applied sciences

Article

An Interactive Learning Network That Maintains Sentiment Consistency in End-to-End Aspect-Based Sentiment Analysis

Musheng Chen, Qingrong Hua, Yaojun Mao and Junhua Wu *

School of Software Engineering, Jiangxi University of Science and Technology, Nanchang 330016, China; dreaminit@163.com (M.C.); 6720210703@mail.jxust.edu.cn (Q.H.); myj_1123@163.com (Y.M.)

* Correspondence: wujunhua1116@163.com

Abstract: Most of the aspect-based sentiment analysis research completes the two subtasks (aspect terms extraction and aspect sentiment classification) separately, and it cannot see the full picture and actual effect of the complete aspect-based sentiment analysis. The purpose of end-to-end aspect-based sentiment analysis is to complete the two subtasks of aspect terms extraction and aspect sentiment classification at the same time, and the current research in this area focuses on the connection between the two subtasks and uses the connection between them to construct the model. However, they rarely pay attention to the connection between different aspects and ignore the sentiment inconsistency within the aspects caused by the end-to-end model. Therefore, we propose an interactive learning network to maintain sentiment consistency, first using the multi-head attention mechanism to achieve the interaction between aspects and subtasks and then using the gate mechanism to design an auxiliary module to maintain sentiment consistency within aspect items. The experimental results on the datasets Laptop14, Restaurant14, and Twitter showed that, compared with the optimal benchmark method, the F1 values of the proposed method increased by 0.4%, 1.21%, and 5.22%, respectively. This indicates that the proposed method can effectively consider the relationships between aspect items and maintain emotional consistency within the aspect items.

Keywords: aspect-based sentiment analysis; multi-head attention mechanism; sentiment consistency; interactive learning

check for updates

Citation: Chen, M.; Hua, Q.; Mao, Y.; Wu, J. An Interactive Learning Network That Maintains Sentiment Consistency in End-to-End Aspect-Based Sentiment Analysis. *Appl. Sci.* **2023**, *13*, 9327. https://doi.org/10.3390/app13169327

Academic Editors: Xiaoke Ma and Jingjing Deng

Received: 10 July 2023
Revised: 2 August 2023
Accepted: 15 August 2023
Published: 17 August 2023

1. Introduction

With the rapid development of information technology, the Internet has become an indispensable part of people's lives, and people increasingly like to leave their own reviews on various platforms. Effective use of these review text data, mining and analyzing the sentiment tendencies they contain, has important practical application value; for example, reviews in the e-commerce platform have important reference value for other users to choose commodity, and also help merchants improve their products and services. As one of the most popular social media platforms worldwide, Twitter contains views, opinions, and ideas on a variety of topics, and in the context of recent works related to Twitter data mining and analysis, a number of works have focused on the sentiment analysis of tweets, which can provide effective information for various fields [1,2].

Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment analysis task; its purpose is to identify the different aspects of a text and judge their corresponding sentiment polarity [3]. ABSA can be divided into two subtasks: aspect terms extraction (ATE) and aspect sentiment classification (ASC) [4]. The purpose of end-to-end ABSA is to complete the two subtasks of ATE and ASC at the same time to achieve the complete ABSA process. Since the ATE task is a sequence-labeling task and the ASC task is a classification task, in order to make the end-to-end model solve the two tasks together, the current research methods are to treat the ASC task as a sequence-labeling task [5].

In the initial study, each subtask was considered separately and was not linked with other tasks [5,6]. Recently, more and more research has begun to focus on the connection between two subtasks and deal with them together by analyzing the connection between the two subtasks [7–9]. However, they do not take into account that there is also a connection between the aspects, such as “nice operating system and keyboard”, the sentiment polarity of the aspect term “operating system” is determined by “nice”, and “operating system” and “keyboard” are connected by the juxtaposition “and”, then the sentiment polarity of these two aspects should be the same and both positive. However, it may be affected by the context and labeled as a different sentiment. Secondly, “operating system” is an aspect item composed of two words; because we treat the aspect sentiment classification as a sequence-labeling task, then there may be a multi-word aspect item with two different sentiment polarities, for example, “operating system” may be labeled as “B-POS E-NEU”, which appears as an aspect item that contains two different sentiment polarities, which is obviously unreasonable.

