MASS: Mitigating Aspect-oriented Semantic Sparsity for Fine-grained Sentiment Analysis

Yanjiang Chen[1], Kai Zhang[1(✉)], Zhe Yang[1], Linan Yue[1], Xinjie Sun[1], Kun Zhang[2], and Qi Liu[1]

[1] State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China, Hefei, China
[2] School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
{yjchen, yz01, lnyue, xinjiesun}@mail.ustc.edu.cn
{kkzhang08, qiliuql}@ustc.edu.cn, zhkun@hfut.edu.cn

Abstract. Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment classification task. Despite significant improvements in this field, progress is hindered by challenges such as the sparsity of contexts for specific aspects, interference from irrelevant words within sentences, and a lack of research on leveraging correlations between samples. In order to address these issues, we propose a novel framework MASS, Mitigating Aspect-oriented Semantic Sparsity for ABSA. Firstly, we employ advanced prompting techniques with Large Language Models (LLMs) to generate nuanced aspect-specific descriptions, thereby enhancing contexts related to the aspect. Subsequently, we design a novel fusion module aimed at seamlessly integrating aspectual insights with the original sentence. Finally, we develop three pioneering contrastive learning strategies to explore and learn complex correlations between samples, which are crucial for fine-grained sentiment analysis. Experiments on six benchmark datasets demonstrate that our MASS substantially outperforms state-of-the-art techniques and provides valuable insights for applying LLMs to downstream tasks.

Keywords: Aspect-based sentiment analysis · Large language models · Contrastive learning.

1 Introduction

Aspect-based sentiment analysis (ABSA) is a popular branch of sentiment analysis which aims to recognize the sentiment polarity of the specific aspect in a sentence. For example, given the sentence "*The restaurant has delicious food, but the atmosphere there is really noisy*", ABSA task aims to predict the sentiment polarity of the two aspects "*food*" and "*atmosphere*", which should be positive and negative, respectively. Recent research [19, 20] exploits attention mechanisms to model the relationships between the aspect and contexts. In addition, various methods [17, 3, 9] leverage syntactic information by processing dependency relationships with graph neural networks.

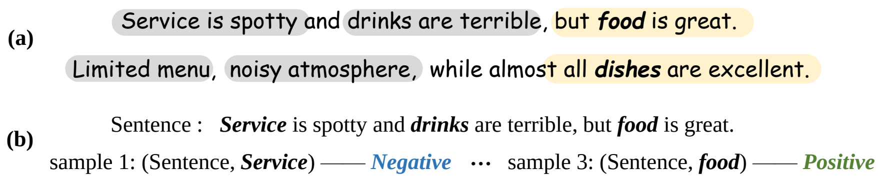

Fig. 1. (a) two example samples with multiple aspects, contexts related to "*food*" and "*dishes*" are marked in yellow, irrelevant contexts are marked in grey; (b) There are three samples, which contain the same sentence but different aspects and labels.

Despite the advancements made by previous approaches, these methods still struggle with sentences that contain multiple aspects and opinions. Taking the example in Fig 1 (a) for illustration, for the aspect "food", the only pertinent description is "but food is great". To reduce the influence of irrelevant contexts, existing methods utilize attention mechanisms and introduce syntactic information to enhance constraints. These methods merely focus on processing original sentences, ignoring to address data sparsity of specific aspects from the reverse perspective, such as directly enriching aspect-related contexts. In addition, few approaches attempt to leverage similarities and differences within different samples. Nevertheless, how to leverage the correlations between samples for ABSA presents a significant challenge. As illustrated in Fig 1 (b), these samples contain the same sentence but different aspects and labels, thus directly capturing the correlations between samples based on original sentences and labels can cause errors and be unsuitable for ABSA.

Recently, Large Language Models (LLMs) have demonstrated remarkable capabilities in numerous NLP fields, we consider that leveraging LLMs to generate aspect-related information for ABSA holds significant potential. Inspired by this, we propose a novel architecture MASS, Mitigating Aspect-oriented Semantic Sparsity for ABSA. Specifically, we first leverage cutting-edge prompt techniques with LLMs to generate aspect-related descriptions. Subsequently, we design a fusion module based on BERT [4] to fully leverage the generated aspect information while ensuring it does not significantly alter the original sentence semantics. Through the module, we obtain aspect-enhanced semantic features that contain semantics of the original sentence and aspect-oriented information. Finally, in order to utilize correlations between samples and thereby enhancing the performance for ABSA, we introduce three pioneering contrastive learning methods based on the semantic features.

