet), a lightweight and efficient sentiment analysis model based on gated convolutional neural networks. The proposed model first encodes the sentence with gated convolutional networks to control what information is useful in predicting the sentiment polarity. They use the gating units from GCAE (Xue and Li, 2018) to extract aspect-sensitive information. Meanwhile, some researchers argue that the position of aspect and its context are important, Wu et al. (2020) proposed a relative position and aspect attention encode model (RPAEN) for ABSA based on convolutional neural networks. The proposed model first introduces a position encode layer to encode the relative position of the aspect term in the text so that the corresponding relative position information can be incorporated into the model and more conducive to the sentiment analysis of the current aspect term. Then, they use the aspect attention mechanism instead of the general attention mechanism to better capture the dependence between all words in the text and aspect term. Table 4 shows the results of some popular ABSA methods based on the CNN model. Table 5 Table 6

Method Based on Memory Network

Memory network is a general machine learning framework introduced by Weston, Chopra, and Bordes (2015), coupled with multiple-hop attention attempts to explicitly focus only on the most informative context area to infer the sentiment polarity toward the target word. Its central idea is inference with a long-term memory component, which could be read,

Table 3. Performance comparison GRU based model.

No.	Method	Dataset	Accuracy (%)			F1 (%) Twitter
			Laptop	Rest	Twitter	
1	ABAE-Bi-GRU (2019)	SemEval 2014 &Twitter	-	81.2	71.2	-
2	HAPN (2018)	SemEval 2014	77.27	82.23	-	-
3	PBAN (2018)	SemEval 2014	74.12	81.16	-	-
4	TD-biGRU (2017)	Twitter	-	-	72.25	70.47

written to, and jointly learned with the goal of using it for prediction. Formally, a memory network consists of a memory m and four components, I , G , O , and R , where m is an array of objects such as an array of vectors. Among these four components, I converts input to internal feature representation, G updates old memories with new input, O generates an output representation given a new input and the current memory state. R outputs a response based on the output representation. The method based on memory networks explicitly hold the context information through memory and acquire relation between the target and context through attention mechanisms.

Tang, Qin, and Liu (2014) proposed a deep memory network (MemNet) for aspect-level sentiment classification. The method can capture the importance of each context word when inferring the sentiment polarity of an aspect. This approach consists of multiple computational layers with shared parameters.

Attention Mechanism

Attention models have recently gained popularity in training neural networks and have been applied to various natural language processing tasks, including machine translation, sentence summarization, sentiment classification, and question answering. It was proposed in machine translation for the purpose of selecting referential words in original language for words in counterparts language before translation. Rather than using all available information, attention mechanism aims to focus on the most pertinent information for a task. Since the success of applying attention network on translation task (Bahdanau, Cho, and Bengio 2015; Luong, Pham, and Manning 2015), lots of work have designed attention mechanism networks to address the aspect-based sentiment analysis. In ABSA task, attention mechanism able to capture the importance of each context word toward a target by modeling their semantic associations so that can obtain comparable result.

A feature-enhanced attention network for target-dependent sentiment classification (FANS) has been proposed by Yang et al. (2018). They improve the attention model by using multi-view co-attention network (MCN) to learn a better multi-view sentiment awareness and target-specific sentence

representation via interactively modeling the context words, target words, and sentiment words. Multi-attention network (MAN) also proposed by Xu et al. (2020). The proposed model uses intra- and inter-level attention mechanisms. Intra-level employs a transformer encode instead of a sequence model to reduce training time. Meanwhile the inter-level attention mechanism uses a global and local attention module to capture differently grained interactive information between aspect and context. Similar to Zhang and Lu (2019), Zhang et al. (2020) also used BERT as a pre-trained to model the data. The proposed model is called multiple interactive attention network (MIN). After the pre-trained process is done, they use the partial transformer to obtain a hidden state in parallel.

Park, Song, and Shin (2020) proposed a deep learning model for ABSA that combines the location attention and content attention. There are two models proposed: one implementing Holistic Recurrent content attention on target-dependent memories from One-directional one-layered networks (HRT_one), the other from Bi-directional bi-layered networks (HRT_Bi). These two models employ target-dependent LSTM to produce memories from an input sentence and GRU cells for integrating sentence representations generated from different content attention weights on memories.

Method Based on Capsule Network

Capsule network was introduced to improve representation limitation of CNN and RNN models by extracting features in the form of vector. Capsule network in sentiment analysis was proposed by Wang, Sun, and Han (2018). They introduce the RNN-capsule for sentiment analysis, the capsule model based on Recurrent Neural Network (RNN). Each capsule is capable not only predicting the probability of its assigned sentiment, but also reconstructing the input instance representation. Compare with most existing neural network models for sentiment analysis, RNN-capsule model does not heavily rely on the quality of input instance representation. This model does not require any linguistic knowledge. The number of capsule N is the same as the number of

Table 4. Performance comparison CNN based model.

No.	Term	Accuracy					Results (F1)		
		Laptop	Restaurant	Twitter	Amazon	Yelp	Laptop	Restaurant	Twitter
1	RPAEN (2020)	74.1	81.2	73.0	-	-	70.5	72.5	70.1
2	RPAEN-BERT (2020)	80.6	85.2	74.8	-	-	76.8	78.0	72.9
3	DNet-GLU (2020)	73.35	79.46	72.84	-	-	68.86	69.67	69.63
4	Dnet-GTU (2020)	73.98	79.38	72.59	-	-	68.28	68.12	68.27
5	SA-SDCNN (2020)	-	-	74.16	73.57	81.36	-	-	73.24
6	GCAE (2018)	69.4	77.28	-	-	-	-	-	-

Table 5. Performance comparison model with attention mechanism.

No.	Term	Accuracy (%)										F1 (%)				
		Laptop	Restaurant	Twitter	Rest 15	Rest 16	Amazon	Yelp	Laptop	Restaurant	Twitter	Rest 15	Rest 16			
1	MIN (2020)	79.78	82.68	73.84	82.84	89.12	-	-	75.49	74.05	72.37	69.26	68.67			
2	MAN (2020)	78.13	84.38	76.56	82.65	85.87	-	-	73.20	71.31	72.19	69.10	73.28			
3	PAHT (2020)	73.82	80.45	-	82.05	85.58	-	-	68.37	70.07	-	68.85	65.00			
4	MAN-BERT (2019)	76.35	81.43	71.35	-	-	-	-	88.13	91.22	87.92	-	-			
5	HRT_One (2020)	-	-	-	-	-	68.69	72.57	72.08	-	-	-	-			
6	HRT_Bi (2020)	-	-	-	-	-	70.83	74.09	71.98	-	-	-	-			
7	FANS (2018)	-	-	71.2	-	-	-	-	-	-	68.8	-	-			
8	RAM (2017)	74.49	80.23	69.36	-	-	71.35	70.80	67.30	-	-	-	-			

Table 6. Performance comparison based on Capsule and Graph models.