In response to the above problems, an interactive learning network that maintains sentiment consistency is proposed. In this network, the task-sharing layer is used to link the two subtasks preliminarily; then, the multi-head attention mechanism is used to combine the context information to let the model consider the connection between the aspects and realize the interaction between the two subtasks so that the model can fully interact with information and consider the context information of the text. An auxiliary component with a gate mechanism is also designed, which can be used to maintain sentiment consistency within aspect items by considering the results of the previous moments at the current moment and correcting the current results.

Our main purpose and contributions are summarized as follows:

- A new framework is proposed to address the complete ABSA in an end-to-end manner. Use the task-sharing layer to enable interaction between two subtasks and take advantage of the multi-head attention mechanism to consider the connection between aspect items;
- An auxiliary component with a gate mechanism is designed to maintain sentiment consistency within aspect items.

2. Related Work

In the past, ABSA was carried out separately from two subtasks, only studying ATE [10], aiming to extract aspects in the text, or only studying ASC [11], using the aspect items obtained in advance to classify their sentiment, both of which currently have good results, but they cannot see the full picture and actual effect of complete ABSA. There is a pipeline method [12], which first uses one model to extract the aspects contained in the text and then another model to classify the sentiment of these aspects and complete two subtasks simultaneously to obtain a complete ABSA model. The two submodels of this approach may work well when used separately, but this pipe-based approach requires data to flow across multiple models that cannot be backpropagated, and the erroneous predictions of one model are passed on to the next, so the end result must be wrong. In addition, the pipeline method requires training two models separately, which also results in a waste of computing resources. Therefore, end-to-end research is necessary.

To process two subtasks at the same time, Wang et al. [5] proposed a multi-task neural learning framework that can simultaneously process ATE and ASC subtasks and use the attention mechanism to learn the joint representation of aspect and affective relationship. Li et al. [6] used stacked LSTMs to predict the boundary and unified prediction of aspect terms and designed three auxiliary components to correct the prediction results. These end-to-end models can achieve multiple tasks in one step, avoid the conduction of false predictions in multiple models, and at the same time, can make the error backpropagate within the model, ultimately reducing the error rate and improving the accuracy rate. However, they only complete the two subtasks at the same time without considering the connection between them, so the effect is also poor. Luo et al. [13] and He et al. [14] proposed that the previous research did not make full use of the interconnection of the two subtasks and proposed a shared network and a multi-task learning network, respectively, using different methods to achieve the interaction between the two subtasks. However, the study of the two of them only linked the two subtasks and did not expand on other content, so the results were more general.

With the continuous development of deep learning, the pre-trained model BERT [15] gradually replaced the traditional word vector model and was widely used in various research fields of natural language processing. Li et al. [16] first applied BERT to end-to-end ABSA when they constructed a series of simple but effective neural baselines for sequence labeling problems and used BERT to fine-tune task-specific components. The results showed that the pre-trained model BERT was very effective; just BERT plus a simple linear classification layer had better results than before.

Luo et al. [17] proposed that the imbalance between labels would affect performance, so they extended the gradient coordination strategy to alleviate this problem and used virtual adversarial training and post-training of domain datasets to improve collaborative extraction performance. They came to the conclusion that alleviating the label imbalance problem is more important for sequence labeling. Oh et al. [18] proposed a deep-context relationship-aware network DCRAN based on aspect sentiment analysis, which allows implicit interaction between subtasks in a more efficient way and allows two explicit self-monitoring strategies for deep context and relationship-aware learning; the results show that their two explicit self-monitoring strategies are very effective.

The above research improves the performance of the model by exploring different contents but does not solve the problem of the connection between the aspects and maintaining the sentiment consistency within aspect items. Hence, an interactive learning network that maintains sentiment consistency is proposed.

3. An Interactive Learning Network That Maintains Sentiment Consistency

The model structure in this paper is shown in Figure 1. Firstly, the text is input into BERT to obtain the context representation of the text, and the front l layer of BERT is used as the task-sharing layer to obtain the shared features of the two subtasks. Then, the remaining layer of BERT is used as the aspect terms extraction layer to obtain the aspect term feature representation, and the aspect term extraction results are obtained through the classification layer. Then, the feature representation of the aspect item is input into the transformer decoder module with multiple multi-headed attention, and the interaction between the aspect items and the interaction between the two subtasks is realized in combination with the shared features. Finally, the sentiment consistency component is used to maintain the sentiment consistency of the aspect items, and the final result is obtained through the classification layer.

3.1. Task Definition

In order to merge the two subtasks of ATE and ASC into one task simultaneously, it is necessary to combine the labeling task of ATE and the classification task of ASC into a unified labeling task. Given a text C of length n, denoted by $C = \{w1, w2, w3, \ldots, wn\}$, label each word with a unified tagging scheme: Ys = *{B-POS, I-POS, E-POS, S-POS, B-NEU, I-NEU, E-NEU, S-NEU, B-NEG, I-NEG, E-NEG, S-NEG, O}*. Each label contains two parts of label information, "B, I, E, S, and O", representing the beginning, the middle, the end, a single word aspect, and non-aspect words, respectively. "POS, NEG, and NEU" indicate positive, negative, and neutral sentiments. The predicted final marker sequence is $Y = \{y1, y2, y3, \ldots, yn\}$, where $y_i \in Y_s$.

3.2. Encoding Layer

This model uses the pre-trained language model BERT as the encoding layer to construct a contextual representation of the text. In addition to the input layer and output layer, each of the remaining layers of the BERT model is composed of multiple transformer modules, and the input of each transformer module in each layer comes from the output of the previous layer. Given a text C, after preprocessing it, such as word segmentation, the processed data are input into the pre-trained model BERT, and the output H is the context representation sequence of the text, defined as shown in Equation (1).

$$H^{[1:L]} = BERT(C) \tag{1}$$

where $H^{[1:L]}$ represents each layer of BERT output,. There are a total of L layers; for example, H^L represents the last layer of BERT, which is the largest layer, $H^L \in R^{m*d}$, where m represents the encoded length of C and d represents the dimension of the word vector.

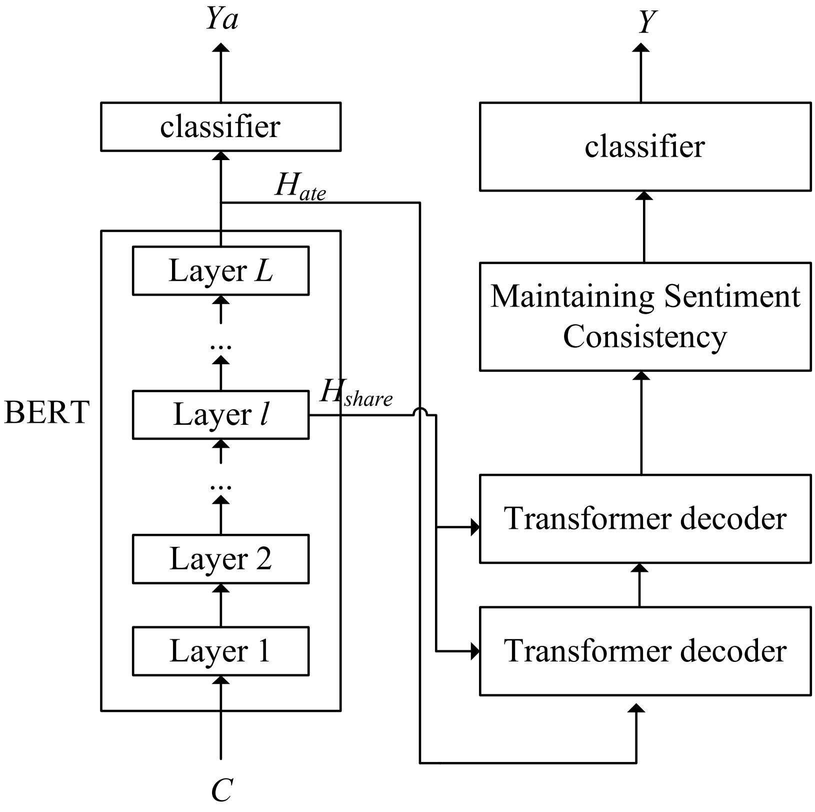

Figure 1. A framework for interactive learning network models that maintains sentiment consistency.

3.3. Task-Sharing Layer

Jawahar et al. [19] believe that different layers of BERT can capture different levels of information; for example, the underlying network captures surface information, the middle network captures syntactic information, and the higher network captures semantic information. The research of this paper requires not only semantic information but also syntactic information, but as the number of layers of BERT increases, the syntactic information in the output features will decrease, so we can take the front l layer of BERT as the shared layer, that is, the output of the l layer is the shared feature representation of the two subtasks H_{share}, which contains context information such as aspect item information, sentiment information, syntactic information and semantic information of text, and the l layer to the L layer of BERT is used for the aspect item extraction subtask. That is, the L-layer output is the feature representation of the aspect extracted by the H_{ate}. The definitions are shown in Equations (2) and (3).

$$H_{ate} = H^L \tag{2}$$

$$H_{share} = H^l \tag{3}$$

The feature representation of the ATE task is input into the classification layer, and the aspect entity sequence Y_a is obtained. The formula is as follows: W_a is the weight matrix of the linear layer, and b_a is the bias term.

$$Y_a = softmax(W_a H_{ate} + b_a) \tag{4}$$

The results produced here are not used as the final aspect items to extract the task results but only play an auxiliary role for later use of the multi-head attention mechanism to consider the interaction between the aspect items.

3.4. Interaction Layer

Then, use two transformer decoder [20] modules to consider the relationship between aspect items and between two subtasks. The transformer decoder module is formed by stacking layers of multiple multi-headed attention mechanisms on top of each other, which can learn the behavior of different tasks and then combine different behaviors as knowledge, which can take into account multiple aspects. The transformer decoder module in this article consists of a layer of multi-head self-attention, a layer of multi-head cross-attention, and a feedforward neural network (FNN); the specific structure of the transformer decoder module is shown in Figure 2. In the multi-head attention mechanism, Q represents the query matrix, K represents the key matrix, V represents the value matrix, and Q, K, and V are obtained by different linear transformations based on different weight matrices, W_Q, W_K, and $W_{V,}$ respectively. The operating principle is to calculate the attention distribution of the current Q under all K, add weight to the distribution and Softmax, attach to V in the form of a score to obtain the attention calculation result, and then set multiple heads. Each head performs the same operation to make the model pay attention to different aspects of information, and finally, the attention calculation results of each head are spliced together and multiplied with the output weight to obtain the final multi-head attention mechanism result. When Q, K, and V are transformed from the same input, the multi-head attention mechanism at this time is called the multi-head self-attention mechanism, and when their inputs are different, it is the multi-head cross-attention mechanism. FNN is a simple neural network containing an input layer, an intermediate layer, and an output layer. Using the GELU activation function, the specific formula is shown in Equation (5), where $\varnothing(x)$ refers to the cumulative function of the Gaussian normal distribution of x.

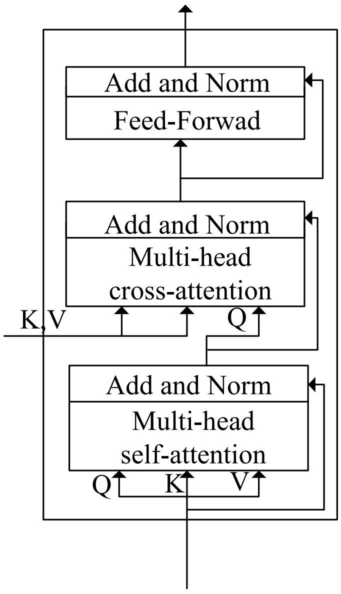

Figure 2. Transformer decoder structure.

The module represents H_{ate} as input using the features of the generated aspect entity sequence Y_a, and in the first layer, the multi-head self-attention mechanism is used, and H_{ate} is operated as Q, K, and V. In the second layer, the bull cross-attention is used, H_{share} as K and V, and the output H_1 of the first layer as Q input into it. Finally, the output H_2 of the second layer is input into the feedforward neural network layer, and finally, a preliminary result S is obtained. Each layer in the transformer decoder is followed by a residual connection [21] and layer normalization [22] to process the result. The specific formula is shown in Equations (6)–(8), where LN denotes layer normalization. The main function of this module is to realize the interaction between aspect items, consider the relationship between aspect items, use the previously obtained aspect items to extract the results, and combine the context information of the text to obtain a preliminary prediction result.

In the second transformer decoder module, the output of the previous module is taken as input, and the rest remains unchanged. This module uses preliminary prediction information, then combines the context and sentiment of the text to produce more accurate results.

$$GELU(x) = x\varnothing(x) \tag{5}$$

$$H_1 = LN(H_{ate} + SelfAtten(H_{ate}, H_{ate}, H_{ate})) \tag{6}$$

$$H_2 = LN(H_1 + CrossAtten(H_1, H_{share}, H_{share})) \tag{7}$$

$$S = LN(H_2 + FNN(H_2)) \tag{8}$$

3.5. Maintaining Sentiment Consistency

The issue of sentiment inconsistencies for aspect items arises from end-to-end modeling using sequence annotation. In order to deal with this problem, this paper constructs a consistent sentiment component with a gate mechanism; in this component, the input is a preliminary sentiment prediction result S for each word. Use Equation (10) to calculate the threshold value g_t, which represents the importance of the current moment prediction result, then, combine the prediction result of the previous moment through Equation (9) so that the prediction of the current moment will inherit the characteristics of the previous moment, to reduce the situation of different sentiment polarities in the same aspect item. The internal structure of this component is shown in Figure 3. In Equations (9) and (10), S_t represents the preliminary prediction result at the current moment, S'_{t-1} represents the final prediction result at the previous moment, W_g is the weight matrix, b_g is the bias term, and σ is the Sigmoid function.

$$S'_t = S_t g_t + S'_{t-1}(1 - g_t) \tag{9}$$

$$g_t = \sigma(W_g S_t + b_g) \tag{10}$$

3.6. Output Layer

The results from the sentiment consistency component are normalized by the Softmax classifier to obtain the probability of a uniform tag corresponding to each word, and the maximum probability corresponds to the final result of the word, which contains the results of two tasks: ATE and ASC. The calculation formula is shown in Equation (11), W is the weight matrix of the linear layer, and b is the bias term.

$$Y = soft\max(WS' + b) \tag{11}$$

3.7. Model Training

This model uses the cross-entropy loss function as the loss function to train the model. The formula of the cross-entropy loss function is shown in Equation (12), where N is the number of samples, M is the number of classes, y_{ij} is the sign function, if the true category of sample i is equal to j, take 1, otherwise take 0, and p_{ij} indicates the predicted probability that sample i belongs to class j, which is the Softmax value.

$$L = -\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log p_{ij} \tag{12}$$

In this study, all tasks are jointly trained, so the objective function is the sum of the loss functions of each task; that is, the objective function is equal to the sum of the loss function of the ATE and the loss function of the final unified prediction, as shown in Equation (13).

$$Loss = L_{ate} + L_{absa} \tag{13}$$

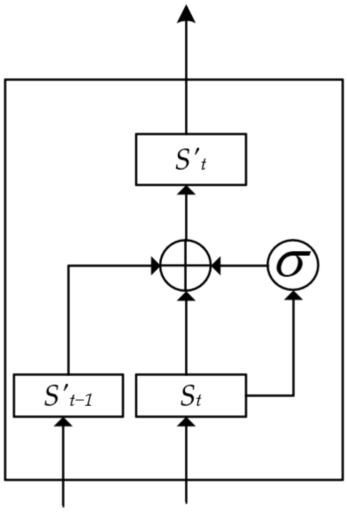

Figure 3. Sentiment consistency component structure diagram.

4. Experiment and Results Analysis

4.1. Dataset

In this paper, we evaluate our model on three widely used sentiment analysis datasets, namely Laptop14 [23], Restaurant14 [23], and Twitter [24] datasets; the statistical results of data information of each dataset are shown in Table 1.

Table 1. Experimental data statistics.

Datasets		**Train**	**Dev**	**Test**
Laptop14	POS	881	104	339
	NEG	754	106	130
	NEU	406	46	165
Restaurant14	POS	1956	213	728
	NEG	735	64	195
	NEU	575	52	197
Twitter	POS	549	69	73
	NEG	212	24	30
	NEU	1811	203	233

4.2. Model Parameters

The experiment uses the "BERT-base-uncased" model of the pre-trained model BERT, where the number of transformer layers L = 12, the embedding dimension d = 768, using the Adam optimizer, all dropout rates of the model are set to 0.1, use Precision, Recall, and F1 as metrics to evaluate the model, the calculation formula is shown in Equation (14). For each dataset, the batch size is set to 16, the learning rate is set to 3×10^{-5}, the number of training steps is set to 3000 steps, and a result is obtained every 100 steps, with the best result as the experimental result. For each dataset, set the number of layers *l* for different task sharing layers, Laptop14 set to 10, Restaurant14 to 11, and Twitter to 12.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{14}$$