2 Related Works

2.1 Aspect-based Sentiment Analysis

ABSA is a fine-grained sentiment analysis task that analyzes the specific aspect in the sentence, rather than simply allocating a general sentiment polarity at the document-level or sentence-level. With the development of deep learning, various works apply attention mechanisms to model the semantic relationship [19, 16]. Another research trend is to leverage syntactic knowledge to model syntactic dependency [17, 9]. More recently, many approaches armed with PLMs have achieved remarkable results. However, these methods do not address the issue of sparse contexts related to the specific aspect and fail to fully leverage the knowledge between samples.

2.2 Large Language Models

Large Language Models (LLMs) have exhibited excellent performance on various natural language understanding and generation tasks. Many studies prove that the performance of LLMs is significantly influenced by the In-context Learning (ICL) demonstrations [1, 10]. ICL is widely applied in various tasks, such as machine translation, relation extraction and data generation. In our approach, we thoroughly harness the semantic understanding and generative capabilities of LLMs, and then enrich aspect-related information within each sentence to address aspect-oriented data sparsity.

2.3 Contrastive Learning

Contrastive Learning (CL) has achieved remarkable performance in the field of NLP. The main goal of CL is to learn representations between samples by contrasting positive pairs and negative pairs [7]. In summary, CL can be divided into two major categories, unsupervised CL and supervised CL. Unsupervised CL attempts to contrast grouped instances to produce more robust representations of unlabeled data, and supervised CL is label-aware and aims to learn distinct representations for differently labeled data [6].

3 Methodology

Problem Definition. In ABSA, a sentence $s = \{w_1, w_2, ..., w_n\}$ and a specific aspect term $a = \{a_1, a_2, ..., a_m\}$ are given, and the goal of ABSA is to precisely predict the sentiment polarity C_a of the given aspect a in the sentence s.
Overall Architecture. As illustrated in Fig 2, MASS consists of three parts: 1) Aspect Information Generation; 2) Fusion and Re-weighting Module; 3) Semantics-based Contrastive Learning.

3.1 Aspect Information Generation

Data augmentation technique is widely used to solve data sparsity in many tasks. However, conventional data augmentation (i.e., Crop, Mask) may be unsuitable for ABSA because they probably delete aspect-related contexts or increase irrelevant words, which instead hurt prediction in sentiment polarity.

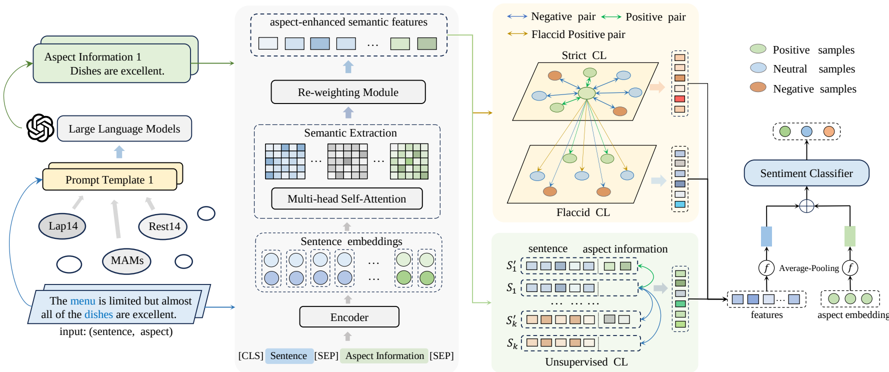

Fig. 2. The overall architecture of MASS. From left to right, the modules are Aspect Information Generation, Fusion and Re-weighting Module, Semantics-based Contrastive Learning and Sentiment Classifier.

As shown in Fig 3, we innovatively propose to generate aspect-related information through LLMs. Aspect information represents how the specific aspect is described in the original sentence. To obtain high quality aspect information, we explore a prompt template, which mainly consists of three parts: task description, output requirements, and input examples. Specifically, we first briefly introduce the core of the ABSA task and give the definition of "*aspect*", "*sentiment polarity*" and "*aspect information*" to LLMs. Then, we input the sentence-aspect pair (s, a) into LLMs and obtain aspect information t for each sentence s:

$$t = \text{LLM}\ (prompt; (s, a)), \tag{1}$$

where $t = \{t_1, t_2, ..., t_l\}$ and l is the length of the aspect information. Specifically, the aspect information generated by LLMs is diverse, depending on the complexity of the sentences and the LLMs' comprehension ability.