No.	Method	Accuracy (%)					F1 (%)				
		Laptop	Restaurant	Twitter	Rest 15	Rest 16	Laptop	Restaurant	Twitter	Rest 15	Rest 16
1	IACapsNet (2019)	76.80	81.79	75.01	-	-	73.29	73.40	73.81	-	-
2	TransCap (2019)	72.65	78.84	-	-	-	68.77	69.70	-	-	-
3	SDGCN (2020)	81.35	83.57	-	-	-	78.34	76.47	-	-	-
4	GCNAs (2019)	70.53	78.12	70.09	-	-	-	-	-	-	-
	GCNAc (2019)	69.74	77.85	70.66	-	-	-	-	-	-	-
5	TD-GAT-GloVe (2019)	73.4	81.2	-	-	-	-	-	-	-	-
	TD-GAT-BERT (2019)	80.1	82.8	-	-	-	-	-	-	-	-
6	SA-GCN (2019)	81.7	85.8	-	-	-	78.8	79.7	-	-	-
7	ASGCN-DT (2019)	74.14	80.86	71.55	79.34	88.69	69.24	72.19	69.68	60.78	66.64
	ASGCN-DG (2019)	75.55	80.77	72.15	79.89	88.99	71.05	72.02	70.40	61.89	67.48
8	AEGCN-GloVe (2020)	75.91	81.04	73.16	79.95	87.39	71.63	71.32	71.82	60.87	68.22
	AEGCN-BERT (2020)	78.73	82.58	75.04	82.71	89.61	74.22	73.40	73.68	69.00	73.93

sentiment categories to be modeled, each corresponding to one sentiment category. For example, three capsules are used to model three fine-grained sentiment categories: "positive," "neutral," and "negative."

Since the capsule model used in sentiment analysis task, Wang et al. (2019) modified the RNN-capsule to suit the ABSA task. They propose the aspect-level sentiment capsule model (AS-Capsule), which is capable of performing aspect detection and sentiment classification simultaneously, in a joint manner. In order to solve the problem of lacking aspect-level labeled data, Chen and Qian (2019) proposed a Transfer Capsule Network (TransCap) model for transferring document-level knowledge to aspect-level sentiment classification. Document-level labeled data like reviews are easily accessible from online websites. Meanwhile, the publicly available dataset for ABSA often contains limited data number of training examples.

Another modification of the capsule network was proposed by Du, Sun, and Wang (2019). The proposed model combines the capsule network with the attention mechanism. This capsule network in aspect-level sentiment analysis is used to tackle the overlapped features by feature clustering. They use EM routing algorithm to cluster features and to construct vector-based feature representation. Furthermore, interactive attention mechanism was introduced in the capsule routing procedure to model the semantic relationship between aspect terms and context.

Method Based on Graph Neural Network

Graph Neural Networks (GNNs) were introduced in (Sperduti and Starita 1997) and (Gori, Monfardini, and Scarselli 2005) as a generalization of recursive neural networks that can directly deal with a more general class of graphs, e.g., cycling, directed, and undirected graphs. GNNs consist of an iterative process, which propagates the node states until equilibrium; followed by a neural network, which produces an output for each node based on its state. Jeon et al. (2019) proposed graph-based aspect and rating classification, which utilizes multi-modal word co-occurrence network to solve aspect and sentiment classification. The graph-based aspect ratings classification framework builds word co-occurrence network from a given corpus, defining words as different models if their source document is labeled with different aspects or sentiment categories. Then, the model computes word-aspect dispersion score and word-rating dispersion score from the network, which are then concatenated and used as input for a feedforward neural network for aspects and rating classifications.

Another approach based on graph methods proposed by Zhao, Hou, and Wu (2019) is a method to model Sentiment Dependencies with Graph Convolutional Networks (SDGCN). For every node in the graph, GCN encodes relevant information about its neighborhoods as a new feature representation vector. An aspect is treated as a node, and an edge represents the sentiment dependency relation of two nodes. The model learns the sentiment dependencies of aspects via that graph.

Graph Convolutional Networks (GCNs) often show the best performance with two layers, and deeper GCNs do not bring additional gain due to the over-smoothing problem. Hou, Huang, and Wang (2019) designed selective attention-based GCN block (SA-GCN) to find the most important context words and directly aggregate this information into the aspect-term representation.

The success of RNN and CNN in ASBA tasks also has shortcomings namely, they do not take full account of the entire text structure and the relation between words in a given document. To overcome these shortcomings, Chen (2019) proposed a novel neural network method (GCNSA) in which the text is treated as a graph and the aspect in the specific area of the graph. They performed the convolutional operation on the text graph to obtain a full-text hidden state and introduced an extended structural attention model implemented by LSTM to capture certain information. Unlike the previous methods, 2019Huang et al. (2019) represented a sentence as a dependency graph instead of a word sequence. They proposed a novel target-dependent graph attention network (TD-GAT), which explicitly utilizes the dependency relationship among words. In their experiment, they try to use two different pre-trained data; GloVe and BERT.

The use of the attention mechanism in a graph convolutional network (GCN) was proposed by Hou et al. (2019). They designed selective attention-based GCN block (SA-GCN) to find the most important context words and directly aggregate this information into the aspect-term representation by applying GCN on the dependency tree. A similar method using dependency tree was proposed by Zhang, Li, and Song (2019) called Aspect-specific Graph Convolutional Network (ASGCN). It starts with a bi-LSTM layer to capture contextual information regarding word orders. Then, a multi-layer graph convolution structure is implemented on top of the LSTM output to obtain aspect-specific features. Recently, Xiao et al. (2020) proposed targeted sentiment classification based on attention encoding and graph convolutional networks (AEGCN). The proposed model with the BERT pre-trained model, composed of a multi-head self-attention improved graph convolutional network built over the dependency tree of a sentence.

Word Embeddings

Word embeddings are the representation of document vocabulary. It allows words with similar meaning to have a similar representation. Word embeddings give the impressive performance of deep learning methods on challenging natural language processing problem. Word embeddings are a class of techniques where individual words are represented as real-values vectors in a predefined vector space. Each word is mapped to one vector and the vector values are learned in a way that resembles a neural network, and hence the technique is often lumped into the field of deep learning. In this paper, we will provide a brief discussion of several word embedding methods used for the ABSA task.

Word2vec

Word2vec (Mikolov et al. 2013) is a statistical method for natural language processing uses a neural network model to learn word associations from a large corpus of text. Word2vec represents each distinct word with a particular list of number called a vector. Two different models were introduced as a part of the word2vec approach to learn the word embedding, namely Continuous Bag-of-Words (CBOW) and Skip Gram. When we reviewed several papers related to ABSA, we found several models using word2vec. We present several proposed models that use word2vec as word embedding in Table 7.

Glove

Glove, Global Vectors for Word Representation is an extension to the word2vec method for efficiently learning word vectors, developed by (Pennington, Socher, and Manning 2014). It is based on matrix factorization techniques on the word-context matrix. Glove mapping words into a meaningful space where the distance between words is related to semantic similarity. Many researchers in the NLP field use Glove as their word embedding, especially in the ABSA task. We describe several studies on ABSA using GloVe as word embedding in [Table 8](#).

BERT

BERT, Bidirectional Encoder Representations from Transformers, is a language representation model designed to pretrained deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. BERT can be fine-tuned with just one additional output layer to create a state-of-the art model for a wide rank task (Kenton, Kristina, and Devlin 2019). Because of these advantages, several studies have tried to use it on ASBA tasks. We summarize some of these studies in [Table 9](#).

We can say that the use of BERT as a pre-trained word embedding has not been done much in the ABSA field when compared to the use of Glove. So that BERT is still very broad to be developed in solving ABSA problems.

Dataset for ABSA

For the dataset section, we use the most widely used datasets for ABSA tasks, namely SemEval 2014 task 4 and Twitter dataset. We chose the SemEval 2014 task 4 dataset because this dataset was used for the competition in making models for ABSA where the competition was divided into several subtasks, namely aspect term extraction, aspect category polarity, aspect category detection, and aspect category polarity. The dataset consists of two domains, namely the restaurant domain and the laptop domain, which contains comments about products, namely restaurants and laptops. So, this dataset is the most widely used dataset for modeling ABSA problems.

Meanwhile, apart from the SemEval 4014 task 4 dataset, the second most used dataset is the Twitter dataset. Unlike the dataset on SemEval, the Twitter dataset uses comments contained on Twitter written by users. So, there is no special domain used in the Twitter dataset.

SemEval 2014 Dataset

SemEval 2014 dataset were proposed by Pontiki and Pavlopoulos (2014) that contained manually annotated reviews of restaurants and laptops. Here we will describe the details about SemEval dataset. Table 10 describes the size of the dataset in the sentence and Table 11 describes the aspect terms and their polarities per domain.

SemEval 2014 dataset with laptop and restaurant domain has become a popular dataset among researchers in ABSA tasks. We combine the above proposed methods and put them into a chart. We divide the chart with two evaluation performances. Figure 8 shows the accuracy performance and Figure 9 shows the F1 performance.

As we can see from the figure, the highest accuracy performance on both laptop and restaurant domain come from AS-LSTM-CNN. We can say that the best performance is in the combination of convolutional neural networks (CNNs) and Recurrent Neural Networks (RNNs). Meanwhile, the highest F1 performance comes from MAN-BERT, and we can say pre-trained BERT can give better performance than others. From this result, we can develop more combination methods and pre-trained models to get better performance.

Twitter Dataset

Not only product reviews that are used in aspect-based sentiment analysis but also Twitter datasets get much attention to solve the ABSA task. Tan et al. (2014), manually annotated Twitter post comments. They used keywords (such as “bill gates,” “taylor swift,” “xbox,” “windows 7,” “google”) to query the Twitter API. Figure 10 shows the chart performance of the Twitter dataset in accuracy and F1 performance. From those methods, we can see that Gated Recurrent Unit gets the best performance.

From the figure above, we can see that MAN gets the best accuracy performance when we use the Twitter dataset. Meanwhile, if we talk about F1 score performance, MAN-BERT gets the highest score. Still, we can say that the use of pre-trained model BERT can have more impact.

Classification Performance Evaluation Metrics

To measure the performance of classification methods, most researchers used standard evaluation metrics (Yang 1999). Four standard metrics are used in most of the ASBA tasks. In this section, we will describe the four evaluations metrics.

Table 7. Comparison of results using word2vec.

No.	Year	Task	Proposed Method	Dataset	Base Method	F1 (%)			Accuracy (%)		
						Laptop	Restaurant	Twitter	Laptop	Restaurant	Twitter
1	2019	Aspect Term Extraction	Transfer Learning (Lee et al. 2019)	SemEval 2014	CNN + Transfer Learning	87.15	-	-	-	-	-
2	2019	Aspect Term Extraction	MCNN (Da'u and Salim 2019)	SemEval 2014 SemEval 2015 (Restaurant)	CNN + POS Tag	80.63	86.89	72.65	-	-	-
3	2018	Aspect Term Sentiment Classification	JAT-LSTM (Cai and Li 2018)	SemEval 2014 (Restaurant)	LSTM	-	75.71	-	-	-	-
4	2019	Aspect Term Sentiment Classification	SA-SDCNN (Gan et al. 2019)	Twitter	CNN	-	-	73.57	-	-	74.16
5	2019	Aspect Term Sentiment Classification	AS-LSTM-CNN (Jiang et al. 2019)	SemEval 2014	LSTM & CNN	81.13	87.42	-	86.86	90.02	-



Table 8. Comparison of results using GloVe.

No.	Year	Task	Proposed Method	Dataset	Base Method	Results (F1)			
						Laptop	Restaurant	Twitter	Accuracy
1	2020	Aspect Term Sentiment Classification	End-to end (Akhtar, Garg, and Ekbal 2020)	SemEval 2014	BiLSTM + CNN	78.57	83.36	-	-
2	2019	Aspect Term Extraction	CTRL (Shu, Xu, and Liu 2017)	SemEval 2014 (Laptop) SemEval 2015 (Restaurant)	CNN	82.73	75.64	-	-
3	2018	Aspect Term Extraction	DE-CNN (Xu et al. 2018)	SemEval 2014 (Laptop) SemEval 2016 (Restaurant)	Double Embedding + CNN	81.59	-	-	-
4	2017	Aspect Term Sentiment Classification	MTNA(Xue et al. 2017)	SemEval 2014 SemEval 2015 SemEval 2016	BiLSTM + CNN	-	83.65 67.73 72.95	-	-
5	2020	Aspect Term Sentiment Classification	HRT_One (Park, Song, and Shin 2020)	SemEval 2014 Twitter	LSTM	-	-	68.69	72.08
6	2020	Aspect Term Sentiment Classification	HRT_Bi (Park, Song, and Shin 2020)	SemEval 2014 Twitter	Bi-LSTMGRU	-	-	70.83	71.98
7	2019	Aspect Term Sentiment Classification	ABAE-Bi-GRU(Penghua and Dingyi 2019)	SemEval 2014 Twitter	Bi-GRU	-	81.2	71.2	-
8	2018	Aspect Term Sentiment Classification	HAPN (Li, Liu, and Zhou 2018)	SemEval 2014	Bi-GRU	77.27	82.23	-	-
9	2018	Aspect Term Sentiment Classification	LSTM+SynATT+TarRep (He et al. 2018)	SemEval 2014 SemEval 2015 SemEval 2016	LSTM	71.94	80.63	69.23	71.32
10	2018	Aspect Term Sentiment Classification	PBAN (Gu et al. 2018)	SemEval 2014	Bi-GRU	74.12	81.16	-	-
11	2018	Aspect Term Sentiment Classification	AOA-LSTM (Binxuan, Ou, and Carley 2018)	SemEval 2014	Bi-LSTM	74.5	81.2	-	-
12	2017	Aspect Term Sentiment Classification	TD-bigRU (Jabreel, Hassan, and Moreno 2017)	Twitter	Bi-GRU	-	-	72.25	70.47
13	2017	Aspect Term Sentiment Classification	RAM (Sun et al. 2017)	SemEval 2014 Twitter	Bi-LSTM GRU	74.49	80.23	69.36	71.35
14	2017	Aspect Term Sentiment Classification	IAN (Ma et al. 2017)	SemEval 2014	LSTM	72.1	78.6	-	-
15	2016	Aspect Term Sentiment Classification	ATAE-LSTM (Wang et al. 2016)	SemEval 2014	LSTM	68.7	77.2	-	-

(Continued)



Table 8. (Continued).

No.	Year	Task	Proposed Method	Dataset	Base Method	Results (F1)			Accuracy		
						Laptop	Restaurant	Twitter	Laptop	Restaurant	Twitter
16	2015	Aspect Term Sentiment Classification	TD-LSTM (2015)	Twitter	LSTM	-	-	70.8	-	-	69.0
		Aspect Term Sentiment Classification	TC-LSTM (2015)			-	-	71.5	-	-	69.5
17	2020	Aspect Term Sentiment Classification	RPAEN (Wu et al. 2020)	SemEval 2014Twitter	CNN	74.1	81.2	73.0	-	-	-
18	2019	Aspect Term Sentiment Classification	SA-SDCNN (Gan et al. 2019)	Twitter	CNN	-	-	74.16	-	-	73.24
19	2018	Aspect Term Sentiment Classification	GCAE (Xue and Li, 2018)	SemEval 2014	CNN	69.4	77.28	-	-	-	-
20	2018	Aspect Term Sentiment Classification	Tnet (Li, Liu, and Zhou 2018)	SemEval 2014Twitter	CNNBi-LSTM	76.54	80.69	74.97	71.75	71.27	73.60
21	2020	Aspect Term Sentiment Classification	MAN (Xu et al. 2020)	SemEval 2014SemEval 2015SemEval 2016Twitter	Attention	78.13	84.38	76.56	73.20	71.31	72.19
22	2020	Aspect Term Sentiment Classification	HRT_One (Park, Song, and Shin 2020)	SemEval 2014Twitter	LSTM	-	-	-	72.08	-	-
		Aspect Term Sentiment Classification	HRT_Bi (Park, Song, and Shin 2020)		Bi-LSTMGRU	-	-	-	71.98	-	-
23	2017	Aspect Term Sentiment Classification	RAM (Sun et al. 2017)	SemEval 2014Twitter	Bi-LSTM GRU	74.49	80.23	69.36	67.30	-	-

**Table 9.** Comparison of results using BERT.

No.	Year	Proposed Method	Dataset	Accuracy (%)						F1 (%)					
				Laptop	Restaurant	Twitter	Rest 15	Rest 16	Laptop	Restaurant	Twitter	Rest 15	Rest 16		
1	2020	RPAEN-BERT (Wu et al. 2020)	SemEval 2014Twitter	-	-	80.6	85.2	74.8	-	-	76.8	78.0	72.9		
2	2019	MAN-BERT (Zhang and Lu 2019)	SemEval 2014Twitter	76.35	81.43	71.35	-	-	88.13	91.22	87.92	-	-		
3	2019	SDGCN (Zhao, Hou, and Wu 2019)	SemEval 2014	81.35	83.57	-	-	-	78.34	76.47	-	-	-		
4	2019	TD-GAT-BERT (Huang and Carley 2020)	SemEval 2014	80.1	82.8	-	-	-	-	-	-	-	-		
5	2019	SA-GCN (Hou, Huang, and Wang 2019)	SemEval 2014	81.7	85.8	-	-	-	78.8	79.7	-	-	-		
6	2020	AEGCN-BERT (Xiao et al. 2020)	SemEval 2014SemEval 2015SemEval 2016Twitter	78.73	82.58	75.04	82.71	89.61	74.22	73.40	73.68	69.00	73.93		

Table 10. Sizes (sentences) of the datasets.

Domain	Train	Test	Total
Restaurants	3041	800	3841
Laptops	3045	800	3845
Total	6086	1600	7686

Table 11. Aspect terms and their polarities per domain.

Dataset	Positive	Negative	Conflict	Neutral	Total
Laptop-Train	987	866	45	460	2358
Laptop-Test	341	128	16	169	654
Restaurant-Train	2164	805	91	633	3693
Restaurant-Test	728	196	14	196	1134

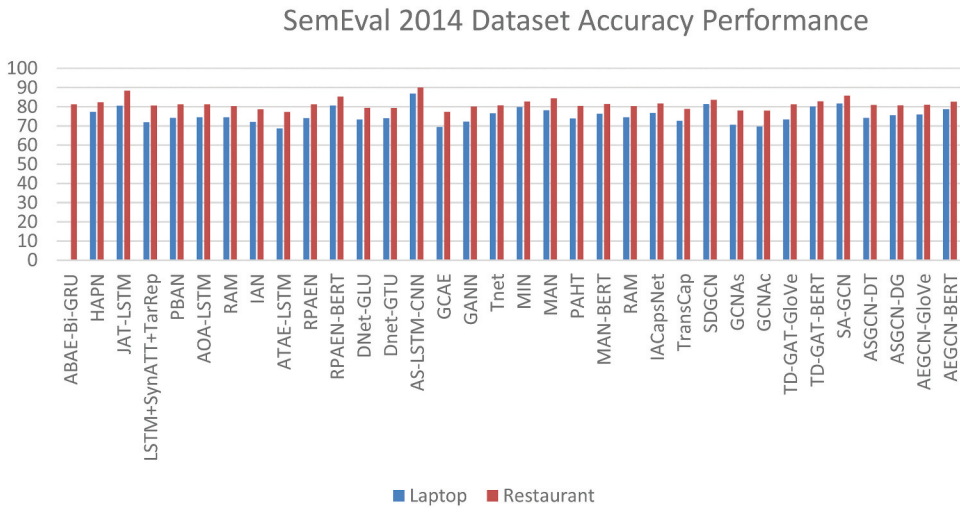


Figure 8. Accuracy comparison in SemEval 2014 dataset.

Precision

Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives. False positive are cases the model incorrectly labels as positive that are actually negative.

$$= \frac{TP}{TP + FP} \tag{11}$$

Recall

A recall is defined as the number of true positives divided by the number of true positives plus false negatives.

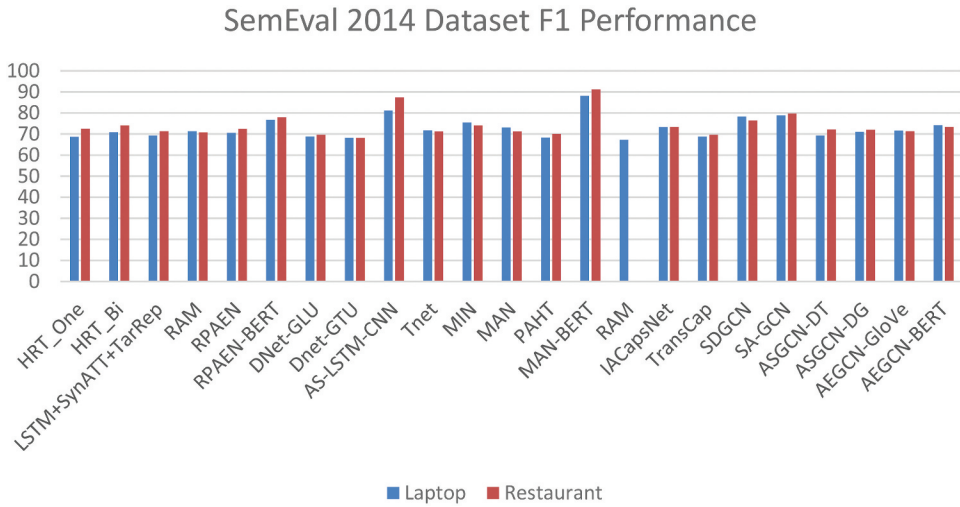


Figure 9. F1 comparison in SemEval 2014 dataset.

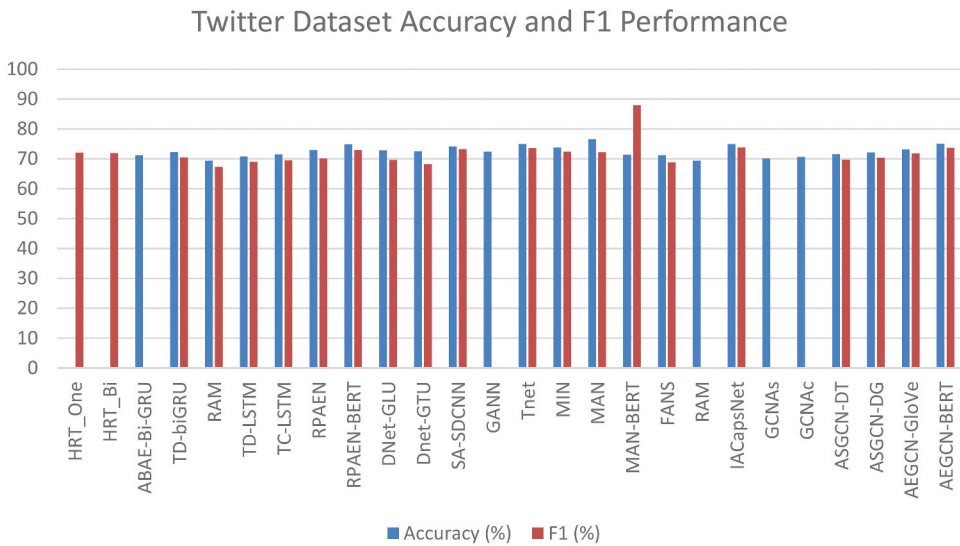


Figure 10. Accuracy and F1 performance in Twitter dataset.

$$r = \frac{TP}{TP + FN} \tag{12}$$

F1-Score

F1-Score (also called F-Score or F-measure) is a measure of a test’s accuracy. It considers both the precision p and the recall r . It is the weighted average of p and r .

$$F_1 = \frac{2pr}{p+r} \quad (13)$$

Accuracy

Accuracy simply measures how often the classifier makes the correct prediction. It is the ratio between the number of correct predictions and the total number of predictions (the number of data points in the test set)

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

Here TP stands for the number of True Positive, TN stands for the number of True Negative, FP stands for the number of False Positive, and FN stands for the number of False Negative.

Conclusion

In this survey, we describe aspect-based sentiment analysis (ABSA). First, we described the task of ABSA, there are three subtasks that we describe namely, aspect term extraction, aspect term categorization, and aspect term sentiment analysis. We provide several models for each task from ABSA. Second, we describe the deep learning methods that are used to solve the ABSA tasks. Finally, we describe two popular datasets that used in the ABSA task.

From the survey that we conducted, we can see that ABSA problem is still an interesting and very broad matter to study because of the many methods that can be used, and effective methods are still needed to solve the ABSA problem. Many deep learning models have been used for ABSA, however, it is still very challenging to build an efficient model with high accuracy in ABSA tasks. From the review, we can say that graph and capsule models are still few to be developed and can be considered in the future to build models based on these two models.

We can see that the use of word embedding is also very influential on the level of accuracy. For example, in one model, the use of different word embedding will result in different levels of accuracy such as the use of Glove and BERT in one model. So, the choice of word embedding becomes very important. BERT here can be considered to be developed as a pre-trained embedding use. From the two datasets, we also found that every dataset has a different method to get better performance. It is still challenging to find the method that can be flexible to use in several datasets.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

- Akhtar, M. S., T. Garg, and A. Ekbal. 2020. Multi-task learning for aspect term extraction and aspect sentiment classification. *Neurocomputing* 398:247–56. Elsevier B.V. doi:[10.1016/j.neucom.2020.02.093](https://doi.org/10.1016/j.neucom.2020.02.093).
- Bahdanau, D., K. Cho, and Y. Bengio. 2015. Neural machine translation by jointly learning to align and translate. Conference paper at International Conference on Learning Representation (ICLR), 1–15.
- Binxuan, H., Y. Ou, and K. M. Carley. 2018. Aspect level sentiment classification with attention-over-attention neural networks. International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, Washington, DC, USA. Springer International Publishing, 197–206. doi: [10.1007/978-3-319-93372-6](https://doi.org/10.1007/978-3-319-93372-6).
- Cai, G., and H. Li. 2018. Joint attention LSTM network for aspect-level sentiment analysis. In China Conference on Information Retrieval. Guilin, China: Springer International Publishing, 147–57. doi: [10.1007/978-3-030-01012-6](https://doi.org/10.1007/978-3-030-01012-6).
- Carley, K. M. 2019. Syntax-aware aspect level sentiment classification with graph attention networks. *Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 5469–5477. November 2019.
- Chen, H.-Y., and H.-H. Chen. 2016. Implicit polarity and implicit aspect recognition in opinion mining. 20–25.
- Chen, J. 2019. Graph convolutional networks with structural attention model for aspect based sentiment analysis. 2019 International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary: IEEE, July, 1–7.
- Chen, Z., and T. Qian. 2019. Transfer capsule network for aspect level sentiment classification. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 547–56.
- Cho, Kyunghyun, Bahdanau, Dzmitry, Bougares, Fethi, Schwenk, Holger, Bengio, Yoshua. 2014. Learning phrase representations using RNN encoder–decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 1724–1734.
- Cui, Yiming, Chen, Zhipeng, Wei, Si, Wang, Shijin, Liu, Ting, Hu, Guoping, 2017 Attention-over-attention neural networks for reading comprehension. ACL 2017-55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers) July, 2017 Vancouver, Canada, 1, 593–602. doi: [10.18653/v1/P17-1055](https://doi.org/10.18653/v1/P17-1055).
- Da’u, A., and N. Salim. 2019. Aspect extraction on user textual reviews using multi-channel convolutional neural network. *PeerJ Computer Science* (5):1–16. doi:[10.7717/peerj-cs.191](https://doi.org/10.7717/peerj-cs.191).
- Ding, X., Liu, Bing, Yu, Philip. 2008. A holistic lexicon-based approach to opinion mining. Proceedings of the 2008 International Conference on Web Search and Data Mining, California, USA, 231–39.

- Du, C., H. Sun, and J. Wang, Qi, Q., Liao, J., Xu, T., Liu, M. 2019. Capsule network with interactive attention for aspect-level sentiment classification. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Hong Kong, China, Association for Computational Linguistics, 5488–5498 <https://aclanthology.org/D19-1551> doi:10.18653/v1/D19-1551.
- Gan, C., Wang, L., Zhang, Z., Wang, Z. 2019. Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis. *Knowledge-Based Systems* 188: 104827. Elsevier B.V. doi:10.1016/j.knosys.2019.06.035.
- Ganu, G., N. Elhadad, and A. Marian. 2009. Beyond the stars: improving rating predictions using review text content. *WebDB* 9:1–6.
- Gori, M., G. Monfardini, and F. Scarselli. 2005. A new model for learning in graph domains. In Proceedings. 2005 IEEE International Joint Conference on Neural Networks, Montreal, QC, Canada: IEEE, 729–34 vol. 2. doi: 10.1109/IJCNN.2005.1555942.
- Gu, S., Zhang, L., Huo, Y., Song, Y., et al. 2018. A position-aware bidirectional attention network for aspect-level sentiment analysis Proceedings of the 27th International Conference on Computational Linguistics August, 2018 Santa Fe, New Mexico, USA. 774–784 <https://aclanthology.org/C18-1066>.
- He, R., Lee, W., Ng, H., Dahlmeier, D. 2017. An unsupervised neural attention model for aspect extraction. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics , Vol 1, Vancouver, Canada, (Association for Computational Linguistics), 388–397 <https://aclanthology.org/P17-1036> doi:10.18653/v1/P17-1036.
- He, R., Lee, W., Ng, H., Dahlmeier, D. 2018. Effective attention modeling for aspect-level sentiment classification Proceedings of the 27th International Conference on Computational Linguistics August, 2018 Santa Fe, New Mexico, USA. 1121–1131.
- Hochreiter, S., and J. Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 1780 (8):1735–80. doi:10.1162/neco.1997.9.8.1735.
- Hou, X., Huang, J., Wang, G., Qi, P., He, X. 2019. Selective attention based graph convolutional networks for aspect-level sentiment classification Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15) June, 2021 (Association for Computational Linguistics) 83–93 Mexico City, Mexico. <https://aclanthology.org/2021.textgraphs-1.8> doi:10.18653/v1/2021.textgraphs-1.8.
- Hou, X., J. Huang, and G. Wang, Qi, P., He, X., Zhou, B. 2019. Selective attention based graph convolutional networks for aspect-level sentiment classification Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15) June, 2021 (Association for Computational Linguistics)83–93, Mexico City, Mexico. <https://aclanthology.org/2021.textgraphs-1.8> doi:10.18653/v1/2021.textgraphs-1.8
- Hu, M., B. Liu, and S. M. Street. 2004. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Seattle WA, USA (New York, United States: Association for Computing Machinery), 168–177.
- Huang, B., and K. M. Carley. 2020. Syntax-aware aspect level sentiment classification with graph attention networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing November 3-7, 2019 Hong Kong, China, 5469–5477. doi: 10.18653/v1/d19-1549.
- Jabreel, M., F. Hassan, and A. Moreno. 2017. Target-dependent sentiment analysis of tweets using bidirectional gated recurrent neural networks. In *Advance in hybridization of intelligent methods* (WEBIST), 39–55.

- Jabreel, M., F. Hassan, and A. Moreno. 2018. Target-dependent sentiment analysis of tweets using bidirectional gated recurrent neural networks. *Smart Innovation, Systems and Technologies* 85:39–55. doi:10.1007/978-3-319-66790-4_3.
- Jeon, S. W., Lee, H. J., Lee, H., Cho, S. 2019. Graph based aspect extraction and rating classification of customer review data. In International Conference on Database Systems for Advanced Applications Chiang Mai, Thailand 11448 , 186–99 doi:10.1007/978-3-030-18590-9.
- Jiang, L., Yu, M., Zhou, M., Liu, X., Zhao, T. 2011. Target-dependent Twitter sentiment classification. In Proceeding of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies June, 2011 (Association for Computational Linguistics) Portland, Oregon, USA, 151–160.
- Jiang, M., W. Zhang, M. Zhang, J. Wu, T. Wen, et al. 2019. An LSTM-CNN attention approach for aspect-level sentiment classification. *Journal of Computational Methods in Sciences and Engineering* 19 (4):859–868. doi:10.3233/JCM-190022.
- Kaur, R. 2021. Naive Bayes: A text classifier based on machine learning. *International Journal of Research Publication and Reviews* 2:260–66.
- Kenton, M. C., L. Kristina, and J. Devlin. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding, Proceedings of NAACL-HLT 2019 June 2 - June 7 4171–4186, Minneapolis, Minnesota.
- Kessler, J. S. 2009. Targeting sentiment expressions through supervised ranking of linguistic configurations. In International AAAI Conference on Web and Social Media San Jose, California, USA 3 , 90–97 <https://ojs.aaai.org/index.php/ICWSM/article/view/13948>.
- Kim, Y. 2014. Convolutional neural networks for sentence classification, Conference on Empirical Methods in Natural Language Processing (EMNLP) October, 2014 (Association for Computational Linguistics)1746–1751 Doha, Qatar.
- Le, H. T., C. Cerisara, and A. Denis. 2017. Do convolutional networks need to be deep for text classification ? AAAI Workshop on Affective Content Analysis Feb 2018 hal-01690601, New Orleans, United States. <https://hal.archives-ouvertes.fr/hal-01690601>
- Lee, Y., M. Chung, S. Cho, J. Choi. 2019. Extraction of product evaluation factors with a convolutional neural network and transfer learning. *Neural Processing Letters* 50 (1):149–64. Springer US. doi:10.1007/s11063-018-9964-8.
- Li, L., Y. Liu, and A. Zhou. 2018. Hierarchical attention based position-aware network for aspect-level sentiment analysis. Conference of Computational Natural Language Learning (CoNLL), Brussels, Belgium, Association for Computational Linguistics, 181–189. <https://aclanthology.org/K18-1018> doi:10.18653/v1/K18-1018.
- Liu, Q., Gao, Z., Liu, B., Zhang, Y. July 2015. Automated rule selection for aspect extraction in opinion mining. In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015) July 25 - 31, 2015 (AAAI Press) Buenos Aires Argentina, 1291–1297.
- Liu, Q., Z. Gao, Liu, B., Zhang, Y. 2016. Automated rule selection for opinion target extraction. *Knowledge-Based Systems* Elsevier B.V. doi:10.1016/j.knosys.2016.04.010.
- Luong, M. T., H. Pham, and C. D. Manning. 2015. Effective approaches to attention-based neural machine translation. Conference on Empirical Methods in Natural Language Processing September, 2015 (Association for Computational Linguistics) Lisbon, Portugal, 1412–1421. doi: 10.18653/v1/d15-1166.
- Ma, D., Li, S., Zhang, X., Wang, H., et al. 2017. Interactive attention networks for aspect-level sentiment classification, International Joint Conference on Artificial Intelligence, August 19 - 25, 2017, 4068–4074, Melbourne Australia.

- R, M. C. M. V. K, Manek, A. S., and P. D. Shenoy. 2016. Aspect term extraction for sentiment analysis in large movie reviews using Gini index feature selection method and SVM classifier. *World Wide Web*. doi:10.1007/s11280-015-0381-x.
- Mikolov, T., Chen, K., Corrado, G., Dean, J. 2013. Efficient estimation of word representations in vector space. 1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings May 2 - 4, 2016 Caribe Hilton, San Juan, Puerto Rico, 1–12.
- Mukherjee, A., and B. Liu. 2012. Aspect extraction through semi-supervised modeling. In Proceedings of the 50 Annual Meeting of the Association for Computational Linguistics July, 2012 (Association for Computational Linguistics) Jeju Island, Korea, 339–348. <https://aclanthology.org/P12-1036>.
- Pang, B., and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2 (1–2):1–135. doi:10.1561/15000000001.
- Pang, B., Lee, L., Vaithyanathan, S. 2002. Thumbs up? Sentiment classification using machine learning techniques. In Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia (Association for Computational Linguistics), 79–86.
- Park, H.-J., M. Song, and K.-S. Shin. 2020. Deep learning models and datasets for aspect term sentiment classification: Implementing holistic recurrent attention on target-dependent memories. *Knowledge-Based Systems* 187:104825. Elsevier B.V. doi:10.1016/j.knsys.2019.06.033.
- Penghua, Z., and Z. Dingyi. 2019. Bidirectional-GRU based on attention mechanism for aspect-level sentiment analysis. In Proceedings of the 2019 11th International Conference on Machine Learning and Computing, February 22 - 24, 2019 (Machinery, New York, NY, United States: Association for Computing) Zhuhai China, 86–90.
- Pennington, J., R. Socher, and C. D. Manning. 2014. GloVe: Global vectors for word representation, Conference on Empirical Methods in Natural Language Processing (EMNLP) October, 20 2014 (Association for Computational Linguistics) Doha, Qatar. 1532–1543.
- Pontiki, M., D. Galanis, H. Papageorgiou *et al.* 2015. SemEval-2015 task 12: Aspect based sentiment analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) Denver, Colorado (Association for Computational Linguistics), 486–495 <https://aclanthology.org/S15-2082> doi:10.18653/v1/S15-2082.
- Pontiki, M., Galanis, D., Papageorgiou, H *et al* 2016. SemEval-2016 task 5: Aspect based sentiment analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2016) June, 2016 (Association for Computational Linguistics) San Diego, California, 19–30 <https://aclanthology.org/S16-1002> doi:10.18653/v1/S16-1002.
- Pontiki, M., J. Pavlopoulos *et al.* 2014 SemEval-2014 Task 4: Aspect Based Sentiment Analysis Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) August, 2014 (Association for Computational Linguistics) Dublin, Ireland. In , 27–35. <https://aclanthology.org/S14-2004> doi:10.3115/v1/S14-2004.
- Poria, S., Cambria, E., Ku, L. W., Gui, C., Gelbukh, A., et al. 2014. A rule-based approach to aspect extraction from product reviews. In Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP) Dublin, 28–37.
- Poria, S., E. Cambria, and A. Gelbukh. 2016. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems* 108:42–49. Elsevier B.V. doi:10.1016/j.knsys.2016.06.009.
- Rana, T. A., and Y. Cheah. 2017. A two-fold rule-based model for aspect extraction Expert Systems with Applications: An International Journal . 89 C :273–85. Elsevier Ltd. doi:10.1016/j.eswa.2017.07.047.

- Ren, F., Feng, L., Xiao, D., Cai, M., Cheng, S. 2020. DNet: A lightweight and efficient model for aspect based sentiment analysis. *Expert Systems with Applications* 151:113393. Elsevier Ltd. doi:10.1016/j.eswa.2020.113393.
- Senarath, Y., N. Jihan, and S. Ranathunga. 2019. A hybrid approach for aspect extraction from customer reviews International Journal on Advances in ICT for Emerging Regions (ICTer) . 12 (September):1–8.
- Shah, K., H. Patel, D. Sanghvi, M. Shah. 2020. A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augmented Human Research* 5(1). Springer Singapore. doi:10.1007/s41133-020-00032-0.
- Shu, L., H. Xu, and B. Liu. 2017. Controlled CNN-based sequence labeling for aspect extraction.
- Shu, L., H. Xu, and B. Liu. no date. Lifelong learning CRF for supervised aspect extraction.
- Sperduti, A., and A. Starita. 1997. Supervised neural networks for the classification of structures. *IEEE Transactions on Neural Networks* 8 (3):714–35. doi:10.1109/72.572108.
- Sun, Z., Bing, L., Yang, W. 2017. Recurrent attention network on memory for aspect sentiment analysis. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 452–461.
- Syahrul, M., and M. Dwi. 2017. Aspect-based sentiment analysis to review products using Naïve Bayes. International Conference on Mathematics: Pure, Applied and Computation, Surabaya-Indonesia, 020060. doi: 10.1063/1.4994463.
- Tan, C., Dong, L., Wei, F., Tang, D., Zhou, M., Xu, K. 2014. Adaptive recursive neural network for target-dependent Twitter sentiment classification. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland, Vol. 2: Short P, 49–54.
- Tang, D., B. Qin, and T. Liu. 2014. Aspect level sentiment classification with deep memory network, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, November, 2 214–224. <https://aclanthology.org/D16-1021> . doi:10.18653/v1/D16-1021 .
- Tang, D., Qin, B., Feng, X., Liu, T., et al. 2016. Effective LSTMs for target-dependent sentiment classification. COLING 2016-26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers, Osaka, Japan (The COLING 2016 Organizing Committee), 3298–3307.
- Tang, D., Qin, B., Feng, X., Liu, T.W, et al. 2015. Target-dependent sentiment classification with long short term memory. *Arxiv Preprint*.
- Toqir, A. R., and Y. Cheah. 2016. Aspect extraction in sentiment analysis: Comparative analysis and survey. *Artificial Intelligence Review*. 46 (4):459–83. Springer Netherlands. doi:10.1007/s10462-016-9472-z.
- Vo, D., and Y. Zhang. 2015. Target-dependent Twitter sentiment classification with rich automatic features. International Joint Conference on Artificial Intelligence Buenos Aires, Argentina, Ijcai, 1347–53.
- Wang, Y., A. Sun, and J. Han. 2018. Sentiment analysis by capsules. In Proceedings of the 2018 Word Wide Web Conference, Lyon, France, 1165–74.
- Wang, Y., Huang, M., Zhu, X., Zhao, L., et al. 2016. Attention-based LSTM for aspect-level sentiment classification. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing Austin, Texas (Association for Computational Linguistics), 606–615.
- Wang, Y., Sun, A., Huang, M., Zhu, X., et al. 2019. Aspect-level sentiment analysis using AS-capsules. International World Wide Web Conference Committee May 13 - 17, 2019 (Association for Computing) San Francisco CA USA, 2033–44.
- Weston, J., S. Chopra, and A. Bordes. 2015. Memory networks. *arXiv Preprint* 1–15.

- Wu, C., , , Xiong, Q., Gao, M., Li, Q., Yu, Y., Wang, K. 2020. A relative position attention network for aspect-based sentiment analysis. *Knowledge and Information Systems* 62 (9):3615–40. Springer London. doi:10.1007/s10115-020-01512-w.
- Xiao, L., Hu, X., Chen, Y., Xue, Y., Gu, D., Chen, B., Zhang, T. 2020. Targeted sentiment classification based on attentional encoding and graph convolutional networks. *Applied Sciences (Switzerland)* 10(3). doi: 10.3390/app10030957.
- Xu, H., Liu, B., Shu, L., Yu, P.S (2018) 'Double embeddings and CNN-based sequence labeling for aspect extraction', *ACL 2018-56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)* July, 2018 Melbourne, Australia, 2, pp. 592–598 doi:10.18653/v1/P18-2094.
- Xu, Q., L. Zhu, T. Dai, C. Yan 2020. Aspect-based sentiment classification with multi-attention network. *Neurocomputing* 388:135–43. Elsevier B.V. doi:10.1016/j.neucom.2020.01.024.
- Xue, W. 2017. MTNA: A neural multi-task model for aspect category classification and aspect term extraction on restaurant reviews. Proceedings of the The 8th International Joint Conference on Natural Language Processing, Taipei, Taiwan, 151–156 <https://aclanthology.org/I17-2026>.
- Xue, W., and T. Li. 2018. Aspect based sentiment analysis with gated convolutional networks. *Arxiv Preprint*.
- Yang, M., Qu, Q., Chen, X., Guo, C., Shen, Y., Lei, K. 2018. Feature-enhanced attention network for target-dependent sentiment classification *Neurocomputing* . 307:91–97. Elsevier B.V. doi:10.1016/j.neucom.2018.04.042.
- Yang, Y. 1999. An evaluation of statistical approaches to text categorization. *Information Retrieval* 90 (1–2):69–90. doi:10.1023/A:1009982220290.
- Zhang, C., Q. Li, and D. Song. 2019. Aspect-based sentiment classification with aspect-specific graph convolutional networks. *arXiv Limitation* 2:4568–78.
- Zhang, D., Zhu, Z., Lu, Q., Pei, H., Wu, W., Guo, Q. 2020. Multiple interactive attention networks for aspect-based sentiment classification. *Applied Sciences (Switzerland)* 10 (6):1–15. doi:10.3390/app10062052.
- Zhang, Q., and R. Lu. 2019. A multi-attention network for aspect-level sentiment analysis.
- Zhao, P., L. Hou, and O. Wu. 2019. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *arXiv Preprint, arXiv: 1906*.
- Zhao, W. X., Jiang, J., Yan, H., Li, X. 2010. Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. Proceeding of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, MA (Association for Computational Linguistics), (October), 56–65 <https://aclanthology.org/D10-1006>.