4.3. Baseline Methods

We compare our model with the following models:

- LM-LSTM-CRF [25]: It is a language model-enhanced LSTM-CRF model, which achieved competitive results on several sequence-labeling tasks;
- E2E-TBSA [6]: Two stacked LSTMs were used to perform two tasks, target boundary detection and complete ABSA, respectively, and two auxiliary components were designed;
- DOER [13]: A double-cross shared RNN framework that jointly trains ATE and ASC for two tasks, considering the relationship between aspect and polarity;
- IMN [14]: An interactive multi-task learning model for the joint extraction of joint aspect items and opinion items, as well as ASC, and introduces a novel messaging mechanism that allows information interaction between tasks;
- BERT-E2E-ABSA [16]: Applying BERT to ABSA, they constructed a series of simple but effective neural baselines for this problem, using the best-performing BERT + GRU as a reference;
- SPAN [26]: A pipelined approach in which one model is used for ATE tasks, and then another model is used for ASC tasks;
- DREGCN [27]: An end-to-end interaction architecture based on multi-task learning relying on syntactic knowledge enhancement, the model uses well-designed dependency-embedding graph convolutional networks to make full use of syntactic knowledge and also designs a simple and effective messaging mechanism to realize multi-task learning;
- DCRAN [18]: A deeply contextualized relationship-aware network that allows implicit interaction between subtasks in a more efficient way and allows two explicit self-supervised strategies for deep context and relationship-aware learning.

4.4. Experimental Results

The comparison results between the model and each model in this paper are shown in Table 2. Glove indicates that the model is based on the Glove [28] word vector model, and BERT indicates that the model is based on the "BERT-base-uncased" model of the pre-trained model BERT. P stands for Precision, and R stands for Recall.

Table 2. Compare experimental results (%).

	Model	**Laptop14**			**Restaurant14**			**Twitter**		
		P	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
GLOVE	LM-LSTM-CRF	53.31	59.40	56.19	68.46	64.43	66.38	43.52	52.01	47.35
	E2E-TBSA	61.27	54.89	57.90	68.64	71.01	66.60	53.08	43.56	48.01
	IMN	-	-	57.66	-	-	68.32	-	-	51.31
	DOER	61.43	59.31	60.35	80.32	66.54	72.78	55.54	47.79	51.37
BERT	BERT-E2E	61.88	60.47	61.12	72.92	76.72	74.72	57.63	54.47	55.94
	SPAN	66.19	58.68	62.21	71.22	71.91	71.57	60.92	52.24	56.21
	DREGCN	-	-	63.04	-	-	72.60	-	-	-
	DCRAN	-	-	65.18	-	-	75.77	-	-	-
	Our method	67.73	63.56	65.58	76.92	77.05	76.98	62.22	60.66	61.43

Through the comparative analysis in Table 2, it can be seen that the effect of the Glove-based model is significantly lower than that of the BERT model, indicating that the BERT model is better than the Glove in this task.

On the Laptop14 dataset, our model is the best on all three metrics and 0.4% higher than the best DCRAN on F1. On the Restaurant14 dataset, our model has the best results on the Recall and F1 metrics, with F1 values 1.21% higher than the best DCRAN and DOER the best on Precision, but its Recall is low and lower than most models, indicating that their model can predict a small number of correct ones. On the Twitter dataset, our model is the best on three indicators, especially on Recall, which is 6.19% higher than the best BERT-E2E and 5.22% higher than the best SPAN on F1 values, which indicates that our model has good predictive performance, not only predicting more but also correctly.

Compared with DREGCN, using dependent syntactic knowledge enhancement, the results of this model are greatly improved, which shows that we can make good use of the context information and dependencies of text using the multi-head attention mechanism.

In addition, we have also completed experiments on the Restaurant15 dataset; the results are in line with our expectations, and all indicators are higher than these models.

4.5. Ablation Study

In order to verify the effectiveness of the interaction between the aspects and the components that maintain sentiment consistency, an ablation experiment is performed on three datasets with consistent settings.

For verifying the validity of the interaction between aspect items, the redundant BERT layer is no longer used to extract the aspect items, and the final output of BERT is directly input into the transformer decoder module, and a transformer decoder module is reduced; only the interaction between subtasks is performed, and the objective function becomes a loss function with only the final prediction result.

For verifying the validity of the component that maintains sentiment consistency, the component is removed, and the representation of the sentiment features that come out of the transformer decoder module is classified directly through the softmax classification layer.

The results of the ablation experiment are shown in Table 3, and it can be seen from the table that after removing the interaction of the aspect items, the proposed model performance has a certain degree of degradation on the three datasets, with F1 decreasing by 2.21, 1.72, and 0.72%. After removing the sentiment consistency component, the degradation in model performance was small, with F1 decreasing by 1.68, 0.86, and 0.31%. From these two experiments, it can be shown that the interaction between the aspect items is effective with the emotionally consistent components. The term w/o means to remove the module, ATI represents the interaction between aspect items, and MSC represents the sentiment consistency component.

Table 3. Results of ablation experiments (%).

Model	**Laptop14**			**Restaurant14**			**Twitter**		
	P	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
Full model	67.73	63.56	65.58	76.92	77.05	76.98	62.22	60.66	61.43
w/o ATI	65.01	61.83	63.37	75.00	75.54	75.26	61.43	60.01	60.71
w/o MSC	65.78	62.15	63.90	76.53	75.71	76.12	61.95	60.31	61.12

4.6. The Number of Task-Sharing Layers

In order to study the optimal value of the number of shared layers l when BERT is used as the task-sharing layer, this paper conducts a comparative experiment. The experiment is carried out on three datasets using the BERT model, except that the number of task-sharing layers l is inconsistent, and other parameter settings are consistent. In this paper, the number of layers shared by tasks is experimented on from 1 to 12, the performance of the model under the number of layers shared by different tasks is tested, and the optimal value is determined by comparing the F1 values. The results are shown in Table 4.

Table 4. The performance comparison results of different shared layer numbers (F1%).

l	Laptop14	Restaurant14	Twitter
1	65.30	75.41	60.74
2	64.45	76.21	60.67
3	64.75	76.08	60.25
4	64.99	75.92	61.13
5	63.96	75.78	60.64
6	65.00	75.94	60.13
7	65.20	76.36	60.50
8	63.82	76.01	60.93
9	64.78	76.19	61.11
10	65.58	75.44	61.43
11	64.89	76.98	60.63
12	65.28	76.03	60.89

As can be seen from Table 4, the optimal number of task-sharing layers *l* is not the same for different datasets. On the Laptop14 dataset, the performance is best at layer 10; on Restaurant14, the performance is best at layer 11; and on the Twitter dataset, the performance is best at layer 10. From this experiment, it can be concluded that the output of the later layers of BERT contains more information features required for this task, but it is not that the higher the number of layers, the better the performance. However, this optimal number of shared layers is susceptible to various influences and cannot find a common value.

5. Conclusions

In this paper, an interactive learning network that maintains sentiment consistency is proposed to handle sentiment analysis tasks at the end-to-end aspect based. This paper uses the task-sharing layer and multi-head attention mechanism to realize the interaction of the two subtasks of aspect item extraction and aspect sentiment classification, consider the relationship between aspect items, and use the context information of the whole text. Auxiliary components are also built using gate mechanisms to maintain sentiment alignment. The experimental results show that our model has better performance than previous studies, especially on the Twitter dataset, which is a big improvement. Future work could add auxiliary modules to the model, such as modules to improve the performance of extracting aspects in intermediate processes or build a more efficient interaction module.

Author Contributions: Conceptualization, M.C. and J.W.; Methodology, Q.H. and Y.M.; Software, Q.H.; Validation, Q.H.; Investigation, Y.M.; Resources, J.W.; Writing—original draft, Q.H.; Project administration, M.C.; Funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Scientific research project of Jiangxi Provincial Department of Education [GJJ200839] and the Doctoral startup fund of Jiangxi University of Technology [205200100402].

Data Availability Statement: In this paper, we evaluate our model on three widely used sentiment analysis open datasets, namely Laptop14, Restaurant14, and Twitter datasets which are detailed introduced in Refs. [23,24] respectively.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Thakur, N. Sentiment Analysis and Text Analysis of the Public Discourse on Twitter about COVID-19 and MPox. *Big Data Cogn. Comput.* **2023**, *7*, 116. [CrossRef]
2. Fellnhofer, K. Positivity and higher alertness levels facilitate discovery: Longitudinal sentiment analysis of emotions on Twitter. *Technovation* **2023**, *122*, 102666. [CrossRef]
3. Truşcă, M.M.; Frasincar, F. Survey on aspect detection for aspect-based sentiment analysis. *Artif. Intell. Rev.* **2023**, *56*, 3797–3846. [CrossRef]

4. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 168–177.
5. Wang, F.; Lan, M.; Wang, W. Towards a One-Stop Solution to Both Aspect Extraction and Sentiment Analysis Tasks with Neural Multi-Task Learning. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–8.
6. Li, X.; Bing, L.; Li, P.; Lam, W. A unified model for opinion target extraction and target sentiment prediction. In Proceedings of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33.
7. Peng, H.; Xu, L.; Bing, L.; Huang, F.; Lu, W.; Si, L. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 7–12 February 2020; Volume 34, pp. 8600–8607.
8. Chen, Z.; Qian, T. Relation-aware collaborative learning for unified aspect-based sentiment analysis. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Virtual, 5–10 July 2020; pp. 3685–3694.
9. Mao, Y.; Shen, Y.; Yu, C.; Cai, L. A joint training dual-mrc framework for aspect based sentiment analysis. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 13543–13551.
10. Yin, Y.; Wei, F.; Dong, L.; Xu, K.; Zhang, M.; Zhou, M. Unsupervised word and dependency path embeddings for aspect term extraction. *arXiv* **2016**, arXiv:1605.07843.
11. Tang, D.; Qin, B.; Feng, X.; Liu, T. Effective LSTMs for target-dependent sentiment classification. *arXiv* **2015**, arXiv:1512.01100.
12. Zhang, M.; Zhang, Y.; Vo, D.T. Neural networks for open domain targeted sentiment. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015; pp. 612–621.
13. Luo, H.; Li, T.; Liu, B.; Zhang, J. DOER: Dual cross-shared RNN for aspect term-polarity co-extraction. *arXiv* **2019**, arXiv:1906.01794.
14. He, R.; Lee, W.S.; Ng, H.T.; Dahlmeier, D. An interactive multi-task learning network for end-to-end aspect-based sentiment analysis. *arXiv* **2019**, arXiv:1906.06906.
15. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
16. Li, X.; Bing, L.; Zhang, W.; Lam, W. Exploiting BERT for end-to-end aspect-based sentiment analysis. *arXiv* **2019**, arXiv:1910.00883.
17. Luo, H.; Ji, L.; Li, T.; Duan, N.; Jiang, D. GRACE: Gradient harmonized and cascaded labeling for aspect-based sentiment analysis. *arXiv* **2020**, arXiv:2009.10557.
18. Oh, S.; Lee, D.; Whang, T.; Park, I.N.; Seo, G.; Kim, E.; Kim, H. Deep context-and relation-aware learning for aspect-based sentiment analysis. *arXiv* **2021**, arXiv:2106.03806.
19. Jawahar, G.; Sagot, B.; Seddah, D. What does BERT learn about the structure of language? In Proceedings of the ACL 2019–57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July 2019.
20. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inform. Process. Syst.* **2017**, *30*.
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016.
22. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *arXiv* **2016**, arXiv:1607.06450.
23. Pontiki, M.; Papageorgiou, H.; Galanis, D.; Androutsopoulos, I.; Pavlopoulos, J.; Manandhar, S. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. *SemEval* **2014**, *27*, 2014.
24. Mitchell, M.; Aguilar, J.; Wilson, T.; Benjamin, V.D. Open domain targeted sentiment. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, WA, USA, 18–21 October 2013.
25. Liu, L.; Shang, J.; Ren, X.; Xu, F.; Gui, H.; Peng, J.; Han, J. Empower sequence labeling with task-aware neural language model. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LO, USA, 2–7 February 2018; Volume 32.
26. Hu, M.; Peng, Y.; Huang, Z.; Li, D.; Lv, Y. Open-domain targeted sentiment analysis via span-based extraction and classification. *arXiv* **2019**, arXiv:1906.03820.
27. Liang, Y.; Meng, F.; Zhang, J.; Chen, Y.; Xu, J.; Zhou, J. A dependency syntactic knowledge augmented interactive architecture for end-to-end aspect-based sentiment analysis. *Neurocomputing* **2021**, *454*, 291–302. [CrossRef]
28. Pennington, J.; Socher, R.; Christopher, D.M. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.