End-to-end aspect category sentiment analysis based on type graph convolutional networks[①]

SHAO Qing(邵 清)[②], ZHANG Wenshuang, WANG Shaojun

(School of Optoelectronic Information and Intelligent Engineering, University of Shanghai
for Science and Technology, Shanghai 200093, P. R. China)

Abstract

For the existing aspect category sentiment analysis research, most of the aspects are given for sentiment extraction, and this pipeline method is prone to error accumulation, and the use of graph convolutional neural network for aspect category sentiment analysis does not fully utilize the dependency type information between words, so it cannot enhance feature extraction. This paper proposes an end-to-end aspect category sentiment analysis (ETESA) model based on type graph convolutional networks. The model uses the bidirectional encoder representation from transformers (BERT) pre-training model to obtain aspect categories and word vectors containing contextual dynamic semantic information, which can solve the problem of polysemy; when using graph convolutional network (GCN) for feature extraction, the fusion operation of word vectors and initialization tensor of dependency types can obtain the importance values of different dependency types and enhance the text feature representation; by transforming aspect category and sentiment pair extraction into multiple single-label classification problems, aspect category and sentiment can be extracted simultaneously in an end-to-end way and solve the problem of error accumulation. Experiments are tested on three public datasets, and the results show that the ETESA model can achieve higher Precision, Recall and F1 value, proving the effectiveness of the model.

Key words: aspect-based sentiment analysis (ABSA), bidirectional encoder representation from transformers (BERT), type graph convolutional network (TGCN), aspect category and sentiment pair extraction

0 Introduction

With the rapid development of Internet shopping platforms, online and offline integrated marketing strategies have gradually become a new trend, and the research on sentiment analysis and opinion mining[1] has also made great progress, becoming a hot research field with significant research value in natural language processing. Traditional document-level sentiment analysis[2] and sentence-level sentiment analysis[3-4] assume in advance that a document or sentence contains only one sentiment tendency, namely positive, neutral or negative. However, due to the complexity and variety of sentiments expressed in the texts, such coarse-grained sentiment analysis cannot effectively extract all sentiments contained in different aspects of sentences. Therefore, more fine-grained aspect-based sentiment analysis (ABSA)[5-6] has been paid attention to by more and more researchers.

Aspect-based sentiment analysis usually consists of two subtasks: aspect term sentiment analysis (ATSA)[7] and aspect category sentiment analysis (ACSA)[8]. The ATSA task focuses more on detecting the sentiment polarity corresponding to explicit aspect terms in the sentences, while aspects in realistic review texts are often implicit and do not appear directly in sentences. The number of aspect categories corresponding to a domain is often fixed, and the annotations of aspect categories are often easier to obtain than aspect term annotations. Therefore, aspect category sentiment analysis is more suitable for application in real-world scenarios, and this paper mainly studies aspect category sentiment analysis.

Aspect category sentiment analysis includes two subtasks: aspect category extraction (ACE) and aspect

category corresponding sentiment analysis. Most of the current research extracts these two sub-tasks separately and conducts research of the two tasks in a pipelined way, which will cause the accumulation of errors and make the model performance not high. This paper transforms aspect category sentiment analysis into the problem of aspect-sentiment pair extraction (ASPE), trained in an end-to-end manner.

In recent years, methods based on deep learning have been widely used in ABSA tasks and achieved good results due to their ability to learn and extract features autonomously. Currently, model frameworks combining recurrent neural networks and attention mechanisms are more common in aspect-level sentiment analysis research. These models can deal with long text dependencies, mine the time sequence information and semantic information in the text, and imitate human neurons to give different weights to different words, which can more accurately capture the emotional tendencies in sentences. However, most of these studies do not consider the dependencies relationships and types of dependencies among words in sentences, resulting in underutilization of the results of syntactic analysis and making it impossible to achieve satisfactory results in sentiment analysis.

Therefore, a graph convolutional network (GCN) with joined word dependency types is proposed to enhance the learning of feature extraction by combining the dependencies types between words in a text with the dependencies. For the aspect-sentiment pair extraction task, this paper proposes a method that combines all aspect categories in the dataset with sentences as input, and then individually outputs the sentiment polarity of each aspect category. The output labels are converted into a four-category problem, that is, four categories including positive, neutral, negative, and none. For example, in the sentence 'while it was large and a bit noisy, the drinks were fantastic, and the food was superb', to extract all aspect-sentiment pairs contained in a sentence, all aspect categories (service, price, ambience, food, miscellaneous) in this dataset need to be fed into the network model jointly with the sentence, respectively, and output the sentiment polarity or no sentiment label of each aspect category. Regarding the imbalance in the dataset caused by this input, the focal loss function is used to solve the problem. This approach enables to solve the problem of end-to-end aspect category sentiment analysis (ETESA) and avoid the error accumulation problem generated by the pipelined method. The main contributions of this paper are as follows.

(1) Through data preprocessing, the aspect-senti-

ment pair extraction is transformed into multiple single-label classification problems, and the loss values obtained from aspect extraction and sentiment extraction are fed forward simultaneously, which can realize the simultaneous judgment of aspect category and its corresponding sentiment polarity and can solve the error accumulation problem caused by the traditional pipeline approach.

(2) The dependency type matrix is constructed in the type graph convolutional network (TGCN), and the word vector and the dependency type tensor are fused to distinguish the importance of the dependency type. At the same time, combined with the binary weight matrix, the text feature representation ability can be enhanced during semantic analysis.

(3) To realize the secondary interaction between aspect categories and contextual semantic information, firstly, the association information between aspect categories and contextual semantics is obtained through the bidirectional gating recurrent unit (BiGRU) network. On this basis, the attention module is designed to capture the sentiment words in the context that are closely related to the aspect in the context, thereby describing more complex contextual relationships of aspect words.

The rest of this paper is organized as follows. Section 1 introduces the related research work. Section 2 details the information of ETESA model proposed in this paper. Section 3 discusses the experimental details and conclusions. Section 4 is a summary of this paper and future directions.

1   Related work

In this section, it presents the state of the application of aspect term sentiment analysis, aspect category sentiment analysis, and multi-task joint learning in the field of sentiment analysis, which is included in aspect-level sentiment analysis.

1.1   Aspect term sentiment analysis

For aspect term sentiment analysis, many researchers use the basic framework of neural network and attention mechanism for sentiment recognition, and perform information interaction processing between aspect terms and context semantics to extract sentiment words related to the term. For example, Ma et al.[9] proposed an interactive attention network based on the long short term memory (LSTM) and attention to perform sentiment analysis by interactively learning the target and context representations through the attention mechanism to obtain their corresponding feature weight representations, respectively. Tian et al.[10] proposed a

method model for TGCN to ABSA with the help of a graph convolutional network of dependencies types. For edges of different dependency types, attention is used to give different weights to learn the graph neural network combining edges and nodes to obtain context semantic information, and attention mechanism is used to learn the weights of different TGCN layers. Wang et al.[11] proposed a relational graph attention network model, which obtained an aspect-oriented dependency structure by pruning the dependency graph, and solved the problem of incorrect matching of aspect information and sentiment information caused by multiple aspects in a sentence. Wang et al.[12] proposed a syntactic information-aware aspect-level sentiment classification model, which uses a memory network that combines an attention mechanism, part-of-speech (POS), position, text semantics and aspect information, and uses a GCN to improve the model classification performance.

1.2　Aspect category sentiment analysis

For aspect category sentiment analysis, since aspect categories are abstractions of terms and do not necessarily appear in the text, it is essential to focus on category-specific sentiment information in the context. For example, Wang et al.[8] proposed an attention-based LSTM network model that applies the aspect embedding twice consecutively, which can fully model the interdependence between the aspect information and the context input, thereby obtaining better sentiment analysis effect. Wang et al.[13] proposed a context-aware network model, which solved the problem of polysemy by introducing a multi-dimensional attention mechanism, and combined sentence semantic information and aspect category semantic information to enhance the interaction between aspect words and sentiment words, further improving the performance of sentiment classification. Xue and Li[14] proposed a network model based on gating mechanism and convolution, using Tanh gate and ReLU gate respectively to process word embedding and aspect embedding from the upper convolutional neural network (CNN) layer to achieve sentiment analysis of a given aspect, the model can be processed in parallel and is more efficient. Cai et al.[15] proposed a hierarchical graph convolutional network model to extract the relationship between the two subtasks of aspect category judgment and sentiment recognition. First, GCN was used to model the internal relationship between multiple categories, and then GCN was used to identify the sentiment of the category extracted by the previous layer.

1.3　Multi-task joint learning

In the current research, most scholars focus on predicting the sentiment of a given aspect term or category. However, a large amount of comment information in real life often does not give specific aspect information, so it is necessary to extract aspect information first, and then perform sentiment analysis. This method of sentiment classification is easy to lead to the accumulation of errors, that is, if the aspect extraction is wrong, it will lead to an error of the sentiment extraction of a specific aspect. Therefore, a model framework of multi-task joint extraction has been proposed recently, which is more suitable for application in real-world scenarios. For example, Wang et al.[16] proposed a hierarchical multi-task learning framework for end-to-end aspect-level sentiment analysis, which first performed aspect term extraction and sentiment word detection, and then performed aspect sentiment detection with the help of attention mechanism, and finally used LSTM and conditional random field (CRF) to extract aspect sentiment, and realized end-to-end aspect-level sentiment analysis by making full use of the information of lower-level auxiliary tasks by the upper-level main task. Zeng et al.[17] proposed an end-to-end network model based on joint learning, and used convolutional neural network CNN to obtain word-level and character-level feature information as embedding information for aspect category detection and aspect emotion recognition. Fu et al.[18] proposed a model framework for a multi-perspective attention mechanism based on a two-layer LSTM, which transformed aspect category sentiment analysis into an aspect category sentiment pair extraction task, and utilized two layers of LSTM to extract contextual semantics and aspect context semantic information respectively, and finally through the attention mechanism extracted the sentiment of multiple categories respectively. Due to the imbalance of the dataset, they proposed a loss function method to solve the problem of the imbalance of the data caused by the transformation of category sentiment to improve the model effect.

To sum up, the research on aspect term sentiment analysis tasks and aspect category sentiment analysis tasks is highly dependent on well-labeled aspect datasets. However, in real application scenarios, the aspect categories and terms of each sentence are often unknown, and it takes a lot of human and material resources that can only be obtained by labeling. In real life, it is often necessary to extract both aspect categories and their corresponding sentiment information. Therefore, in the case where the aspect categories are not given in advance in the dataset, studying aspect category sentiment pair extraction task via multi-task joint learning is challenging. Based on this, this paper

proposes an end-to-end aspect category sentiment analysis model based on the type graph convolutional network. A joint dependent type GCN network is used to enhance feature extraction, by converting the input format and using focal loss to solve the data imbalance problem. This framework provides a more accurate model approach for the aspect category sentiment extraction task.

2   Model description

This section will introduce in detail the end-to-end aspect category sentiment analysis model ETESA based on type graph convolutional network. The structure of the model is shown in Fig. 1. The ETESA model consists of an embedding layer, a BiGRU encoding layer, a type graph convolutional network layer, an aspect attention layer, and an output layer. The model uses the bidirectional encoder representation from transformers (BERT) pre-training model to vectorize aspects and

words, and then uses the BiGRU network layer to initially connect the aspect and context semantic information, learn the features between aspects and contexts, and then send the learned context information to TGCN. The network layer further strengthens the extraction of contextual features, and then uses the attention mechanism to give high weight to aspect-related sentiment words, and finally through the fully connected layer (FC) employs softmax for sentiment analysis.

2.1   BERT-based embedding layer

In this paper, the BERT pre-training model is used to implement the vector representation of words, which is different from the common aspect category sentiment analysis input format. All aspect categories contained in the dataset are taken as input together with sentences, and then the sentiment labels of each aspect category are output separately, which is transformed into multiple single-label classification problem.
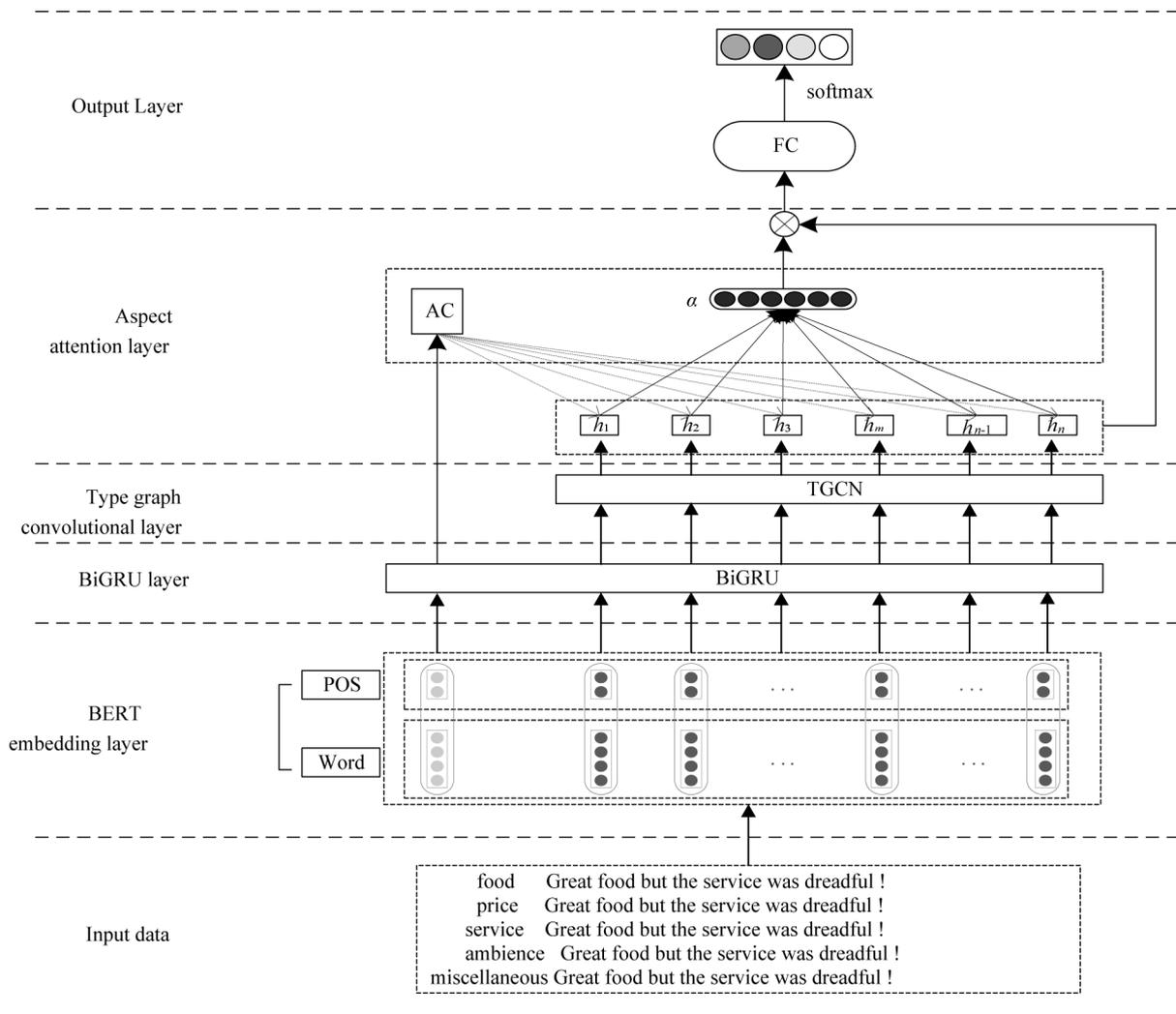
Fig. 1   Architecture of ETESA model