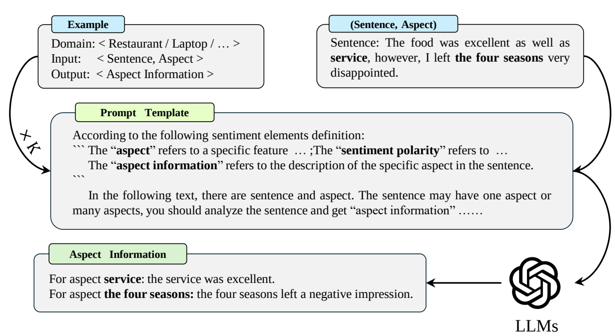

Fig. 3. The process of Aspect Information Generation through LLMs.

3.2 Fusion and Re-weighting Module

Encoder. We adopt BERT as the encoder to get contextual embeddings. We construct input as "$[CLS]$ s $[SEP]$ t $[SEP]$" to transform it into hidden state H, and H contains sub-sequence H_a, which is the representation of the aspect.
Semantics Extraction and Fusion. We first leverage attention mechanism on H to obtain the overall semantics of the new sentence, and then use a Feed-Forward Network (FFN) to facilitate learning nonlinear features:

$$H_m = \text{softmax}\left(\frac{(HW_h^q)(HW_h^k)^T}{\sqrt{d_k}}\right) HW_h^v, \tag{2}$$

$$H_f = \max(0,\ H_m W_1 + b_1) W_2 + b_2, \tag{3}$$

where W_h^q, W_h^k, W_h^v are learnable parameters in the attention mechanism, and W_1, b_1, W_2, b_2 are weight parameters in FFN.
Aspect-enhanced Semantic Features. Aspect information may change overall semantics of original sentence while its content is excessive. Thus, we re-weight aspect information based on its length, treating the length as a penalty. Moreover, we emphasize the importance of contexts that are near the aspect to reduce the noise that naturally arisen from attention weights. Formally, we design a function P applied to H_f for extracting aspect-enhanced semantic features z:

$$q_i = \begin{cases} 1 - \frac{r+1-i}{n} & 1 \le i < r+m, \\ 1 - \frac{i-r-m}{n} & r+m < i \le n, \\ \frac{m}{l} & n < i \le n+l, \end{cases} \tag{4}$$

$$z = P(H_f) = qH_f, \tag{5}$$

where q_i is the weight value for i-th token, r is the start position of the aspect, m is the length of aspect term and l is the length of aspect information.

3.3 Semantics-based Contrastive Learning

Suppose a batch which contains n samples $B = \{(S_1, C_1), (S_2, C_2), ..., (S_n, C_n)\}$, S_i is the i-th sample with label C_i, and $I = \{1, ..., n\}$ is the sample index set. Notably, to distinguish different features, the aspect-enhanced semantic features of S_i are denoted as z_i, while original sentence features are denoted as $z_i^{'}$.
Strict Label Contrastive Learning. Strict Label Contrast (SLC) is a supervised method which treats all samples of the same sentiment polarity as positive pairs and different ones as negative pairs. As a result, SLC makes features of samples with the same label closer than others. Given sample S_i in a batch, all other samples which have the same sentiment polarity as S_i make up the set $P(i) \equiv \{p \in I \setminus \{i\} : \boldsymbol{C}_p = \boldsymbol{C}_i\}$, and we define the set $O(i) \equiv I \setminus \{i\}$. The SLC function is defined as:

$$L_S = \sum_{i \in I} \frac{-1}{|P(i)|} \sum_{p \in P_{(i)}} \log \frac{e^{(z_i \cdot z_p / r)}}{\sum_{j \in O(i)} e^{(z_i \cdot z_j / r)}}. \tag{6}$$

Flaccid Label Contrastive Learning. Different from traditional categorization problems, sentiment analysis should consider sentiment intensity. Specifically, the distance of sentiment features between positive and neutral (or negative and neutral) samples tends to be closer compared to those positive and negative samples. We design Flaccid Label Contrast (FLC) to model this distance relationship. First, we construct matrix $M \in R^{n \times n}$. The elements of $M \in R^{n \times n}$ are initialized to zero, where m_{ij} represents the weight between sample i and sample j. The FLC function is denoted as:

$$m_{ij} = \begin{cases} \alpha & \text{if } C_i = C_j, \\ \beta & \text{if } C_i \neq C_j,\ C_i \text{ or } C_j \text{ is Neu,} \end{cases} \tag{7}$$

$$L_F = \sum_{i \in I} \frac{-1}{|I|} \sum_{j \in O_{(i)}} \log \frac{m_{ij} \cdot e^{(z_i \cdot z_j / r)}}{\sum_{k \in O(i)} e^{(z_i \cdot z_k / r)}}. \tag{8}$$

Unsupervised Contrastive Learning. Unsupervised Contrastive Learning (UCL) often obtains positive pairs by data augmentation. In our approach, we use features z_i and z_i' to construct positive pairs. Through this construction, UCL helps distill the LLM's understanding and knowledge. On the other hand, it can effectively prevent aspect information from excessively altering semantics of original sentences. In addition, z_i and z_k (features of other sentences) constitute negative pairs. The UCL function can be formulated as :

$$L_U = \sum_{i \in I} \frac{-1}{|I|} \log \frac{e^{(z_i \cdot z_i^{'} / r)}}{\sum_{k \in I} e^{(z_i \cdot z_k / r)} + e^{(z_i \cdot z_k' / r)}}. \tag{9}$$

3.4 Model Training

We obtain the final classification features of each sample by concatenating original aspect representation H_a and the aspect-enhanced semantic features z, then we map it to the probability distribution over the three sentiment polarities:

$$H_a^{final} = [\mathrm{Avg}(H_a), \mathrm{Avg}(z)],\ \hat{y} = \mathrm{softmax}(W_s H_a^{final} + b_s), \tag{10}$$

$$\mathcal{L} = -\sum_{i=1}^{|D|} \sum_{j=1}^{|C|} y_i^j \log \hat{y}_i^j + \lambda_1 L_U + \lambda_2 L_S + \lambda_3 L_F, \tag{11}$$

where y_i is the ground truth, D contains all training samples and C contains all sentiment polarities, and λ_1, λ_2, λ_3 are contrastive learning coefficients.

4 Experiments

4.1 Experimental Setup

Datasets. We conduct our experiments on six benchmark datasets, including Lap14 and Rest14 [14], Rest15 [13], Rest16 [12], Twitter [5] and MAMs [8]. All

Table 1. Experiment results (%) on six benchmark datasets.

Models	Lap14		Rest14		Rest15		Rest16		Twitter		MAMs	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
BERT-SPC [15]	78.99	75.03	84.46	76.98	83.03	63.92	90.75	74.00	73.41	72.38	82.82	81.90
R-GAT [18]	78.21	74.07	86.60	81.35	83.22	69.73	89.71	76.62	76.15	74.88	82.71	82.21
DGEDT [17]	79.80	75.60	86.30	80.00	84.00	71.00	91.90	79.00	77.90	75.40	-	-
kumaGCN [3]	81.98	78.81	86.43	80.30	<u>86.35</u>	70.76	92.53	79.24	77.89	77.03	-	-
DualGCN [9]	81.80	78.10	87.13	81.16	82.33	68.12	90.91	77.86	77.40	76.02	83.83	83.47
DR-BERT [20]	81.45	78.16	87.72	82.31	-	-	-	-	77.24	76.10	-	-
TF-BERT [21]	81.80	78.46	87.09	81.15	-	-	-	-	78.43	77.25	-	-
APARN [11]	81.96	79.10	87.76	<u>82.44</u>	-	-	-	-	<u>79.76</u>	<u>78.79</u>	<u>85.59</u>	<u>85.06</u>
CEIB [2]	<u>82.92</u>	<u>79.50</u>	<u>87.77</u>	82.08	86.16	<u>72.97</u>	<u>92.86</u>	<u>81.08</u>	-	-	84.95	84.41
GPT-4 (0-shot)	77.37	74.70	82.75	74.33	80.54	62.28	87.18	68.75	73.70	72.75	64.27	64.51
GPT-4 (8-shot)	78.15	75.79	84.62	76.32	82.04	68.54	88.83	71.90	74.13	73.23	65.15	65.06
Our MASS	**84.17**	**81.16**	**89.73**	**84.27**	**88.15**	**74.17**	**94.31**	**82.56**	**81.21**	**80.61**	**85.64**	**85.11**

datasets consist of three sentiments, including positive, neutral and negative. Each data item includes a sentence, an aspect and its sentiment polarity.

Implementation Details. During the implementation, we use *GPT-4* to generate aspect information and build our framework based on *bert-base-uncased*. The hidden size of attention mechanism is set to 300. The learning rate is tested among {1e-5, 2e-5, 4e-5} and the batch size is adjusted in {16, 32}. The dropout rate is set to 0.5. The hyper-parameter α and β are set to 0.8 and 0.4. The coefficients λ_1, λ_2 and λ_3 have been carefully adjusted, and finally are set to 1, 1 and 4, respectively. We evaluate it with accuracy and macro-F1 value.

Baseline methods. We compare our MASS with a series of advanced ABSA models based on *bert-base-uncased*: **BERT-SPC** [15], **R-GAT** [18], **DGEDT** [17], **kumaGCN** [3], **DualGCN** [9], **DR-BERT** [20], **TF-BERT** [21], **APARN** [11], **CEIB** [2]. In addition to the ABSA models mentioned above, we also test on GPT-4 to explore LLMs' ability to address the ABSA task straightforwardly.

4.2 Main Results

The experiment results of the ABSA methods on six benchmark datasets are reported in Table 1. We can observe that our MASS substantially and consistently outperforms all compared baselines on the overall datasets in terms of both accuracy and macro-F1 score. Specifically, our MASS achieves improvements of 1.25% ~ 1.96 % in accuracy and 1.20% ~ 1.83% in F1 value on five benchmark datasets (i.e., Lap14, Rest14, Rest15, Rest16, Twitter) compared with state-of-the-art baselines (i.e., APARN, CEIB). We attribute these advancements to the aspect information generated from LLMs and contrastive learning methods, which both significantly mitigate the issue of aspect-related data sparsity in ABSA. It could be observed that the performance improvement on the MAMs dataset is relatively modest, and we speculate it's because MAMs is a challeng-

Table 2. Ablation study results (%) on six benchmark datasets.

Models	Lap14		Res14		Res15		Res16		Twitter		MAMs	
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
MASS	**84.17**	**81.16**	**89.73**	**84.27**	**88.15**	**74.17**	**94.31**	**82.56**	**81.21**	**80.61**	**85.64**	**85.11**
w/o Aspect Information	82.13	79.32	87.32	81.63	86.16	70.39	92.68	78.82	78.03	77.09	84.66	84.30
w/o Fusion Module	82.92	79.84	88.78	82.70	86.85	70.76	93.33	80.89	79.48	78.69	84.81	84.16
w/o Strict CL	83.54	80.66	89.20	83.94	87.22	71.69	93.50	81.08	80.49	79.74	85.18	84.66
w/o Flaccid CL	83.39	80.56	88.93	83.38	87.41	73.38	93.50	81.50	80.64	79.76	85.10	84.46
w/o Unsupervised CL	83.23	80.24	89.11	83.35	87.59	72.97	93.66	81.23	79.91	79.10	84.88	84.25

ing dataset with more complex expressions and opinions, increasing challenges for LLMs to generate high-quality aspect information.

4.3 Ablation Study

We conduct extensive ablation studies and report the results in Table 2. The decrease in model *w/o Aspect Information* proves that aspect information generated by LLMs could effectively enrich aspect-oriented semantics, thus help minimize interference from irrelevant contexts. The results in model *w/o Fusion Module* illustrates the importance of devising an effective strategy to incorporate aspect information into the original sentence. Notably, ablation of any contrastive method results in a reduction of accuracy, which demonstrates our contrastive learning methods can effectively leverage correlations between samples.

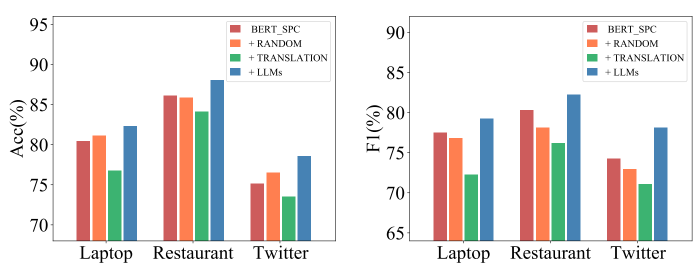

Fig. 4. BERT-SPC with different augmentation methods. The methods include: Add random words; Translate sentences into synonyms; Fuse aspect information.

4.4 Analysis on Aspect Information

Our work exploits aspect information derived from LLMs, which could be viewed as data augmentation. To further demonstrate that it is suitable for ABSA task, we compare it with conditional data augmentation methods on BERT-SPC. The results shown in Fig 4 reveal that our approach obtain substantial improvements, which reveal that aspect information derived from LLMs offers an effective and fine-grained augmentation that effectively mitigates data sparsity of the specific aspect. On top of that, other augmentation methods are unsuitable for ABSA because they may strip away vital information or introduce irrelevant words.

5 Conclusion

In this paper, we propose a novel framework MASS, which could mitigate aspect-oriented semantic sparsity for fine-grained sentiment analysis. Specifically, we leverage cutting-edge prompting techniques to stimulate LLMs to generate aspect information. Subsequently, we integrate the original sentence with the aspect information by employing a fusion and re-weighting module. Finally, we devise three contrastive learning methods to model semantic correlations between samples. Extensive experiments on six benchmarks demonstrate that our MASS surpasses state-of-the-art baselines.

Acknowledgement. This research was partially supported by the National Natural Science Foundation of China (No.62406303,62337001), Anhui Provincial Natural Science Foundation (No. 2308085QF229), Anhui Science and Technology Innovation Plan (No.202423k09020010) and the Fundamental Research Funds for the Central Universities (WK2150110034).

References

1. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., et al., K.: Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. NIPS'20, Red Hook, NY, USA (2020)
2. Chang, M., Yang, M., Jiang, Q., Xu, R.: Counterfactual-enhanced information bottleneck for aspect-based sentiment analysis. Proceedings of the AAAI Conference on Artificial Intelligence **38**(16), 17736–17744 (Mar 2024)
3. Chen, C., Teng, Z., Zhang, Y.: Inducing target-specific latent structures for aspect sentiment classification. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 5596–5607 (Nov 2020)
4. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. pp. 4171–4186 (Jun 2019)
5. Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., Xu, K.: Adaptive recursive neural network for target-dependent Twitter sentiment classification. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). pp. 49–54 (Jun 2014)
6. Gao, T., Yao, X., Chen, D.: SimCSE: Simple contrastive learning of sentence embeddings. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. pp. 6894–6910 (Nov 2021)
7. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (June 2020)

8. Jiang, Q., Chen, L., Xu, R., Ao, X., Yang, M.: A challenge dataset and effective models for aspect-based sentiment analysis. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)
9. Li, R., Chen, H., Feng, F., Ma, Z., Wang, X., Hovy, E.: Dual graph convolutional networks for aspect-based sentiment analysis. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)
10. Liu, J., Shen, D., Zhang, Y., Dolan, B., Carin, L., Chen, W.: What makes good in-context examples for GPT-3? In: Proceedings of Deep Learning Inside Out (DeeLIO 2022): The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures. pp. 100–114 (May 2022)
11. Ma, F., Hu, X., Liu, A., Yang, Y., Li, S., Yu, P.S., Wen, L.: AMR-based network for aspect-based sentiment analysis. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. pp. 322–337 (Jul 2023)
12. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.e.a.: SemEval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)
13. Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., Androutsopoulos, I.: SemEval-2015 task 12: Aspect based sentiment analysis. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)
14. Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.: SemEval-2014 task 4: Aspect based sentiment analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). pp. 27–35 (Aug 2014)
15. Song, Y., Wang, J., Jiang, T., Liu, Z., Rao, Y.: Attentional encoder network for targeted sentiment classification. In: International Conference on Artificial Neural Networks (2019)
16. Tan, X., Cai, Y., Zhu, C.: Recognizing conflict opinions in aspect-level sentiment classification with dual attention networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)
17. Tang, H., Ji, D., Li, C., Zhou, Q.: Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (Jul 2020)
18. Wang, K., Shen, W., Yang, Y., Quan, X., Wang, R.: Relational graph attention network for aspect-based sentiment analysis. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 3229–3238. Online
19. Zhang, K., Liu, Q., Qian, H., Xiang, B., Cui, Q., Zhou, J., Chen, E.: Eatn: An efficient adaptive transfer network for aspect-level sentiment analysis. IEEE Transactions on Knowledge and Data Engineering **35**(1), 377–389 (2021)
20. Zhang, K., Zhang, K., Zhang, M., Zhao, H., Liu, Q., Wu, W., Chen, E.: Incorporating dynamic semantics into pre-trained language model for aspect-based sentiment analysis. In: Findings of the Association for Computational Linguistics: ACL 2022. pp. 3599–3610. Dublin, Ireland (May 2022)
21. Zhang, M., