Knowledge-Based Systems 253 (2022) 109506

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Double embedding and bidirectional sentiment dependence detector for aspect sentiment triplet extraction

Dawei Dai [a], Tao Chen [a], Shuyin Xia [a,*], Guoyin Wang [a], Zizhong Chen [b]

[a] The College of Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Telecommunications and Posts, China
[b] The department of Computer Science and Engineering, University of California, Riverside, United States of America

ARTICLE INFO

Article history:
Received 19 March 2022
Received in revised form 19 July 2022
Accepted 20 July 2022
Available online 2 August 2022

Keywords:
Aspect sentiment triplet extraction
Character-level embedding
Triplet extraction
Double embedding
Bidirectional sentiment-dependence detector

ABSTRACT

Aspect sentiment triplet extraction (ASTE) is a popular subtask related to aspect-based sentiment analysis (ABSA). It extracts aspects and their associated opinion expressions and sentiment polarities from comment sentences. Previous studies have proposed a multitask learning framework that jointly extracts aspect and opinion terms and treats the sentiment analysis task as a table-filling problem. Although the multitask learning framework solves the problem of identifying overlapping opinion triples, the entire model cannot explicitly simulate interactions between aspects and opinions. Therefore, we propose a sentiment-dependence detector based on a dual-table structure that starts from two directions, aspect-to-opinion and opinion-to-aspect, to generate two sentiment-dependence tables dominated by two types of information. These complementary directions allow our framework to explicitly consider interactions between aspects and opinions and better identify triples. Moreover, we use a double-embedding mechanism—character-level and word-vector embeddings—in the model for triplet extraction that enables it to represent contexts at different granularity levels and explore high-level semantic features. To the best of our knowledge, this study presents the first bidirectional long short-term memory (BiLSTM) model based on double embedding used to perform ASTE tasks. Finally, our analysis shows that our proposed bidirectional sentiment-dependence detector and double-embedding BiLSTM model achieve more significant results than the baseline model for triples with multiple identical aspects or opinions.

1. Introduction

Sentiment recognition and polarity detection are basic tasks in affective computing and sentiment analysis [1,2]. The aspect-based sentiment analysis (ABSA) task [3–5] is a fine-grained sentiment analysis task [6] that can identify aspects and their corresponding sentiment from sentences. Various subtasks have been developed to solve ABSA tasks. However, each subtask has been designed with the intent of extracting aspects. For example, aspect term extraction (ATE) [7–9] extracts target entities and phrases in a sentence and primarily focuses on extracting the target of a given aspect. Opinion term extraction (OTE) [10–12] aims to extract words from a given sentence that express subjective opinions about a target entity with sentiment polarity. Aspect-term sentiment analysis (ATSA) [13] and target-oriented opinion word extraction (TOWE) [14–16] are primarily used to extract the feelings and opinions about corresponding aspects from sentences. To perform these subtasks, researchers have designed various neural network models to handle fine-grained sentiment-classification tasks.

In recent years, most researchers have developed joint learning algorithms that can complete each sentiment analysis subtask simultaneously, and the task based on aspect sentiment triplet extraction (ASTE) proposed by Peng et al. [17] has attracted significant attention. The ASTE task [17–19] is a relatively new subtask. In this type of task, not only should the target entity and corresponding opinion be extracted from the sentence, but the sentiment dependence between the target and opinion should also be obtained to form a triad to be extracted. For example, in Fig. 1, the aspect target is a restaurant, the opinions are cute and not upscale, and the corresponding polarities of sentiment are positive and negative. Peng et al. [17] proposed a two-stage model. In the first stage, a long short-term memory (LSTM) model was used to jointly extract the aspect sentiment and opinion pairs. In the second stage, a classifier was used to pair the target with a sentiment label with corresponding opinion words to obtain all valid triples. They explored the task using a unified aspect sentiment labeling. However, this approach did not handle the overlap of certain elements in different triples well. We call this the element overlap problem.

* Corresponding author.
E-mail addresses: dw_dai@163.com (D. Dai), 951405993@qq.com (T. Chen), xiasy@cqupt.edu.cn (S. Xia), wanggy@cqupt.edu.cn (G. Wang), zizhongchen@gmail.com (Z. Chen).

https://doi.org/10.1016/j.knosys.2022.109506
0950-7051/

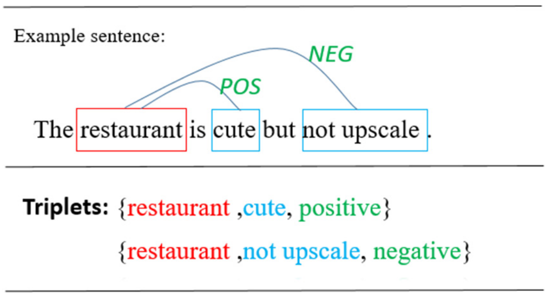


Fig. 1. An example of ASTE.

Recently, the multitask learning framework for ASTE proposed by Zhang et al. [20] has become the most prominent method for this task. To avoid the aforementioned problems, they adopted a multistructure. In the first stage, they used two heads to extract aspects and opinions separately. In the second stage, a word-level dependency parser was used as a third head to transform the sentiment analysis task into a table-filling problem. Although this solution solves the previous problem, it does not explicitly model the interaction between aspects and opinions. In recent years, most models developed for the ASTE task have tended to use only one feature for training. In a supervised depth model, the use of handmade features [21,22] and vocabulary can often achieve better performance and results [23]. To date, most deep learning models have adopted pretrained universal embedding, such as GloVe [24], universal comment embedding [25], and fine-grained domain embedding [26]. However, automatic feature learning is desirable for models, and an important problem involves studying how to make models more competitive without using manual features.

In this study, we propose a model based on double embedding and a bidirectional sentiment-dependence detector, DE-OTE-BISDD. In this model, we adopt a double-embedding mechanism: character-level and word-vector embeddings. As the embedded layer is the first layer of the model, it encodes all information about the word, and the quality of the embedded layer is crucial for subsequent layers to extract the semantic information of the entire sentence. In this manner, we can obtain a representation of the sentence at different granularity levels and then fuse it into high-level semantic features to extract aspects and opinions. We then use a bidirectional sentiment-dependence detector to predict the sentiment connection between aspects and opinions. This yields two dependency tables dominated by different information from the perspectives of aspect-to-opinion and opinion-to-aspect, thereby yielding effective triples according to the decoding strategies we formulated. Fig. 2 shows a table of sentiment dependencies obtained from both directions.

In our observations, the three elements of the triad were highly correlated. When predicting the sentiment dependence of aspects and opinions, the model can fully capture interactions between aspect and opinion words. We consider the aspects and opinions of individual words, as well as the aspects and opinions in span-level interactions; our model independently predicts sentiment relationships between all possible aspect–opinion pairs. We conducted extensive experiments on four SemEval datasets of ABSA and compared the results with those of state-of-the-art methods to validate our proposed method. We also conducted ablation and case studies to demonstrate how the proposed approach can improve the performance of the model.

The contributions of this study are as follows.

We propose a bidirectional LSTM (BiLSTM) model based on double embedding, which enables the model to obtain semantic information at different granularity levels to enhance its understanding of the entire sentence. Our proposed approach can be easily combined with different models to significantly improve their performance.

We propose a bidirectional sentiment-dependence detector that predicts in two directions: aspect-to-opinion and opinion-to-aspect. The triplet extraction results can be significantly improved using these two complementary directions. We also applied the detector to other models, which significantly improved their performance.

We evaluated our model on four SemEval datasets, and the results showed that our model outperformed several strong baselines in the ASTE task.

2. Related work

Contemporary sentiment analysis has evolved into more detailed sentiment analysis and perception [27]. Fine-grained sentiment analysis [28–30] is an important task in natural language processing. This is generally studied as a fine-grained opinion-mining task. In recent years, researchers have begun to notice the dependence and correlation between various subtasks based on aspect-level sentiment classification (ASC); therefore, most research now concerns joint tasks. This section briefly reviews related work on the ABSA task and its derived subtasks and explores the latest trends in explainable artificial intelligence for sentiment analysis.

2.1. Aspect-based sentiment analysis and its related subtasks

Among neural network models, the LSTM model is a sequential structure that can preserve the long-term semantic dependencies of text and flexibly capture the interaction between target words and the context. The LSTM model cannot focus on capturing the local information of sentences; however, for the ASC task, local sentence information is important for analyzing the specific aspect sentiment. To effectively solve this problem, Wang et al. [31] proposed an attention mechanism that focuses on different parts of sentences and uses aspect words as input to the LSTM model for ASC. Zhou et al. [29] proposed a hierarchical transfer model of location awareness, which modeled location information at the embedding, word, segment, and classifier levels and transferred the location information of its hierarchical modeling to ASC [32–34] to improve its performance. Liang et al. [35] proposed a method to construct a graph convolutional network by integrating affective knowledge from SenticNet to fully consider the dependency between context and aspect words, as well as sentiment information between opinion and aspect words.

In recent years, to improve the performance of sentiment classification, most models using LSTM networks have captured the important context information of the target through various attention mechanisms; however, the training is extremely complex and time-consuming. LSTM is difficult to parallelize owing to its limitations, particularly when it requires a large amount of memory to backpropagate over a sequence, which poses difficulties in the long-term memory mode. Therefore, Song et al. [36] proposed an attentional encoder network based on target sentiment classification that can mine rich semantic information and provide a model between the context and target without considering distances between words. Wei et al. [37] proposed a model based on a convolutional neural network (CNN) and gating mechanism. The CNN is not time-dependent, and the gating unit can work independently, facilitating parallelization during training. Valdivia et al. [38] used different sentiment analysis methods to extract sentiments from different corpora and filtered

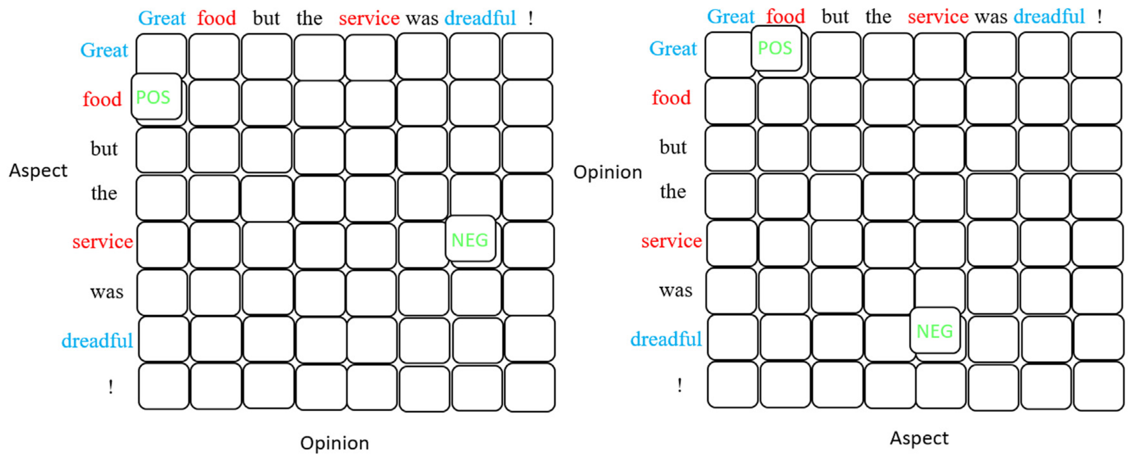


Fig. 2. An example of table filling for aspect joint opinion and opinion joint aspect.

them via consensus detection of neutral comments to improve the classification effect.

With the intensifying research on ABSA tasks, researchers have proposed a variety of related subtasks. For example, for the TOWE task, Fan et al. [14] designed a method for encoding target word information into an LSTM to integrate the left and right context and global context information of the target opinion to extract opinion targets and words in pairs. In the ATE task, Xu et al. [26] proposed a CNN model based on double embedding, which uses universal and domain-specific embeddings to extract aspects without additional supervision. In the OTE task, Wu et al. [39] transferred rich knowledge of the sentiment classification of comments obtained from online websites to a low-resource TOWE task to obtain potential opinions through this effective transformation method.

2.2. Aspect sentiment triplet extraction

Currently, many studies focus on developing a unified framework that can effectively identify specific aspects, opinions, and sentiments in text. Therefore, capturing and using the relationship between these factors is challenging. To solve this problem, Peng et al. [17] proposed a new subtask called ASTE and a two-stage model. In the first stage, the model identifies sentiment labels with attached aspects and corresponding opinion terms. In the second stage, it pairs sentiment labels with attached aspects and opinion labels to form triples. Li et al. [40] adopted a unified marking scheme that uses two stacked recursive neural networks. The upper network predicts the results of the target sentiment analysis, whereas the lower network assists in predicting the target boundary. Considering the high correlation between elements in triples, Xu et al. [41] constructed an end-to-end model and used a position annotation scheme to jointly extract triples. Zhang et al. [20] proposed a multitask learning framework to jointly extract aspect and opinion terms and finally transform the sentiment analysis task into a table-filling problem. Chen et al. [42] adopted a machine-reading comprehension method to perform the ABSA task. They proposed a two-way machine-reading comprehension framework and designed three types of queries—nonrestrictive extraction, restrictive extraction, and sentiment classification queries—to establish the association between different subtasks. Triples were identified in this two-way, multiround question-and-answer process.

2.3. Explainable artificial intelligence for sentiment analysis

Explainable artificial intelligence is an emerging field of deep learning. Trueman et al. [43] proposed a convolutional stacked BiLSTM with a multiplicative attention mechanism for aspect category and affective polarity detection. Cambria et al. [44] proposed a commonsense-based neurosymbolic framework. They employ unsupervised and reproducible subsymbolic techniques to build trusted symbolic representations, translating natural language into protolanguage, and thus extracting polarity from text in an interpretable manner. Lieto et al. [45] proposed an explainable emotion attribution and recommendation system that relies on a recently introduced common sense reasoning framework, $\mathbf{T}^{CL}$ logic. Starting from the formalization of emotion ontology based on the Plutchik model, $\mathbf{T}^{CL}$ logic has been used to automatically generate new common-sense semantic representations of complex sentiments. Perikos et al. [46] proposed an interpretable hidden Markov model approach, introducing explainable models that indicate the sentiment parts of sentences and evolution of the overall sentiment from beginning to end.

3. Our proposed framework

3.1. Task definition

Given a sentence $X = \{x_1, x_2, \dots, x_N\}$, the goal is to represent a sentence with n tokens. The ASTE task identifies the set of all triples in a sentence, i.e., (aspect, opinion, sentiment), where sentiment ∈ {Positive, Negative, Neutral}. We treat the aspect and opinion extraction subtasks as sequence labeling problems and adopt the beginning, inside, outside (BIO) annotation scheme [47,48]. For sentiment analysis, we mainly follow the work of Miwa and Sasaki [49] to formulate the task as a table-filling problem, predict all sets of triples jointly in two directions, and finally combine all triples through a decoding strategy.

3.2. Proposed framework

As shown in Fig. 3, the proposed framework comprises four main modules: sentence coding, semantic fusion, multitask learning, and triplet decoding. The sentence coding stage consists of two embedding layers with different trainable granularity levels. For a given sentence, we first pass it through word-level and word-direction embeddings to obtain two different granularity levels of sentence representation. Then, the two high-level semantic features obtained are fused into the BiLSTM, and the

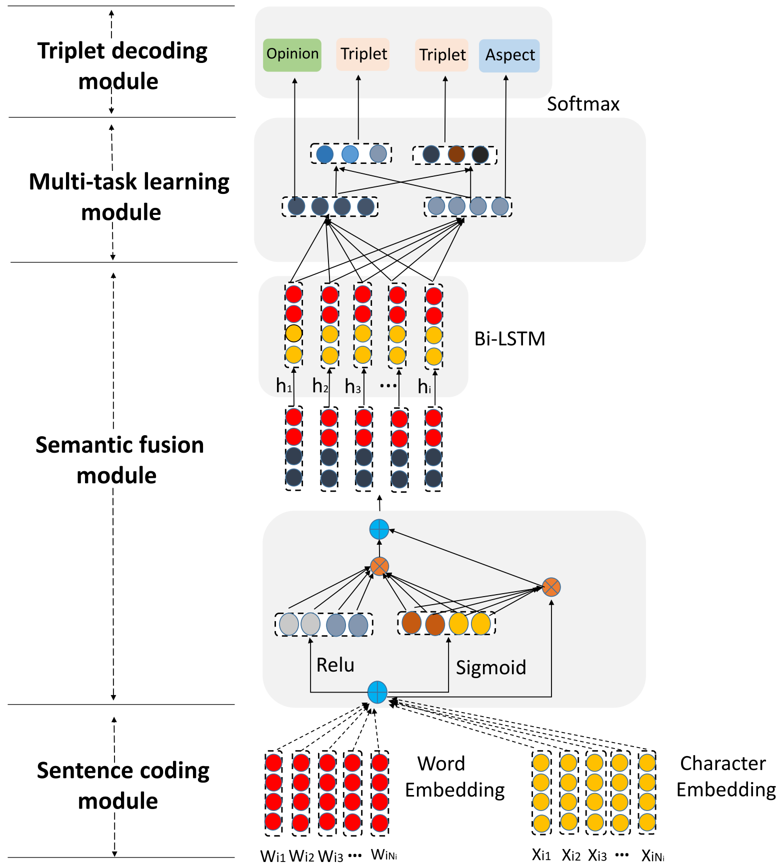


Fig. 3. Architecture of the proposed framework.

abstract representation of the aspect and opinion is learned. The abstract representation is input into the multitask learning module. The multitask learning module consists of an aspect marker layer, opinion marker layer, and bidirectional sentiment-dependence detector. After word-level sentiment dependence is detected from the aspect-to-opinion and opinion-to-aspect perspectives, all triples are generated by the triplet decoding module according to heuristic rules. Finally, all triples are merged according to the formulated decoding strategy.

3.3. Sentence coding module

We used two encoding schemes to obtain a contextualized representation of each word in a sentence: character- and word-level encoding layers.

Character-level encoding layer: This layer uses a CNN to obtain the character-level embedding of each word. We let $\{x_1, x_2, \ldots, x_N\}$ represent the entire input sentence and then map each character to a higher-dimensional vector space to obtain a vector representation of each character. These character vectors are used as the input for the entire CNN. The output of the CNN is maximally pooled by width to obtain a vector of a fixed size for each word.

Word-level encoding layer: This layer maps all words in the sentence into a high-dimensional vector space and then uses the pretrained 300-dimensional GloVe [24] word vector to obtain a vector representation $\{e_1, e_2, \ldots, e_N\}$.

3.4. Semantic fusion module

We splice the vector embedded with characters and words and then pass it to the two-layer highway network [50]. Subsequently, the output of the two-layer highway network is transmitted to the BiLSTM network [51]. Thus, we can obtain semantic representations of the context that fuse different granularity levels using the BiLSTM:

$$h_i = \left[\overrightarrow{h}_i; \overleftarrow{h}_i\right], \tag{1}$$

where $\overrightarrow{h}_i$ and $\overleftarrow{h}_i$ represent the forward and backward LSTM hidden state-vector representations, respectively.

Next, we extract specific aspect and opinion features from the contextual semantic representations that fuse different granularity levels by dimensionally reducing the linear layers and nonlinear functions. In essence, we do not use semantic information for multitask learning to remove possible redundant information from the semantic information. We intend to remove irrelevant features and avoid the risk of overfitting for the multitasking learning module. The calculation formulas are as follows:

$$\mathbf{z}_i^{(\mathrm{ap})} = f\left(W_r{}^{(ap)} h_i + b_r^{(ap)}\right), \tag{2}$$

$$\mathrm{z}_i^{(\mathrm{op})} = f\left(W_r^{(\mathrm{op})} h_i + b_r^{(\circ p)}\right), \tag{3}$$

where $z_i^{(ap)} \in R^{\mathrm{d_r}}$ and $z_i^{(op)} \in R^{\mathrm{d_r}}$ express aspects and views. $W_\mathrm{r}^{(op)}$, $W_\mathrm{r}^{(ap)}$ and $\mathrm{b}_\mathrm{r}^{(op)}$, $\mathrm{b}_\mathrm{r}^{(ap)}$ are the weight and bias terms of the model, respectively. $f(\cdot)$ is a nonlinear function; we used ReLU as the activation function. From Eqs. (2) and (3), we obtain another set of representations: $z_i^{(ap)'} \in R^{\mathrm{d_r}}$ and $z_i^{(op)'} \in R^{d_r}$ for sentiment analysis.

3.5. Multitask learning module

The multitask learning module consists of two main parts: the aspect and opinion marker layers and a bidirectional sentiment-dependence detector.

Aspect and opinion marker layers: In the aspect- and point-marking section, we use the same settings as Zhang et al. [20]

Tags mark each word in the sentence according to the {B,I,O} scheme. In particular, we use $z_i^{(ap)}$ to calculate the probability $P_i^{(ap)}$ of each word being marked as an aspect. $z_i^{(op)}$ is used to calculate the probability $P_i^{(op)}$ of each word being marked as an opinion. $W_{\mathrm{t}}^{(op)}$, $W_{\mathrm{t}}^{(ap)}$ and $\mathrm{b}_{\mathrm{t}}^{(op)}$, $\mathrm{b}_{\mathrm{t}}^{(ap)}$ are the learnable weights and biases, respectively. The calculation formulas are as follows:

$$P_i^{(ap)} = \text{softmax}\left(\mathrm{w}_t^{(ap)} z_i^{(ap)} + b_t^{(ap)}\right), \tag{4}$$

$$P_i^{(op)} = \text{softmax}\left(\mathrm{w}_t^{(op)} z_i^{(op)} + b_t^{(op)}\right). \tag{5}$$

Bidirectional sentiment-dependence detector: In the DE-OTE-BISDD, we use a biaffine scorer to detect sentiment dependence between aspect and opinion pairs. A biaffine scorer was proven to be expressive in syntactic dependency parsing [52]. It captures the interaction between word pairs and significantly reduces learning redundancy. We designed a double-table structure that obtains the sentiment dependence relation table between an aspect and opinion pair from the aspect-to-opinion and opinion-to-aspect directions. We define these two processes as the $A \rightarrow O$ and $O \rightarrow A$ directions, respectively.

We then define the set of sentiment dependency types as {neutral (NEU), negative (NEG), positive (POS), absence of sentiment dependence (NO-DEP)} to obtain various dependencies.

In particular, we first analyze the sentiment dependence process from the perspective of the $A \rightarrow O$ direction design.

$A \longrightarrow O$, aspect-to-opinion sentiment dependence analysis: We match all the words of the sentence (including self-pairings) to form a two-dimensional table. The horizontal direction represents aspects, and the vertical direction represents opinions. When the biaffine scorer analyzes the sentiment dependence between an aspect and opinion, it fills in the corresponding sentiment polarity where the last word of one aspect links to the last word of an opinion. We use Eq. (6) to calculate the score between each word pair and then use the softmax function to determine the probabilities of all word pair scores:

$$\tilde{\mathbf{S}}_{i,j,k}^{A->O} = \left[w^{(k)} z_i^{(ap)'} + b^{(k)}\right]^T z_j^{(op)'}, \tag{6}$$

$$\mathbf{s}_{i,j,k}^{A->O} = \text{softmax}\left(\tilde{\mathbf{s}}_{i,j,k}^{A->O}\right), \tag{7}$$

where $\tilde{\mathbf{S}}_{i,j,k}^{A->O}$ represents the score of the K-th dependency type of the word pair (w_i, w_j) in the direction $A \longrightarrow O$. w^k and b^k are the weights and biases, respectively, that produce the K-th score. $S_{i,j,k}^{A->O}$ denotes the probability of the score for the word pair (w_i, w_j).

$O \longrightarrow A$, opinion-to-aspect sentiment dependence analysis: We match all the words of the sentence (including self-pairings) to form a two-dimensional table. The horizontal direction represents opinions, and the vertical direction represents aspects. When the biaffine scorer analyzes the sentiment dependence between an opinion and aspect, it fills in the corresponding sentiment polarity where the last word of one opinion links to the last word of an aspect. We use Eq. (8) to calculate the score between each word pair and then use the softmax function to determine the probabilities of all word pair scores:

$$\tilde{\mathbf{s}}_{i,j,k}^{O->A} = \left[w^{(k)} z_i^{(op)'} + b^{(k)}\right]^T z_j^{(ap)'}, \tag{8}$$

$$\mathbf{S}_{i,j,k}^{O->A} = \left[\text{softmax}\left(\tilde{\mathbf{S}}_{i,j,k}^{O->A}\right)\right]^T, \tag{9}$$

where $\tilde{S}_{i,j,k}^{O->A}$ represents the score of the K-th dependency type of the word pair (w_i, w_j) in the direction $O \rightarrow A$. w^k and b^k are the weights and biases, respectively, that produce the K-th score. $S_{i,j,k}^{O->A}$ is the probability of the score for the word pair (w_i, w_j).

Finally, we combine the word pairs obtained from the bidirectional sentiment-dependence detector with the span of the aspect and opinion extracted from the aspect and opinion markers to obtain the attachment span level.

3.6. Triplet decoding module

After obtaining the aspect, opinion, and bidirectional sentiment-dependence detection results, we use heuristic rules to decode the triplet. We use a biaffine scorer to consider the sentiment dependencies obtained from both directions as the starting point and perform a reverse traversal of the results generated by the aspect and opinion markers.

For example, given the sentence "Green Tea creme brulee is a must !", with the BIO encoding strategy, we obtain the aspect tag {B,I,I,I,O,O,O,O} and opinion tag {O,O,O,O,O,O,B,O}. We also obtain the lexical sentiment dependence in two directions. In the $A \rightarrow O$ direction, this has the form of an index, (3,6,POS). In the $O \rightarrow A$ direction, if the model is correct, we obtain the same index representation as in the $A \rightarrow O$ direction, (3,6,POS). Here, index 3 represents the last word of the aspect (brulee), and index 6 represents the last word of the opinion (must), which form a positive sentiment together. Finally, we fuse the results predicted in the two directions heuristically. If the results predicted in the $A \rightarrow O$ direction are not predicted, they are added to the final result set, and vice versa. Based on the maximum probability predicted by the two-direction model, we believe that, if the $A \rightarrow O$ and $O \rightarrow A$ directions predict different sentiment dependencies for the same aspect and opinion pair, then the result with the maximum probability of sentiment dependence is the most accurate. The details of the algorithm are shown in 1.

Algorithm 1 Bidirectional decoding triplet

Input: aspect tags $\left\{\mathrm{g}_i^{(ap)}\right\}_{i=1}^{n}$, opinion tags $\left\{\mathrm{g}_i^{(op)}\right\}_{i=1}^{n}$
$A \rightarrow O$ direction sentiment dependency (j_1, k_1, m_1, p_1)
$O \rightarrow A$ direction sentiment dependency (j_2, k_2, m_2, p_2)
Output: Set of triples

function DecodingTriplet(j,k,m)
  $j' \longleftarrow j$
  while $\mathrm{g}_{j'}^{(ap)}$ is I **do** ◁ stop on B and O.
    $j' \longleftarrow j' - 1$
    if $j' \leq 0$ **then** ◁ or exceeding boundary.
      break
  $k' \longleftarrow k$
  while $\mathrm{g}_{j'}^{(ap)}$ is I **do**
    $k' \longleftarrow k' - 1$
    if $k' \leq 0$ **then**
      break
  return $\left[(j', j), (k', k), m\right]$
if $(j_1, k_1) = (j_2, k_2)$ **then**
  if $p_1 > p_1$ **then**
    $t \longleftarrow$ DecodingTriplet(j_1, k_1, m_1)
  else
    $t \longleftarrow$ DecodingTriplet(j_2, k_2, m_2)
else
  $t \longleftarrow$ DecodingTriplet(j_1, k_1, m_1) $\cup$ DecodingTriplet(j_2, k_2, m_2)
return t

3.7. Training

To learn the ASTE subtasks together, we merge the loss functions between the different modules. For aspect and opinion extraction, we minimize the cross-entropy loss function as follows:

Table 1
Statistics of datasets. #s and #t represent the number of sentences with overlapping triples and the number of triples with other triples, respectively.

Datasets	REST14			REST15			REST16			LAPTOP14		
	Train	Val.	Test	Train	Val.	Test	Train	Val.	Test	Train	Val.	Test
Sentence	1300	323	496	593	148	318	842	210	320	920	228	339
Triplet	2409	590	1014	977	160	479	1370	334	507	1451	380	552
#s	437	92	193	151	42	68	208	52	77	263	80	103
#t	578	147	389	189	62	71	256	61	120	365	101	140

$$L_{\text{ap}} = -\frac{1}{|S|} \sum_{i} \sum_{j} \hat{p}_{i,j}^{(ap)} \log\left(p_{i,j}^{(ap)}\right) \tag{10}$$

$$L_{\text{op}} = -\frac{1}{|S|} \sum_{i} \sum_{j} \hat{p}_{i,j}^{(op)} \log\left(p_{i,j}^{(op)}\right) \tag{11}$$

Similarly, for bidirectional affective dependence detection, the minimum cross-entropy loss function is calculated as follows:

$$L_{A->O} = -\frac{1}{|S|^2} \sum_{(i,j)} \sum_{k} \hat{s}_{i,j,k}^{A->O} \log\left(s_{i,j,k}^{A->O}\right) \tag{12}$$

$$L_{O->A} = -\frac{1}{|S|^2} \sum_{(i,j)} \sum_{k} \hat{s}_{i,j,k}^{O->A} \log\left(s_{i,j,k}^{O->A}\right) \tag{13}$$

Finally, we combine these loss functions to obtain the loss target of the entire model:

$$L(\theta) = L_{\text{ap}} + L_{op} + L_{A->O} + L_{O->A} \tag{14}$$

The AdamW [53] approach was adopted for the optimization of Eq. (14).

4. Experiments

4.1. Datasets

Our proposed DE-OTE-BISDD model was evaluated using datasets originally published by SemEval [54–56]. In the following sections, we refer to them as REST14, REST15, REST16, and LAPTOP14. REST14, REST15, and REST16 are datasets in the restaurant domain, whereas LAPTOP14 is in the laptop domain. Peng et al. [17] published the first version of this dataset. We referred to Peng et al. [17] and Wang et al. [57] with the same comments as in the term notes. Note that in the preprocessing step, we retained all overlapping opinion triples (OOTs) in the four datasets, which, on average, accounted for 24.2% of all triples in all four datasets. These settings are the same as those used by Zhang et al. [20]. Each dataset was divided into three subsets: training, validation, and testing. Table 1 presents their statistics.

4.2. Implementation details

In our experiment, we used a dual-embedding mechanism in the character-level encoding layer. We used 100-dimensional 1D filters for CNN character embedding, where each had a width of 5. At the word-direction coding level, word embedding was initialized with pretrained word vectors of 300-dimensional GloVe [24]. The BiLSTM hidden state d_h was set to 300. The regularized coefficient of L2, γ, was 10^5. To avoid overfitting, dropout was adopted with a discard rate of 0.2 for character-level embedding and 0.5 for word-direction embedding. During training, the learning rate was set to 10^3, and the batch size to 32. All parameters were initialized using a uniform distribution and optimized using the Adam optimizer. To measure the performance of our model and the baseline, we used precision, recall, and F1-score to evaluate the results of various subtasks. For testing, we selected the test results with the best validation performance.

4.3. Model

To demonstrate the validity of the DE-MTL-BISDD model, we compared it with the following baselines.

Pipeline [17]: Triplet extraction is divided into two stages. In the first stage, the aspect-sentiment pair and opinion label are extracted through a joint labeling scheme. In the second stage, the aspect tags attached to the sentiment tags in the first stage are paired with opinion tags to form an effective triad.

CMLA+[57]: A pairwise attention mechanism is used to simulate the interaction between aspect and opinion terms.

Unified+[40]: A uniform marking scheme is used to jointly extract aspect terms and associated sentiment.

RENANTE+[58]: This is a cooperative extraction system based on the BiLSTM-CRF model.

OTE-MTL [20]: This is a multitask learning framework proposed to jointly extract aspect and opinion terms and transform sentiment analysis tasks into table-filling problems.

In addition, to demonstrate the validity of our proposed method, we applied it to two other baseline models, CMLA-MTL and HAST-MTL [20], for multitask learning to examine the effects of their various components. The resultant models are as follows: (1) DE-CMLA-MTL and (2) DE-HAST-MTL: the double-embedding method is applied to the model such that the model can obtain the semantic representation of a context integrating different granularity levels; (3) CMLA-BISDD and (4) HAST-BISDD: the bidirectional sentiment-dependence detector is applied to the model; and (5) DE-CMLA-BISDD and (6) DE-HAST-BISDD: both the double embedding and bidirectional sentiment-dependence detector are applied to multitask learning.

5. Results and analysis

5.1. Comparison with the baselines

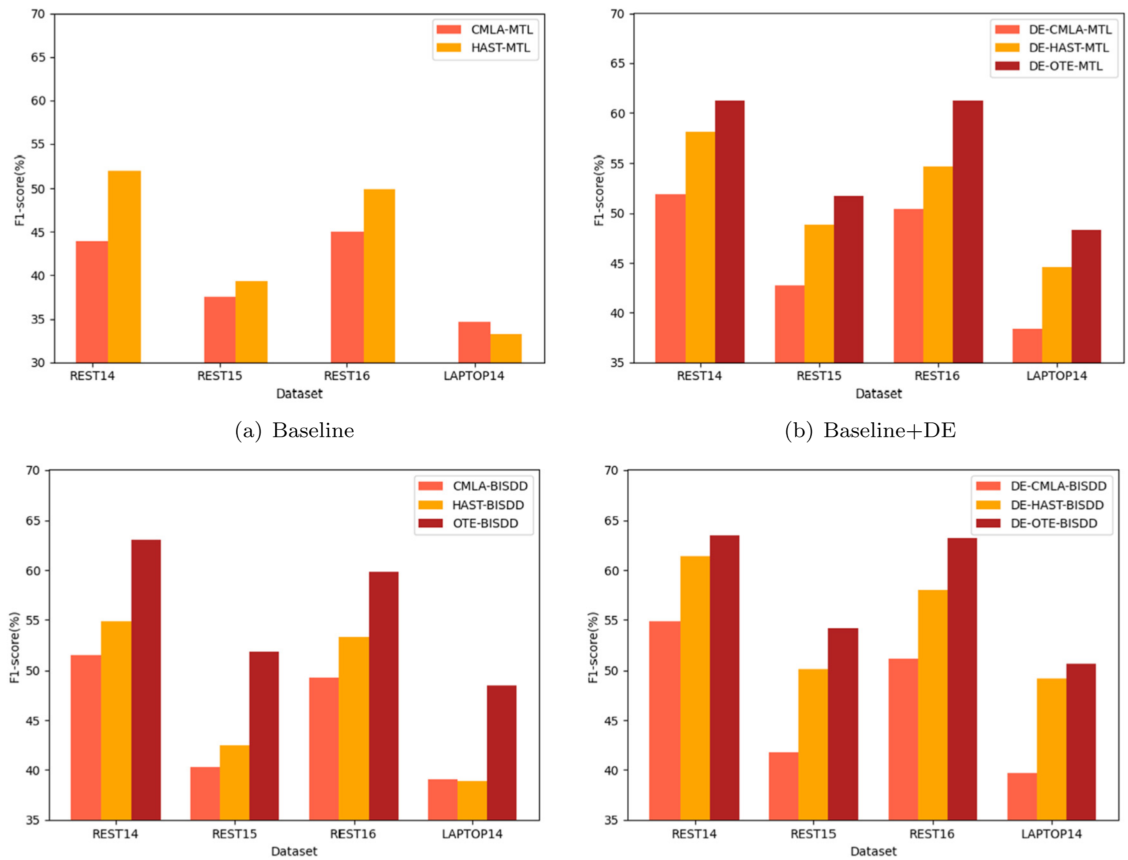
Table 2 lists the comparison results in terms of accuracy, recall rate, and F1-score between DE-OTE-BISDD and the previous model on the four datasets. Fair comparisons were made between datasets with and without OOTs. In particular, our proposed model had the best F1-scores across all datasets and significantly outperformed popular baselines in most cases. Compared with the most advanced model, OTE-MTL, on F1 values, our proposed model obtained 3.08%, 4.45%, 4.24%, and 6.39% absolute gains on REST14, REST15, REST16, and LAPTOP14, respectively, which shows that our proposed method can enhance the model's understanding of the entire sentence and effectively encode the interaction between the aspect and opinion, ultimately identifying more triples and improving the F1 values.

5.2. Ablation study

As shown in Table 3 and Fig. 4, we conducted ablation studies on the four datasets and applied the components to other baseline models to confirm the validity of the proposed approach. In general, in the double-embedding model, an improvement was observed over the previous baseline, particularly in DE-CMLA-MTL and DE-HAST-MTL. The F1-scores on the four datasets are

Table 2
Precision, Recall and F1-score were evaluated. Best results are in bold. DE-OTE-BISDD outperforms all the baselines significantly.

Model	REST14			REST15			REST16			LAPTOP14		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
RENANTE (2019)	30.90	38.30	34.20	29.40	26.90	28.00	27.10	20.50	23.30	23.10	17.60	20.00
CMLA+ (2017)	38.80	47.10	42.50	34.40	37.60	35.90	43.60	39.80	41.60	31.40	34.60	32.90
Unified+ (2019)	43.83	62.38	51.43	43.34	50.73	46.69	38.19	53.47	44.51	42.25	42.78	42.47
Pipeline (2019)	42.29	**64.07**	50.90	40.97	**54.68**	46.79	46.76	**62.97**	53.62	40.40	**47.24**	43.50
OTE-MTL (2020)	66.04	56.25	60.62	57.51	43.96	49.76	64.68	54.97	59.36	50.52	39.71	44.31
DE-OTE-BISDD (Ours)	**68.57**	59.17	**63.53**	**61.54**	48.43	**54.21**	**65.20**	61.34	**63.21**	**56.17**	46.20	**50.70**



(a) Baseline

(b) Baseline+DE

(c) Baseline+BISDD

(d) Baseline+DE+BISDD

Fig. 4. Experimental results of ablation study using the bidirectional sentiment dependence detector(BISDD) or the double embedding(DE) approach to the baseline model.

approximately 4–9 percentage points higher, whereas DE-OTE's performance on the four datasets is only approximately 1–4 percentage points higher. This demonstrates the effectiveness of identifying triples in models that incorporate contextual semantics at different granularity levels. In the CMLA-BISDD and HAST-BISDD models with only bidirectional sentiment-dependence detectors, the F1-scores improved by approximately 4–9 percentage points in most of the datasets, whereas OTE-BISDD improved by approximately 2–6 percentage points in the REST14, REST15, and LAPTOP14 datasets. However, the improvement was smaller for the REST16 dataset. This demonstrates the value of effectively modeling aspects and perspectives in models using a bidirectional sentiment-dependence detector. This method is based on two directions, which addresses the limitation of a single direction and improves the overall performance of the model through their joint decision.

Finally, we applied both the bidirectional sentiment-dependence detector and double-embedding approach to the baseline model and found that with both mechanisms, the effect of the HAST baseline improved significantly, whereas the effect of the OTE and CMLA baselines showed less improvement. We believe that the double-embedding mechanism has a significant positive impact on performance because the model can learn the sentiment dependence between word levels and characters well after obtaining the contextual semantic representation at different granularity levels and identify effective triples only in the one-way process.

5.3. Case analysis

We selected three representative examples from the test set as case studies. Table 4 lists the predictions for triples. Overall, our model yields reasonable results. In the first case, both directions, $A \rightarrow O$ and $O \rightarrow A$, work well and both provide correct results. For the second case, neither direction could predict all triples correctly. Because only part of the triple information is captured in a single direction, the predicted results in each direction are incomplete; however, all correct triples can eventually be obtained using the decoding strategy we developed. In the third

Table 3
According to the results of the ablation study of triplet extraction, the Precision, Recall and F1-score of different models were compared to analyze the influence of different components on the model.

Model	REST14			REST15			REST16			LAPTOP14		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
CMLA-MTL	43.24	44.95	43.97	35.87	39.85	37.55	44.22	46.43	45.01	33.61	36.11	34.68
HAST-MTL	58.97	46.75	52.04	41.48	37.58	39.32	52.32	48.56	49.92	47.70	25.74	33.24
DE-CMLA-MTL	63.49	48.03	51.91	46.24	43.63	42.76	52.43	53.25	50.44	39.37	40.94	38.35
DE-HAST-MTL	62.06	**57.59**	58.11	54.27	45.09	48.86	**64.51**	57.00	54.71	47.84	**48.19**	44.53
DE-OTE-MTL	**69.28**	54.94	**61.20**	**54.48**	**49.41**	**51.73**	61.88	**60.71**	**61.22**	**50.98**	45.97	**48.28**
CMLA-BISDD	50.52	52.47	51.48	38.20	42.59	40.28	49.21	49.31	49.26	37.01	41.30	39.04
HAST-BISDD	60.96	49.90	54.88	45.37	39.87	42.44	61.07	47.34	53.33	52.29	30.98	38.91
OTE-BISDD	**67.92**	**58.88**	**63.07**	**57.51**	**47.18**	**51.83**	**63.90**	**56.21**	**59.81**	**52.77**	**44.93**	**48.53**
DE-CMLA-BISDD	54.68	49.51	54.85	40.23	43.42	41.77	47.43	52.86	51.08	41.29	38.22	39.70
DE-HAST-BISDD	**72.27**	53.45	61.45	50.33	47.60	50.16	57.76	57.99	57.97	48.88	**47.64**	49.19
DE-OTE-BISDD	68.57	**59.17**	**63.53**	**61.54**	**48.43**	**54.21**	**65.20**	**61.34**	**63.21**	**56.17**	46.20	**50.70**

Table 4
Case analysis. Triples marked in red indicate incorrect predictions.

Case	Ground truth	A->O	O->A	DE-OTE-BISDD
The design and atmosphere is just as good.	[(design,good,POS), (atmosphere,good,POS)]	[(design,good,POS), (atmosphere,good,POS)]	[(design,good,POS), (atmosphere,good,POS)]	[(design,good,POS), (atmosphere,good,POS)]
Average to good Thai food, but terrible delivery.	[(Thai food, Average to good,POS), (delivery, terrible,NEG)]	[(Thai food, Average to good,POS)]	[(delivery, terrible,NEG)]	[(Thai food, Average to good,POS), (delivery, terrible,NEG)]
The atmosphere is unheralded , the service impecible , and the food magnificant .	[(atmosphere, unheralded,POS), (service,impecible,POS),(food, magnificant,POS)]	[~~*(atmosphere, unheralded, NEG)*~~, (service,impecible,POS),(food, magnificant,POS)]	[~~*(atmosphere, unheralded, NEG)*~~, (service,impecible,POS), ~~*(food, magnificant, NEG)*~~]	[~~*(atmosphere, unheralded, NEG)*~~, (service,impecible,POS),(food, magnificant,POS)]

case, partially incorrect triples are predicted in both directions. Among triples with the same aspects and views, our decoding strategy ultimately chooses the sentiment polarity with the highest probability as the final triplet. Although this corrects some of the predicted results of the triples, results that are incorrect in both directions remain.

6. Conclusion

In this study, we proposed a method based on double embedding and a bidirectional sentiment-dependence detector for ASTE. Our proposed approach improves the performance of the model by emphasizing the fusion of word vector features at different granularity levels to obtain higher-level contextual semantic vectors and transforming the sentiment prediction problem into a double-table-filling problem to model the interaction between aspects and opinions. We evaluated our model on four SemEval datasets, and the results showed that our model achieved a superior performance compared to several strong baselines for ASTE. In the future, our work will focus on extending these methods to other tasks in the natural language processing field and on developing stronger span-level aspect and opinion extractors.

CRediT authorship contribution statement

Dawei Dai: Conceptualization, Methodology, Writing, Funding acquisition. **Tao Chen:** Methodology, Data curation, Writing – original draft. **Shuyin Xia:** Conceptualization, Methodology, Resources. **Guoyin Wang:** Conceptualization, Review & editing. **Zizhong Chen:** Review & editing.

Declaration of competing interest

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

Acknowledgments

This work was sponsored by China Postdoctoral Science Foundation (2021M70 0562), Natural Science Foundation of Chongqing, China (No. cstc2019jcyj-msxmX0380), National Key Research and Development Program of China under Grant No. 2019QY(Y)0301.

References

[1] E. Cambria, Affective computing and sentiment analysis, IEEE Intell. Syst. 31 (2) (2016) 102–107.

[2] E. Cambria, S. Poria, A. Hussain, B. Liu, Computational intelligence for affective computing and sentiment analysis [guest editorial], IEEE Comput. Intell. Mag. 14 (2) (2019) 16–17.

[3] K. Schouten, F. Frasincar, Survey on aspect-level sentiment analysis, IEEE Trans. Knowl. Data Eng. 28 (3) (2016) 813–830.

[4] F. Tang, L. Fu, B. Yao, W. Xu, Aspect based fine-grained sentiment analysis for online reviews, Inform. Sci. 488 (2019) 190–204.

[5] J. Zhou, J.X. Huang, Q. Chen, Q.V. Hu, L. He, Deep learning for aspect-level sentiment classification: Survey, vision and challenges, IEEE Access PP (99) (2019) 1.

[6] Y. Zhang, Z. Zhang, D. Miao, J. Wang, Three-way enhanced convolutional neural networks for sentence-level sentiment classification, Inform. Sci. 477 (2018).

[7] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, An interactive multi-task learning network for end-to-end aspect-based sentiment analysis, 2019.

[8] Z. Wu, F. Zhao, X.Y. Dai, S. Huang, J. Chen, Latent opinions transfer network for target-oriented opinion words extraction, 2020.

[9] D. Ma, S. Li, F. Wu, X. Xie, H. Wang, Exploring sequence-to-sequence learning in aspect term extraction, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 3538–3547.

[10] Y. Zhou, W. Jiang, P. Song, Y. Su, S. Hu, Graph convolutional networks for target-oriented opinion words extraction with adversarial training, in: 2020 International Joint Conference on Neural Networks, IJCNN, 2020.

[11] B. Yang, C. Cardie, Joint inference for fine-grained opinion extraction, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers, 2013, pp. 1640–1649.

[12] B. Yang, C. Cardie, Extracting opinion expressions with semi-markov conditional random fields, in: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 2012, pp. 1335–1345.

[13] X. Cheng, W. Xu, T. Wang, W. Chu, Variational semi-supervised aspect-term sentiment analysis via transformer, 2018.

[14] Z. Fan, Z. Wu, X. Dai, S. Huang, J. Chen, Target-oriented opinion words extraction with target-fused neural sequence labeling, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Vol. 1, Long and Short Papers, 2019, pp. 2509–2518.

[15] J. Jiang, A. Wang, A. Aizawa, Attention-based relational graph convolutional network for target-oriented opinion words extraction, in: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, 2021, pp. 1986–1997.

[16] Y. Feng, Y. Rao, Y. Tang, N. Wang, H. Liu, Target-specified sequence labeling with multi-head self-attention for target-oriented opinion words extraction, in: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021, pp. 1805–1815.

[17] H. Peng, L. Xu, L. Bing, F. Huang, W. Lu, L. Si, Knowing what, how and why: A near complete solution for aspect-based sentiment analysis, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, (05) 2020, pp. 8600–8607.

[18] Z. Wu, C. Ying, F. Zhao, Z. Fan, R. Xia, Grid tagging scheme for aspect-oriented fine-grained opinion extraction, in: Findings of the Association for Computational Linguistics: EMNLP 2020, 2020.

[19] L. Xu, H. Li, W. Lu, L. Bing, Position-aware tagging for aspect sentiment triplet extract, 2020.

[20] C. Zhang, Q. Li, D. Song, B. Wang, A multi-task learning framework for opinion triplet extraction, 2020, arXiv preprint arXiv:2010.01512.

[21] S. Kiritchenko, X. Zhu, C. Cherry, S. Mohammad, NRC-Canada-2014: Detecting aspects and sentiment in customer reviews, in: Proceedings of the 8th International Workshop on Semantic Evaluation, SemEval 2014, 2014.

[22] S.V. Wawre, S.N. Deshmukh, Sentiment Classification using Machine Learning Techniques.

[23] T. Young, D. Hazarika, S. Poria, E. Cambria, Recent trends in deep learning based natural language processing, IEEE Comput. Intell. Mag. 13 (3) (2018) 55–75.

[24] J. Pennington, R. Socher, C.D. Manning, Glove: Global vectors for word representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2014, pp. 1532–1543.

[25] S. Poria, E. Cambria, A. Gelbukh, Aspect extraction for opinion mining with a deep convolutional neural network, Knowl.-Based Syst. 108 (2016) 42–49.

[26] H. Xu, B. Liu, L. Shu, P.S. Yu, Double embeddings and cnn-based sequence labeling for aspect extraction, 2018, arXiv preprint arXiv:1805.04601.

[27] E. Cambria, A. Kumar, M. Al-Ayyoub, N. Howard, Guest editorial: Explainable artificial intelligence for sentiment analysis, Knowl.-Based Syst. (2021) 107920.

[28] Target-Aware Convolutional Neural Network for Target-Level Sentiment Analysis.

[29] Z.A. Jie, C.B. Qin, C. Jxh, D. Qvh, H.A. Liang, Position-aware hierarchical transfer model for aspect-level sentiment classification - ScienceDirect, Inform. Sci. 513 (2020) 1–16.

[30] Q. Jiangtao, L. Chuanhui, L. Yinghong, L. Zhangxi, Leveraging sentiment analysis at the aspects level to predict ratings of reviews, Inf. Sci. Int. J. (2018).

[31] Y. Wang, M. Huang, X. Zhu, L. Zhao, Attention-based LSTM for aspect-level sentiment classification, in: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016.

[32] R.K. Amplayo, S. Lee, M. Song, Incorporating product description to sentiment topic models for improved aspect-based sentiment analysis, Inf. Ences 454 (2018) 200–215.

[33] M. Zhang, V. Palade, Y. Wang, Z. Ji, Attention-based word embeddings using artificial bee colony algorithm for aspect-level sentiment classification, Inform. Sci. 545 (2021) 713–738.

[34] T.H. Nguyen, K. Shirai, PhraseRNN: Phrase recursive neural network for aspect-based sentiment analysis, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015.

[35] B. Liang, H. Su, L. Gui, E. Cambria, R. Xu, Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks, Knowl.-Based Syst. 235 (2022) 107643.

[36] Y. Song, J. Wang, J. Tao, Z. Liu, Y. Rao, Attentional encoder network for targeted sentiment classification, 2019.

[37] X. Wei, L. Tao, Aspect based sentiment analysis with gated convolutional networks, in: Meeting of the Association for Computational Linguistics, 2018.

[38] A. Valdivia, M.V. Luzón, E. Cambria, F. Herrera, Consensus vote models for detecting and filtering neutrality in sentiment analysis, Inf. Fusion 44 (2018) 126–135.

[39] Z. Wu, F. Zhao, X.Y. Dai, S. Huang, J. Chen, Latent opinions transfer network for target-oriented opinion words extraction, 2020.

[40] X. Li, L. Bing, P. Li, W. Lam, A unified model for opinion target extraction and target sentiment prediction, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, (01) 2019, pp. 6714–6721.

[41] L. Xu, Y.K. Chia, L. Bing, Learning span-level interactions for aspect sentiment triplet extraction, 2021, arXiv preprint arXiv:2107.12214.

[42] S. Chen, Y. Wang, J. Liu, Y. Wang, Bidirectional machine reading comprehension for aspect sentiment triplet extraction, 2021, arXiv preprint arXiv:2103.07665.

[43] T.E. Trueman, E. Cambria, et al., A convolutional stacked bidirectional lstm with a multiplicative attention mechanism for aspect category and sentiment detection, Cogn. Comput. 13 (6) (2021) 1423–1432.

[44] E. Cambria, Q. Liu, S. Decherchi, F. Xing, K. Kwok, SenticNet 7: a commonsense-based neurosymbolic AI framework for explainable sentiment analysis, in: Proceedings of LREC 2022, 2022.

[45] A. Lieto, G.L. Pozzato, S. Zoia, V. Patti, R. Damiano, A commonsense reasoning framework for explanatory emotion attribution, generation and re-classification, Knowl.-Based Syst. 227 (2021) 107166.

[46] I. Perikos, S. Kardakis, I. Hatzilygeroudis, Sentiment analysis using novel and interpretable architectures of hidden Markov models, Knowl.-Based Syst. 229 (2021) 107332.

[47] E.F. Sang, J. Veenstra, Representing text chunks, 1999, arXiv preprint Cs/9907006.

[48] L. Ratinov, D. Roth, Design challenges and misconceptions in named entity recognition, in: Proceedings of the Thirteenth Conference on Computational Natural Language Learning, CoNLL-2009, 2009, pp. 147–155.

[49] M. Miwa, Y. Sasaki, Modeling joint entity and relation extraction with table representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2014, pp. 1858–1869.

[50] R.K. Srivastava, K. Greff, J. Schmidhuber, Highway networks, 2015, arXiv preprint arXiv:1505.00387.

[51] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.

[52] T. Dozat, C.D. Manning, Deep biaffine attention for neural dependency parsing, 2016, arXiv preprint arXiv:1611.01734.

[53] I. Loshchilov, F. Hutter, Fixing weight decay regularization in adam, 2017.

[54] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, Semeval-2014 task 4: Aspect based sentiment analysis, in: Proceedings of the 8th International Workshop on Semantic Evaluation, SemEval 2014, Association for Computational Linguistics, Dublin, Ireland, 2014, pp. 27–35.

[55] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, I. Androutsopoulos, Semeval-2015 task 12: Aspect based sentiment analysis, in: Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval 2015, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 486–495.

[56] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S.M. Jiménez-Zafra, G. Eryiğit, 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, California, 2016, pp. 19–30.

[57] W. Wang, S. Pan, D. Dahlmeier, X. Xiao, Coupled multi-layer attentions for co-extraction of aspect and opinion terms, 2017.

[58] H. Dai, Y. Song, Neural aspect and opinion term extraction with mined rules as weak supervision, in: Meeting of the Association for Computational Linguistics, 2019.