The BERT model is based on the Encoder structure of the bidirectional Transformer for feature extraction, which can distinguish texts with polysemy. To participate in the training of text data, it needs to be vectorized into a trainable data format. This paper uses the pre-training model BERT to vectorize the dataset. For a dataset containing m aspect categories $AC \{ a_i \}_{i=1}^m$, given a sentence $S = \{ w_1, w_2, \cdots, w_n \}$ consisting of n words, make '[CLS] + a_i + [SEP] + S + [SEP]' as the input of the embedding layer BERT, where a_i represents the aspect category word, S represents a piece of data, [CLS] and [SEP] respectively represent the start character and the aspect or sentence end character of the new input to the BERT model. After being encoded by the pre-trained model, the embedded representation of a sentence will be obtained as $I^e \in R^{(n+3+x_a) \times d_w}$, where n represents the number of words in a sentence, and the number of input characters is fixed to 3, x_a represents the number of aspect category words, and d_w represents the dimension of the BERT model embedding. Since the character embedding does not participate in the calculation of the next layer, the word vector input to the next layer is represented by $E^w \in R(n+x_a) \times d_w$. Each word in the sentence can obtain its corresponding part-of-speech tag through the Allennlp tool, but the aspect category has no part-of-speech information. In order to maintain the balance of the feature dimension, a virtual tag pT is added to represent the part-of-speech of the aspect category, so the part-of-speech can be obtained. The embedding is represented as $E^p \in R(n+x_a) \times d_p$, where d_p represents the dimension of the part-of-speech embedding. The feature vector $E^i \in R^{(n+x_a) \times (d_w+d_p)}$ obtained by splicing word embedding and part-of-speech embedding vector is sent to the BiGRU layer.

2.2 BiGRU encoding layer

As a variant of recurrent neural network (RNN), gating recurrent unit(GRU) can capture long-distance semantic information with timing information, and effectively solve the problems of gradient explosion and gradient disappearance during forward and backward propagation. In addition, because it only has update gates and reset gates, compared to LSTM, using GRU training can achieve the same effect as LSTM in performance, but can speed up the training time to a certain extent. Therefore, this paper uses the BiGRU network to initially associate aspect categories and sentences and learns the hidden layer output vector containing aspect and context information. The formula is as follows.

$$\overrightarrow{H} = \text{GRU} \{ a_1, \cdots a_{x_a}, h_1, h_2, \cdots h_n \} \qquad (1)$$

$$\overleftarrow{H} = \text{GRU} \{ h_n, h_{n-1}, \cdots h_1, a_{x_a}, \cdots a_1 \} \qquad (2)$$

$$H = (\overrightarrow{H}, \overleftarrow{H}) \qquad (3)$$

where, $h_t \in R^{1 \times d_{hid}}$ represents the hidden layer output vector of the word part in the sentence, $t \in (1, n)$; $a_j \in R^{1 \times d_{hid}}$ represents the aspect category hidden layer output vector, $j \in (1, x_a)$, this layer outputs $H \in R^{(n+x_a) \times 2d_{hid}}$.

2.3 Type graph convolutional network layer

Graph convolutional neural networks can efficiently extract features from graph structure information, and using the graph structure information derived from syntactically dependent features, and then applying it to GCN for aspect-level sentiment analysis is a common approach used by researchers in recent years. Ref. [10] proposed a TGCN network for aspect term sentiment analysis. At present, no one has applied dependent type information to aspect category sentiment analysis. Therefore, this paper uses a type graph convolution network to enhance feature extraction to realize aspect category sentiment analysis. First, use the Allennlp tool to obtain the dependencies between words, and establish the adjacency matrix $Adj = \{ a_{xy} \}_{n \times n}$ and relation matrix $Rel = \{ r_{xy} \}_{n \times n}$ through this dependency information. For these two matrices, it is generated by the following rules: if there is a dependency between word x and word y, let $a_{xy} = 1$ and $r_{xy} = r_i$ (where $0 < i < 50$, which represents the type of dependency, where 0 represents no dependency relationship). Then each relation matrix is mapped to the embedding representation $e_{xy} \in R^{1 \times 2d_{hid}}$. The detailed steps of the dependency type graph convolution layer are as follows. First, calculate the weight $w_{xy}^{(l)}$ of each edge between x and y in the lth layer TGCN, where $w_{xy}^{(l)} \in R^{1 \times 1}$.

$$s_x^{(l)} = h_x^{(l-1)} \oplus e_{xy} \qquad (4)$$

$$s_y^{(l)} = h_y^{(l-1)} \oplus e_{xy} \qquad (5)$$

$$w_{xy}^{(l)} = \frac{a_{xy} \cdot \exp(s_x^{(l)} \cdot s_y^{(l)})}{\sum_{y=1}^n a_{xy} \cdot \exp(s_x^{(l)} \cdot s_y^{(l)})} \qquad (6)$$

where, $\oplus$ represents the splicing of tensors, $\cdot$ represents the sum of the multiplication of the corresponding positions of the two tensors, $s_x^{(l)} \in R^{1 \times 4d_{hid}}$, and $s_y^{(l)} \in R^{1 \times 4d_{hid}}$.

Then through a learnable parameter matrix $w_R^{(l)}$ align e_{xy} and $h_y^{(l-1)}$, where $w_R^{(l)} \in R^{2d_{hid} \times 2d_{hid}}$, $h_y^{(l-1)} \in R^{1 \times 2d_{hid}}$, the formula is as follows:

$$h_y^{(l-1)} = h_y^{(l-1)} + e_{xy} w_R^{(l)} \qquad (7)$$

where $h_y^{(l-1)} \in R^{1 \times 2d_{hid}}$.

Finally, the hidden layer output of the lth layer of the word x is calculated as

$$h_x^{(l)} = \sigma\left(\sum_{y=1}^n w_{xy}^{(l)}\left(h_y^{(l-1)}w^{(l)} + b^{(l)}\right)\right) \quad (8)$$

where, $h_x^{(l)} \in R^{1 \times 2d_{hid}}$, $w^{(l)} \in R\,2d_{hid} \times 2d_{hid}$ and $b^{(l)}$ are trainable parameters, σ is the activation function ReLU, the final output of this layer is $\{h_1^{(l)}, h_2^{(l)}\cdots h_x^{(l)}\cdots h_n^{(l)}\}$.

2.4 Aspect attention layer

The attention mechanism can quickly extract important features of data, so it has been widely used in natural language processing. In the ACSA task, the model should pay more attention to words that have semantic associations with aspect category words. For example, in the sentence 'Delicious sushi and poor service', semantically, 'Delicious' is used to modify 'sushi', while 'poor' modifies 'service'. Therefore, when judging the sentiment tendencies of the category 'food', 'Delicious' should be given a higher weight. The main purpose of this layer is to capture the dependencies between aspect category words and contextual semantics and to enhance aspect-related feature extraction. This paper uses the bilinear attention score function as

$$Score(k_j, q) = qWk_j^{T} \quad (9)$$

where, $k_j \in R^{n \times (d_{hid} + d_a)}$ is the concatenation of the aspect category high-dimensional map vector after average pooling and the output vector $h_x^{(l)}$ of the TGCN hidden layer, $q \in R^{1 \times (d_{hid} + d_a)}$ is a learnable tensor, and $W \in R^{(d_{hid} + d_a) \times (d_{hid} + d_a)}$ is a learnable weight matrix. The formula for calculating the attention weight $\alpha(k_j, q) \in R^n$ is as follows.

$$\alpha(k_j, q) = \frac{\exp(Score(k_j, q))}{\sum_{j=1}^n \exp(Score(k_j, q))} \quad (10)$$

$$Z(k_j, q) = \alpha(k_j, q)k_j \quad (11)$$

2.5 Output layer

The output feature vector $Z(k_j, q) \in R\,(d_{hid} + d_a)$ of the previous layer is passed through the fully connected layer and the output tensor size is 4, and then the final sentiment prediction result is obtained through softmax function. The specific formula is as follows.

$$\hat{o} = \text{softmax}(Z(k_j, q)WT + b) \quad (12)$$

$$\hat{y} = \text{argmax}(\hat{o}) \quad (13)$$

where $W \in R^{d_c \times (d_{hid} + d_a)}$ and $b \in R^{d_c}$ are the learnable parameter matrix and bias, respectively, here d_c represents the number of categories; and $\hat{o}$ is a set of probability values predicted as four sentiment values; $\hat{y}$ is the final prediction result value. In this paper, the conversion of aspect-category sentiment pairs is used for sentiment analysis. Although the interpretability of the model is improved, it also causes data imbalance. Therefore, the focal loss function is used in this paper to solve the model performance caused by data imbalance. The specific expression is

$$L = -\sum_{t \in \hat{o}} \alpha_t (1 - \hat{o}_t)\gamma \log \hat{o}_t \quad (14)$$

where, $\hat{o}$ indicates that the output is a set of four emotions, α_t and γ are adjustable factors.

3 Experiment

3.1 Data set

This paper conducts a series of experiments on ETESA model on three public datasets to verify the interpretability and effectiveness of the ETESA model.

The first is the SemEval2014 restaurant dataset (Rest14). This dataset comes from the data for restaurant reviews in SemEval2014 task 4[19], which is commonly used for ACSA tasks. It includes 5 aspect categories: 'service', 'price', 'food', 'ambience' and 'miscellaneous'. It contains 4 sentiment categories: 'positive', 'negative', 'neutral' and 'conflict'. Since there are conflicting label items, this experiment removes the data of the 'conflict' label.

The other two datasets are the SemEval2015 and SemEval2016 restaurant datasets (Rest15 and Rest16). These two datasets come from the restaurant review datasets in SemEval2015 task 12[20] and SemEval2016 task 5[21], respectively, with 12 and 13 aspect categories each, and both contain 3 sentiment labels: 'positive', 'negative' and 'neutral'. The detailed statistics of the three datasets are shown in Table 1.

Table 1   The statistics of experimental datasets

Datasets		Aspect	Sentences	Positive	Negative	Neutral	Total
Rest14	train	5	2885	2179	839	500	3518
	test	5	767	657	222	94	973
Rest15	train	13	1117	1083	369	52	1504
	test	13	580	413	329	45	787
Rest16	train	12	1702	1501	697	100	2298
	test	12	585	513	195	42	750

3.2 Experiment setup and evaluation metrics

This paper adopts the PyTorch deep learning framework and uses the 768-dimensional dynamic pre-training model BERT provided by Devlin et al. [22] as the word embedding method. The part-of-speech embedding vector dimension is 332, and the BiGRU hidden layer unit is 250. In addition, the Allennlp tool is used to obtain information such as dependency syntactic relations and part-of-speech tags, and the syntactic dependency tree is visualized as shown in Fig. 2. This paper uses 2-layer TGCN and sets dropout to 0.5 for BiGRU. Using the Adam optimizer, set the learning rate size to 2e-5, the batch size to 8, and the epoch size to 30, and the result is the optimal value among all epochs. In the experiment, Precision, Recall and F1 values were used as the evaluation indicators to verify the model.
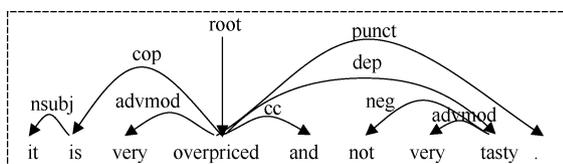
Fig. 2 An example of syntactic dependency tree

3.3 Models comparison

To comprehensively evaluate the performance of the ETESA model proposed in this paper on the ACSA task, this paper selects a series of representative compare experiments for comparison with ETESA, including:

TAN[23] proposes a model network based on GRU encoding and topic attention, which detects different aspect categories by paying attention to distinct parts in sentences based on various topics.

MTNA[24] proposes a multi-task learning model based on LSTM and CNN networks that simultaneously addresses both aspect category and aspect term extraction.

BERT-pair-NLI-B[25] constructs auxiliary sentences based on aspects, converts ABSA to sentence pair classification task, and implements sentiment analysis by fine-tuning the pre-trained BERT model.

TAS-BERT[26] proposes a target-aspect-sentiment triplet joint detection model based on the BERT pre-trained model.

MEJD[27] proposes a model based on BERT, LSTM and graph attention convolutional network GACN to capture the dependencies between aspects and sentences, and realizes the target-aspect-sentiment triplet multi-task sentiment analysis model.

3.4 Experiment results

In order to verify the effectiveness of the proposed method, this paper compares each dataset with related research in recent years. As can be seen from Table 2, on the ASPE task, the REST14 and REST16 datasets have achieved the best F1 values, Precision and Recall, and the F1 value of the ETESA model is 0.41% and 0.70% higher than that of the MEJD model, while the results on the REST15 dataset are higher than those of the MEJD but slightly lower than the TAS-BERT model. Through analysis of the results, it is found that adding dependency type information to the model makes the results of this model better than that of MEJD. At the same time, because the distribution ratio of different sentiment polarities in the training set and test set in the REST15 dataset is quite different, the F1 value of ETESA on this dataset will be slightly lower than the TAS-BERT model based on BERT. On the ACE task, by analyzing the experimental results, it is found that the F1 value of ETESA achieves the current best

Table 2 Performance comparison of different model

Tasks	Models	REST14			REST15			REST16		
		Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
ACE	BERT-pair-NLI-B	93.91	**88.70**	91.23	75.63	66.51	70.78	87.04	74.44	80.25
	TAN	-	-	90.61	-	-	-	-	-	78.38
	MTNA	-	-	88.91	-	-	65.97	-	-	76.42
	TAS-BERT	92.68	88.15	90.36	82.32	71.17	76.34	88.92	75.34	81.57
	MEJD	94.29	88.16	91.12	82.74	72.07	77.04	89.07	**78.80**	83.62
	ETESA	**94.96**	88.39	**91.56**	**83.48**	**74.08**	**78.50**	**90.93**	78.52	**84.27**
ASPE	BERT-pair-NLI-B	82.72	78.85	80.74	66.73	60.88	63.67	77.51	68.45	72.70
	TAS-BERT	84.79	80.20	82.43	71.55	**65.70**	**68.50**	78.96	69.84	74.12
	MEJD	85.82	80.72	83.19	71.28	64.64	67.80	79.73	71.17	75.21
	ETESA	**86.33**	**81.04**	**83.60**	**71.93**	64.86	68.21	**80.40**	**71.90**	**75.91**

results. Analyzing the results of three different datasets, it can be found that the models REST14 and REST16 perform better. By analyzing Table 1, it is found that this is mainly because the data in REST15 is not enough, and the model is not enough to learn better features. Therefore, the performance of the model on this data set is inferior to that of other data sets.

Overall, ETESA achieves a relatively state-of-the-art on the ASPE task, demonstrating the effectiveness of end-to-end type graph convolutional neural network approach.

3.5　Ablation experiment

To investigate the effectiveness of different modules (part-of-speech, BiGRU, type-dependent) in the model, the following ablation experiments were studied and compared with the full experiments. Table 3 shows the experimental results using different modules on 3 datasets. It can be observed that the model proposed in this paper has the highest F1 value, and if any one module is removed from the full model, the model performance on all three data sets will be degraded. From

the comparison results of models 1 and 4, it can be known that the dependency type information in dependency parsing can improve the performance of the model to a certain extent, because the importance of different dependency types in sentences is different. By giving different weights to them through attention, the capture of aspects and related sentiment words can be improved. From the experimental results of models 2 and 4, it can be concluded that the BiGRU module is crucial to the task of this article, as it can extract contextual semantic information and initially obtain the interaction information between aspect categories and sentences. From the results of models 3 and 4 on all datasets, it can be seen that after adding part-of-speech information, the performance of aspect-sentiment pair extraction task is improved more than other related models, which indicates that part-of-speech has a greater impact on the sentiment classification task. To sum up, part-of-speech, BiGRU, and type-dependent all play an important role in enhancing aspect-sentiment pair extraction tasks, and each module has its own unique contribution to the model.

Table 3　Ablation experimental results ("✓" and "✗" represent if a corresponding module is used or not)

	Setting				REST14			REST15			REST16		
	POS	BiGRU	TGCN	GCN	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
1	✓	✓	✗	✓	85.88	80.01	82.84	71.41	61.68	66.19	79.20	68.62	73.53
2	✓	✗	✓	✗	85.26	79.71	82.39	71.54	61.93	66.39	79.07	67.26	72.69
3	✗	✓	✓	✗	84.56	79.52	81.96	68.99	60.88	64.68	77.07	66.84	71.59
4	✓	✓	✓	✗	86.33	81.04	83.60	71.93	64.86	68.21	80.40	71.89	75.91

3.6　Parametric study

This section explores the impact of different hyperparameters on model performance, including the number of layers of TGCN and the size of α_t in the loss function.

Fig. 3 shows the different experimental results for the TGCN layer number from 0 to 4. It can be seen that when the number of layers is 2, the model obtains the optimal F1 value. When the number of layers is 0, that is, the TGCN layer is not used, the effect is the worst, which proves the effectiveness of the TGCN layer. When the number of layers is 1 and 3, the performance of the model is not as good as that of 2 layers. The possible reason is that when the number of layers is 1, the model does not capture the dependencies between all words. When the number of layers is 3, the model is overfitted, so the performance on the test set is reduced.

Fig. 4 shows the study of the hyperparameter α_t for
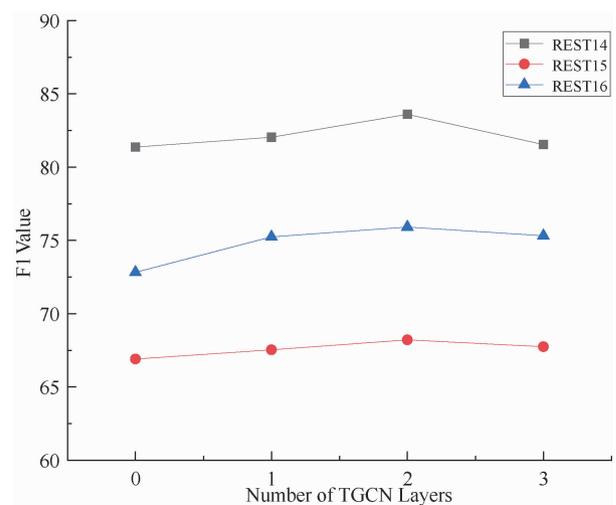
Fig. 3　The influence of different TGCN layers

balanced samples in the loss function, three common values with an interval of 0.25 were chosen for the experiments. The results show that the model has the best performance when the value of α_t is 0.5. Because in

the training phase, the samples are weighted random sampling, so that the emotional and non-emotional samples in each batch are uniform, so the model performance is the best when the value is 0.5.
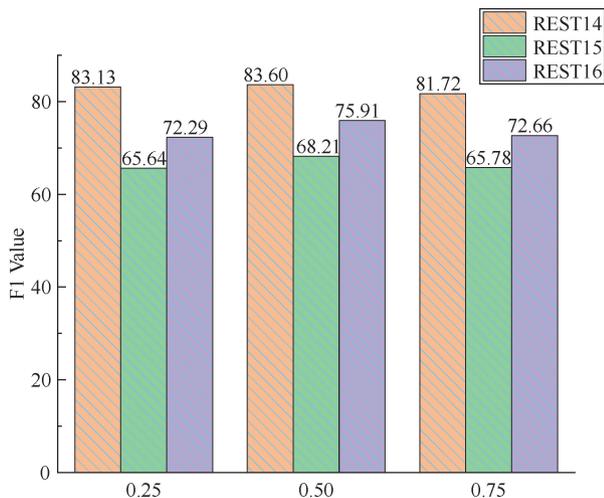
Fig. 4　The influence of balanced sample scale factor α_t

4　Conclusions

Aiming at the problems that syntactic dependency information is not fully utilized in the current research on aspect category sentiment analysis tasks, there is a lack of effective mechanisms to distinguish the important relationships of different dependency types, and the pipeline method used in most studies is prone to error accumulation. This paper designs and proposes an end-to-end aspect category sentiment analysis (ETESA) model based on type graph convolutional networks. The model converts the three-category sentiment problem into a four-category problem. First, the pre-trained BERT model is used to obtain word embedding information, and this model can solve the problem of polysemy in the text. Then, the BiGRU network is used to preliminarily combine the aspect and context semantic features. After that, the feature extraction of the context is further achieved through the TGCN network, and finally the attention mechanism is used to obtain the sentiment related to the aspect category. The comparative experimental results on three public datasets show that the ETESA model proposed in this paper can effectively improve the performance of aspect category sentiment analysis.

The recent related research on sentiment analysis[26-27] can achieve more fine-grained sentiment analysis by jointly extracting three tuples of aspect categories, aspect sentiment, and aspect terms, compared with the two tuples proposed in this paper. Therefore, in future work, the model method of triplet task combination extraction will be further studied, so that it can achieve joint extraction of more tasks.

References

[1]　LIU B, ZHANG L. A survey of opinion mining and sentiment analysis[M]. Boston: Springer, 2012: 415-463.

[2]　MORAES R, VALIATI J F, NETO W G. Document-level sentiment classification: an empirical comparison between SVM and ANN [J]. Expert Systems with Applications, 2013, 40(2): 621-633.

[3]　WU F, ZHANG J, YUAN Z, et al. Sentence-level sentiment classification with weak supervision[C] // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. Tokyo: Association for Computing Machinery, 2017: 973-976.

[4]　XIE R Z, LI Y. Text sentiment classification model based on BERT and dual channel attention[J]. Journal of Data Acquisition and Processing, 2020, 35 (4): 642-652. (In Chinese)

[5]　ZHOU R M, HU X Y, WU K W, et al. Aspect-based sentiment analysis with cross-heads attention[J]. Computer Engineering and Applications, 2023, 59 (9): 190-197. (In Chinese)

[6]　FU C Y, HUANG X Y, LIU H K, et al. Aspect level sentiment analysis of syntactic feature fusion enhancement[J]. Journal of Chinese Computer Systems, 2023, 44 (4): 682-689. (In Chinese)

[7]　CHEN P, SUN Z, BING L, et al. Recurrent attention network on memory for aspect sentiment analysis[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen: EMNLP, 2017: 452-461.

[8]　WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification[C] // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: EMNLP, 2016: 606-615.

[9]　MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification[C] // The 26th International Joint Conference on Artificial Intelligence. Melbourne: IJCAI, 2017: 4068-4074.

[10]　TIAN Y, CHEN G, SONG Y. Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble[C] // Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Mexico: Association for Computational Linguistics, 2021: 2910-2922.

[11]　WANG K, SHEN W, YANG Y, et al. Relational graph attention network for aspect-based sentiment analysis[C] // The 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, 2020: 3229-3238.

[12]　WANG G, LI H Y, QIU Y F, et al. Aspect-based sentiment classification via memory graph convolutional network [J]. Journal of Chinese Information Processing, 2021, 35 (8): 98-106. (In Chinese)

[13]　WANG J J, JIANG M, ZHANG M. Context-aware learning

for aspect category sentiment classification[J]. Application Research of Computers,2021,38(6): 1770-1774. (In Chinese)

[14] XUE W,LI T. Aspect based sentiment analysis with gated convolutional networks[C]// The 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics,2018: 2514-2523.

[15] CAI H,TU Y,ZHOU X,et al. Aspect-category based sentiment analysis with hierarchical graph convolutional network[C]// Proceedings of the 28th International Conference on Computational Linguistics. Barcelona: COLING, 2020: 833-843.

[16] WANG X,XU G,ZHANG Z,et al. End-to-end aspect-based sentiment analysis with hierarchical multi-task learning[J]. Neurocomputing,2021,455: 178-188.

[17] ZENG Z,MA J,CHEN M,et al. Joint learning for aspect category detection and sentiment analysis in Chinese reviews[C]// China Conference on Information Retrieval. Fuzhou:CCIR,2019: 108-120.

[18] FU Y,LIAO J,LI Y,et al. Multiple perspective attention based on double BiLSTM for aspect and sentiment pair extract[J]. Neurocomputing,2021,438: 302-311.

[19] PONTIKI M,D GALANIS,PAVLOPOULOS J,et al. SemEval-2014 task 4: aspect based sentiment analysis[C] // The 8th International Workshop on Semantic Evaluation. Dublin: SemEval,2014: 27-35.

[20] PONTIKI M,GALANIS D,PAPAGEORGIOU H,et al. Semeval-2015 task 12: aspect based sentiment analysis[C] //The 9th International Workshop on Semantic Evaluation. Denver: SemEval,2015: 486-495.

[21] PONTIKI M,GALANIS D,PAPAGEORGIOU H,et al. Semeval-2016 task 5: aspect based sentiment analysis[C] //The 10th International Workshop on Semantic Evaluation. San Diego: SemEval,2016: 19-30.

[22] DEVLIN J,CHANG M W,LEE K,et al. Bert: Pre-training of deep bidirectional transformers for language understanding[C] // 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics,2019: 4171-4186.

[23] MOVAHEDI S,GHADERY E,FAILI H,et al. Aspect category detection via topic-attention network [EB/OL]. (2019-06-21) [2022-11-26]. https: // arxiv. org/pdf/ 1901. 01183pdf.

[24] XUE W,ZHOU W,LI T,et al. MTNA: a neural multi-task model for aspect category classification and aspect term extraction on restaurant reviews[C]// Proceedings of the 8th International Joint Conference on Natural Language Processing. Taibei: IEEE,2017: 151-156.

[25] SUN C,HUANG L,QIU X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence [C] // 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics,2019: 380-385.

[26] WAN H,YANG Y,DU J,et al. Target-aspect-sentiment joint detection for aspect-based sentiment analysis[C] // Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI Press,2020: 9122-9129.

[27] WU C,XIONG Q,YI H,et al. Multiple-element joint detection for aspect-based sentiment analysis [J]. Knowledge-Based Systems,2021,223: 107073.

SHAO Qing,born in 1970. She received her Ph. D degree in Donghua University in 2010. Her research interests include intelligent computing and fault tolerant computing.