Aspect-Based Sentiment Analysis using Local Context Focus Mechanism with DeBERTa

Tianyu Zhao  
Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia  
Beijing University of Posts and Telecommunications  
Beijing, China  
zhaoty1021@bupt.edu.cn

Junping Du*  
Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia  
Beijing University of Posts and Telecommunications  
Beijing, China  
junpingdu@126.com

Yingxia Shao  
Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia  
Beijing University of Posts and Telecommunications  
Beijing, China  
shaoyx@bupt.edu.cn

Ang Li  
Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia  
Beijing University of Posts and Telecommunications  
Beijing, China  
david.lee@bupt.edu.cn

Abstract—Text sentiment analysis, often termed as opinion mining, delves into quantifying individuals’ opinions, evaluations, attitudes, and emotions conveyed about entities. Sentiment analysis of text can be categorized into text-level, sentence-level, and aspect-level analyses. Aspect-Based Sentiment Analysis (ABSA) represents a detailed sub-discipline within sentiment analysis, with its primary goal being to ascertain the sentiment polarity of specific aspects. The research of pre-training neural models has significantly improved the performance of many natural language processing tasks. In recent years, pre-training model (PTM) has been applied in ABSA. Therefore, there has been a question, which is whether PTMs contain sufficient syntactic information for ABSA. In this paper, we explored the recent DeBERTa model (Decoding-enhanced BERT with disentangled attention) to solve Aspect-Based Sentiment Analysis problem. DeBERTa is a kind of neural language model based on transformer, which uses self-supervised learning to pre-train on a large number of original text corpora. Based on the Local Context Focus (LCF) mechanism, by integrating DeBERTa model, we propose a multi-task learning model for aspect-based sentiment analysis. The experiments result on the most commonly used laptop and restaurant datasets of SemEval-2014 and the ACL twitter dataset show that LCF mechanism with DeBERTa has significant improvement.

Keywords—Aspect-Based Sentiment Analysis, DeBERTa, Local Context Focus

I. INTRODUCTION

Sentiment Analysis (SA) is a vibrant domain within Natural Language Processing that delves into the perspective conveyed in text [1-4]. With the development of science and technology, more and more people like to express opinions and comments on products after consumption, which has formed a large number of comment texts [5-9]. There are lots of entity attributes in the comment texts. Aspect-Based Sentiment Analysis (ABSA) [10] aims to do the fine-grained sentiment analysis towards aspects, which is the sub-task of Sentiment Analysis. Traditional methods explore the general emotions of texts, but ABSA is a more challenging task because opinions can contain several aspects. Take the sentence “Its size is ideal and the weight is acceptable” for example, the customer had a positive sentiment for both size and weight. The task is to predict the emotions towards the underlined aspects. Typically, ABSA’s core research directions are on pinpointing different elements related to aspect-level sentiment. These elements include aspect terms, aspect categories, opinion terms, and sentiment orientations[11]. ABSA encompasses four primary tasks: namely, Aspect Term Extraction (ATE), Aspect Category Detection (ACD), Opinion Term Extraction (OTE), and Aspect Sentiment Classification (ASC). The ASC task aims to determine the sentiment direction associated with a specific aspect in a sentence. In this paper, we only focus on ASC tasks. Initial ASC tasks often relied on handcrafted features like term frequency[12]. Recently, deep learning-driven ASC tasks have garnered significant interest. BERT-centric deep language models have become prevalent in ABSA. [13][14]. BERT utilizes the recent advancements in deep bidirectional encoder representations from transformers, analyzing text by accounting for the context both to the left and right of a word across all layers[15]. BERT can produce richer textual semantic representations, where each word is translated into an embedding vector that is influenced by its context within the sentence. Pre-training methods (PTMs) such as BERT have brought significant performance improvements of the ASC task. By extending the training duration, using larger batches, and leveraging more data, RoBERTa (A Robustly Optimized BERT Pre-training Approach) has the potential to outperform the conventional BERT technique [16]. In 2020, He et al. [17] unveiled a novel structure for training BERT-based language models named Decoding-enhanced BERT with Disentangled Attention (DeBERTa). Two innovative techniques are employed to enhance the prior best-performing PLMs: a disentangled attention approach and a bolstered mask decoder.

In this paper, we present a method for Aspect-Based Sentiment Analysis using Local Context Focus (LCF)
Mechanism with DeBERTa. LCF mechanism combines local context features and global context features to predict sentiment polarity of targeted aspect [18]. Through this approach, the model can identify previously undetected aspects and place greater emphasis on local context words related to a specific aspect. It is very important for sentiment analysis based on domain specific aspects. Because the words representing aspects and sentiments are position dependent in the viewpoint text, they are usually close to each other. LCF can compute local context features. By further studying the application of PTMs in ABSA tasks, we made fine-tuning and mechanism enhancement on it. Therefore, the main contributions of this paper are as follows:

- This paper adds the latest PTM to the LCF design model for the first time. DeBERTa improves previous state-of-the-art PLMs using two novel techniques: a disentangled attention mechanism, and an enhanced mask decoder. This significantly improves the performance of LCF design.
- On the basis of preprocessing, a new mechanism is added to enhance the relationship between local context features and global context features. Contextual features make the model better predict the polarity in aspects of the target.
- Through the improvement and adjustment of the model, the performance of the ASC task is significantly improved. We conducted experiments with three datasets, which is the laptop and restaurant datasets of SemEval-2014 and the ACL twitter dataset. The experimental results show that the model is enhanced to some extent on three datasets, especially the Restaurant dataset.

1 Related Works

ABSA tasks include four tasks. ASC task is designed to predict the sentiment polarity for a particular aspect within a sentence. Because ASC task is the focus of this paper. And in the method, we use the PTMs. Accordingly, we separate our discussion of related work into two areas: Firstly, methods and related research on ASC tasks in recent years. Secondly, development and application of PTM in Aspect-Based Sentiment Analysis tasks.

1.1 Aspect Sentimental Classification

Aspect Sentiment Classification is a crucial subtask within ABSA. Typically, aspects can manifest as either aspect terms or aspect categories. This distinction gives rise to two ASC challenges: sentiment classification based on aspect terms and sentiment classification rooted in aspect categories. In fact, the main research issues of these two subtasks are the same. What they explore is how to use aspects and context to classify sentiments at aspect level.

Lately, deep learning-driven ASC has captured widespread attention. Numerous neural network-based ASC models [19-23] have been introduced, leading to significant performance enhancements. [24-28]. Models like TC-LSTM [29] were developed to capture the interplay between the aspect and the surrounding sentence context. Different segments of a sentence play varied roles concerning specific aspects; thus, the attention mechanism is commonly employed to derive representations for particular aspects [30]. As a representative work, Wang et al. [31] purposed Attention-based LSTM with Aspect Embedding (ATAE-LSTM) model. The research appended aspect embeddings to every word vector within the input sentence to determine the attention weight. This approach enables the computation of aspect-specific sentence embeddings, which are then utilized for sentiment classification. Subsequently, designs for intricate attention mechanisms were introduced with the intent to learn more refined aspect-specific representations. Ma et al. proposed the IAN, which separately produces representations for both aspect and sentence attentions interactively[32]. Besides, other network structures like the gated network [33] also had a good effect. Recently, the development of preprocessing models has greatly improved the performance of tasks. For instance, Sun et al. [34] utilize the ability of BERT by transforming the task as a pair classification problem. There is another method of the ASC research modeling the syntactic structure of the sentence to infer the polarity of the sentient aspects. As the improvement of dependency analysis based on neural network [35][36] in recent years, better parse tree brought significant improvements to the dependent ASC models. Sun et al. [37] and Zhang et al. [38] utilized word dependencies and the syntactical information to model the dependency tree. What they employ is the graph neural network (GNN) [39][40].

In contrast to text-level and sentence-level sentiment analysis, aspect-level sentiment analysis encounters unique challenges. Aspect-level sentiment analysis technology not only analyzes the explicit language expression structure, but also deeply understands the implicit semantic expression. Besides, Aspect-level sentiment analysis needs to determine the context range in which sentiment is expressed for each evaluation aspect. Aspect-level sentiment analysis technology needs to correctly understand the semantic information of the text word-level and sentence-level.

1.2 Pre-Training Models

Pre-trained word representation models, like Word2Vec [41] and GloVe [42] were used in the conventional neural ABSA models, coupling with the task-specific neural architecture. When contrasted with earlier feature-based models, they exhibit a certain degree of efficacy. However, due to their context-independent word embeddings, they struggle to grasp the intricate sentiment dependencies within sentences. Lately, pre-trained language models like BERT [43] and RoBERTa [44] have ushered in notable advancements in NLP endeavors. Thanks to the extensive knowledge acquired during the pre-training stage, merely utilizing these contextualized embeddings has resulted in significant performance enhancements. For example, Li et al. [45] tried using several stacked standard prediction layers on PTM for the E2E-ABSA tasks. On the basis of the original Bert model, Through experiments, RoBERTa has demonstrated that augmenting the volume of pre-training data enhances the model's effectiveness. Similarly, prolonging the pre-training duration or amplifying the count of pre-training iterations also bolsters the model's performance.

For PTMs, most of the current mainstream models use Transformer as a feature extractor. At this stage, the potential of Transformer has not been fully tapped, and there is still a
lot of potential to tap. Not only do PTMs serve as the foundation for the ABSA model, but they also offer advantages when managing ABSA tasks in various manners. For instance, the language modeling tasks employed during the PTM’s pre-training phase often equip the model with capabilities for generative data augmentation. Li et al. [46] harnessed PTM to realize generative, semantic-preserving augmentation, securing noticeable advancements over foundational techniques across multiple ABSA tasks. At present, it is generally believed in the NLP community that PTMs can accurately reflect the semantics of input words [47]. Yet, the contextual embeddings derived from the self-attention mechanism, which grasp the complete word interdependencies in a sentence, might be somewhat superfluous for ABSA tasks. Because ABSA often doesn’t need to capture as many dependencies, doing so is wasteful. The pursuit of merging significant sparse relational structures with PLM, refining the inherent comprehensive self-attention, and deriving ABSA-pertinent representations more efficiently warrants greater focus and research efforts.

2 DeBERTa-LCF

In this section, we will outline our method for the Aspect-Target Sentiment Classification task, which is executed in a two-step process. In the first step, we use the pre-trained model DeBERTa as a basis. We first briefly introduce the model structure of BERT and DeBERTa. In the second step, BERT shared layer is adopted as the embedding layer and feature extractor layer. And integrate it into the local context focusing mechanism. The main framework of DeBERTa-LCF model is shown in Fig. 3.

![Fig. 1. The main framework of DeBERTa-LCF model.](image)

### 2.1 BERT and DeBERTa

BERT and DeBERTa both take Transformers as backbone architecture. The disentangled attention introduced by DeBERTa suggests distinguishing between the content and the position of the text. The central concept of DeBERTa revolves around learning attention weights for each of these components. This sets it apart from other suggestions that add the position vector directly to the content vector[48]. Such a distinct division enables the model to more effectively differentiate between the positional and content elements of the data. While positional embeddings generate syntactic features, content embeddings handle the semantic aspects.

$$A_{\phi} = \{H, \Pi_{i|j}\} \cdot \{H_\phi, \Pi_{i|j}\}^T$$

Equation 1 defines the cross-attention matrix used in DeBERTa. In models based on BERT, tokens are depicted using both content and position vectors. The existing relative position coding methods use a separate embedding matrix to calculate the relative position deviation when calculating the attention weight [49]. It is similar to calculating the attention weight using only the content to content and position term in equation 1.

In ABSA tasks, the goal is to determine the sentiment of each aspect of the entity. We employed the pre-trained parameters from DeBERTa to initialize our model for subsequent tasks. Further, we implemented fine-tuning using ABSA labeled data to refine the model parameters. He et al. [50] suggested that the disentangled attention mechanism can be integrated into the BERT model without significant alterations to the rest of the neural network architecture. We adhered to the conventional fine-tuning approach, as the disentangled attention mechanism inherently encapsulates these features. The disentangled mechanism already took into account the content and relative positions of the context words. So we do not do much fine-tuning on the DeBERTa model. DeBERTa only needs half the data and is better than BERT and RoBERTa. We use the standard softmax layer with categorical cross-entropy loss function, which is the output of the language model, to provides downstream sentiment classification tasks. The BERT-Shared layer is considered as embedding layers, and the fine-tuning is carried out individually based on the joint loss function derived from multi-task learning.

2.2 Local Context Focus

Semantic-relative distance (SRD) is purposed for helping models capture local context. SRD is a principle centered on token-aspect pairings, delineating the gap between a token and its aspect. Upon computing the tagged outputs from the attention layer, any output features at locations exceeding the SRD threshold are either obscured or diminished. Meanwhile, the output features of the local context words remain wholly intact.

We introduced two structures, CDM and CDW, to emphasize local contexts. Our model centers on the local context through the integration of a local context focus layer. The LCF design’s input sequence predominantly draws from the global context. The DeBERTa layer possesses the prowess to...
assimilate context features. Leveraging its self-attention mechanism, the Multi-Head Self-Attention (MHSA) executes various attention functions to gauge the attention score for every contextual word. By MHSA, the features of each code are more closely related to itself. Multiple self-attention mechanisms operate concurrently. The derived results are then processed, ensuring that the acquired information is holistic, facilitating the extraction of profound semantic features from the context. MHSA helps mitigate the adverse implications of long-range context dependencies during feature learning.

3 Experiments

3.1 Datasets

To evaluate our model, we used the SemEval-2014 and the ACL twitter dataset. There are three datasets for this task: the laptop, restaurants, and twitter reviews. Rest14 and Laptop14 are from SemEval 2014 task 4 [1] containing sentiment reviews from restaurant and laptop domains. Twitter is from Dong et al. [51] which is processed from tweets. In these datasets, users rate their experiences with laptops and restaurants and comments made on Twitter in various aspects. The polarity of each aspect on these datasets may be positive, neutral, and negative. Therefore, the datasets we used contain 12184 total samples and three sentiments (positive, negative, and neutral).

The statistics of these datasets are presented in Table 1.

| Table 1. Detail of benchmark datasets |
|--------------------------------------|
| Laptops  | Restaurants  | Twitter  |
| Positive | 994          | 2164     | 1561   |
| Negative | 870          | 807      | 1560   |
| Neutral  | 464          | 637      | 3127   |

3.2 Compared Models

In order to prove the excellent reliability of the DeBERTa-LCF model described in this paper, it was necessary to show that this model is superior to other models. We compared DeBERTa-LCF with the following state-of-the-art models: BERT-BASE [52] is the baseline of BERT-based models. ATAELSTM [53] is a classic LSTM based model, which uses attention-based LSTM to explore the relationship between an aspect and sentence content. GCAE [54] is a CNN based model which based on convolutional neural networks and gating mechanisms. BERT-ADA [55] is a model built upon BERT, tailored for domain adaptation. It refines the BERT-BASE model using a task-specific corpus. IAN [56] generates the representation of target aspect and context through two LSTM networks respectively, which learns the representation of target aspects and contexts interactively.

RAM [57] is a novel framework based on neural networks to identify the sentiment of opinion targets with a RNN for sentence representation.

Table 2. Experimental results of performance.

| Models        | Laptop Acc | Laptop F1 | Restaurant Acc | Restaurant F1 | Twitter Acc | Twitter F1 |
|---------------|------------|-----------|----------------|----------------|-------------|------------|
| BERT-BASE     | 78.52      | 75.5      | 82.12          | 74.53          | 73.4        | 70.22      |
| ATAE-LSTM     | 68.7       | 67.23     | 77.1           | 66.84          | 69.22       | 67.4       |
| GCAE          | 78.05      | 69.93     | 78.31          | 68.74          | 71.42       | 69.3       |
| BERT-ADA      | 79.7       | 75.21     | 81.23          | 72.5           | 74.36       | 71.53      |
| IAN           | 72.14      | 70.6      | 78.64          | 70.12          | 70.3        | 67.6       |
| RAM           | 74.45      | 71.32     | 80.27          | 71.02          | 69.72       | 67.2       |
| DeBERTa-LCF*  | 79.54      | 75.64     | 83.46          | 75.4           | 70.83       | 68.64      |

The results with “*” are derived from our model. We highlight the best results on bold.

3.3 Analysis of Overall Performance

Table 2 presents a summary of the experimental outcomes, measured using ACC and F1 metrics for the Laptop, Restaurant and Twitter datasets. The experiments result on the most commonly used the laptop and restaurant datasets of SemEval-2014 and the ACL twitter dataset show that LCF mechanism with DeBERTa has been improved in different degrees, especially in restaurant datasets.

In Restaurant dataset, our proposed DeBERTa-LCF model achieves the best results in terms of both macro-F1 scores and accuracy scores. It gets an improvement of 1.34, 0.87.

In Laptop dataset, our model has also been improved to some extent on macro-F1 scores. And its accuracy scores are only slightly worse than the base-lines.

In Twitter dataset, our model is second only to the best one at present, and the difference is very slight. Comparing with BERT-BASE models indicates that the local context focus mechanism is adept at identifying undiscovered aspects and determining sentiment polarity.

4 Conclusion

In this paper, we propose a DeBERTa model with LCF mechanism for aspect-based sentiment classification tasks. We introduce LCF mechanism, which is of great significance for aspect item extraction. LCF designs focus on the local context and learn global context representations in parallel. At the same time, we introduce the DeBERTa model, which is the latest pre-training model. This greatly increased the performance of the model. We applied the DeBERTa model and integrate it and LCF mechanisms for the first time. We conduct a set of experiments on three datasets. The results prove that our model achieves certain performance.

References

[1] Schouten, K., Frasincar, F.: Survey on aspect-level sentiment analysis. IEEE Transactions on Knowledge and Data Engineering, 28(3), pp. 813-830 (2015).
[2] Yawen Li, Di Jiang, Rongzhong Lian, Xueyang Wu, Conghui Tan, Yi Xu, and Zhiyang Su. Heterogeneous latent topic discovery for semantic
text mining"[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(1): 533-544

[3] Xunpu Yuan, Yawen Li, Zhe Xue, Feifei Kou. Financial sentiment analysis based on pre-training and textcn. Chinese Intelligent Systems Conference, CISC, 2020, 48-56.

[4] Feifei Kou, Junping Du, Congxian Yang, Yansong Shi, Wanqiu Cui, Meiyu Liang, and Yue Geng[J]. HashTag recommendation based on multi-features of microblogs. Journal of Computer Science and Technology (2018): 711-726

[5] Ang Li, Junping Du, Feifei Kou, Zhe Xue, Xin Xu, Mingying Xu, Yang Jiang: Scientific and Technological Information Oriented Linguistic-adversarial and Media-adversarial Cross-media Retrieval[J]. arXiv preprint arXiv:2203.08615, 2022

[6] Xinlei Wei, Junping Du, Meiyu Liang, and Lingfei Ye. Boosting deep learning via support vector regression for fast moving crowd counting[J]. Pattern Recognition Letters, 2019, 119: 12-23

[7] Ang Li, Yawen Li, Yingxia Shao and Bingyan Liu. Multi-View Scholar Clustering With Dynamic Interest Tracking"[J]. IEEE Transactions on Knowledge and Data Engineering, 2023: 1-14

[8] Shitao Xiao, Yingxia Shao, Yawen Li, Hongzhi Yin, Yanyan Shen, Bin Cui. LECF: recommendation via learnable edge collaborative filtering[J]. Science China Information Sciences, 65(1):1-15, 2022.

[9] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S. Semeval-2014 task 4: SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 27–35. Association for Computational Linguistics, Dublin, Ireland (2014).

[10] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Mohammad, A.S., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., et al: Semeval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), pp. 19–30. Association for Computational Linguistics, San Diego, CA, USA (2016).

[11] Zhang, W.X., Deng, Y., Li, X., Yuan, Y.F., Bing, L.D., Lam, W.: Aspect sentiment quad prediction as paraphrase generation. EMNLP 2021 Main Conference. pp. 9209–9219 (2021).

[12] Caroline, B., Diana, N.P., Claude, R.: XRCE: Hybrid Classification for Aspect-based Sentiment Analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 838–842. Association for Computational Linguistics, Dublin, Ireland (2014).

[13] Song, Y., Wang, J., Jiang, T., Liu, Z., Rao, Y.: Targeted Sentiment Classification with Attentional Encoder Network. In: Tetko, I., Kůrková, V., Karpov, P., Theis, F. (eds) Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series. ICANN 2019. LNCS, vol. 11730, pp. 93–103. Springer, Cham (2019).

[14] Rietzler, A., Stabinger, S., Opitz, P., Engl, S. Adapt or get left behind: Domain adaptation through bert language model finetuning for aspect-target sentiment classification. In: Proceedings of The 12th Language Resources and Evaluation Conference, pp. 4933–4941. arXiv, Marseille (2020).

[15] Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. In Conference of the North American Chapter of the Association for Computational Linguistics, pp. 4171–4186. Association for Computational Linguistics, New Orleans, Louisiana (2019).

[16] Liu, Y.H., Myle, O., Naman, G., Du, J.F., Mandar, J., Chen, D.Q., Omer, L., Mike, L., Luke, Z., and Veselin, S.: RobERTa: A Robustly Optimized BERT Pretraining Approach. CoRR, abs/1907.11692 (2019).

[17] He, P., Liu, X., Gao, J., and Chen, W. DeBERTa: Decoding-enhanced BERT with Disentangled Attention. In: International Conference on Learning Representations 2021. Vienna, Austria (2021).

[18] Zeng, B., Yang, H., Xu, R., Zhou, W., Han, X. LCF: A local context focus mechanism for aspect-based sentiment classification. Applied Sciences 9, 3389 (2019).

[19] Li, Wenling, Yingmin Jia, and Junping Du. Resilient filtering for nonlinear complex networks with multiplicative noise. IEEE Transactions on Automatic Control, 64(6): 2522-2528, 2018

[20] Jian Cao, Dianhui Mao, Qiang Cai, Haisheng Li, and Junping Du. A review of object representation based on local features. Journal of Zhejiang University SCIENCE C, 14(7): 495-504, 2013.

[21] Meng, Deyuan, Yingmin Jia, Junping Du, and Jun Zhang. High-precision formation control of nonlinear multi-agent systems with switching topologies: A learning approach. International journal of robust and nonlinear control, 25(33): 1993-2018, 2015.

[22] Yingxia Shao, Shiyue Huang, Yawen Li, Xupeng Miao, Bin Cui, Lei Chen. Memory-aware framework for fast and scalable second-order random walk over billion-edge natural graphs[J]. The VLDB Journal, 30(5), 769-797, 2021

[23] Yawen Li, Ye Yuan, Yishu Wang, Xiaolan Li, Yuangia Ma, Guoren Wang. Distributed Multimodal Path Queries[J]. IEEE Transactions on Knowledge and Data Engineering, 34(7):3196-3211, 2022

[24] Vo, D., Yue, Z.: Target-dependent twitter sentiment classification with rich automatic features. In: International Conference on Artificial Intelligence, pp. 1347–1353. AAAI Press. Buenos Aires Argentina (2015).

[25] Wenling Li, Yingmin Jia, and Junping Du. Tobit Kalman filter with time-correlated multiplicative measurement noise. IET Control Theory & Applications, 11(1): 122-128, 2017

[26] Deyuan Meng, Yingmin Jia, and Junping Du. Robust iterative learning protocols for finite-time consensus of multi-agent systems with interval uncertain topologies. International Journal of Systems Science, 46(5): 857-871, 2015

[27] Li, Wenling, Jian Sun, Yingmin Jia, Junping Du, and Xiaoayan Fu. Variance-constrained state estimation for nonlinear complex networks with uncertain coupling strength. Digital Signal Processing, 67: 107-115, 2017.

[28] Yawen Li, Isabella Yunfei Zeng, Ziheng Niu, Jiahao Shi, Ziyang Wan and Zeli Guan, Predicting vehicle fuel consumption based on multi-view deep neural network, Neurocomputing, 502:140-147, 2022

[29] Tang, D.Y., Qin, B., Feng, X.C., Liu, T.: Effective LSTMs for Target-Dependent Sentiment Classification. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 3298–3307. The COLING 2016 Organizing Committee, Osaka, Japan (2016).

[30] Tay, Y., Loo, A.T., Hui, S.C.: Learning to Attend via Word-Aspect Associative Fusion for Aspect-based Sentiment Analysis. In The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), pp. 5956–5963. Association for the Advancement of Artificial Intelligence, New Orleans, Louisiana, USA (2018).

[31] Wang, Y.Q., Huang, M.L., Zhao, L., Zhu, X.Y.: Attention-based LSTM for Aspect-level Sentiment Classification. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 606–615. Association for Computational Linguistics, Austin, Texas (2016).

[32] Ma, D., Li, S., Zhang, X., & Wang, H., “Interactive attention networks for aspect-level sentiment classification,” in Twenty-Sixth International Joint Conference on Artificial Intelligence, pp. 4068–4074. International Joint Conferences on Artificial Intelligence, Melbourne, Australia (2017).

[33] Wei, X., Tao, L.: Aspect based sentiment analysis with gated convolutional networks. In Meeting of the Association for Computational Linguistics, pp. 2514–2523. Association for Computational Linguistics, Melbourne, Australia (2018).

[34] Sun, C., Huang, L., Qi, X.: Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In Human Language Technology: Conference of the North American Chapter of the
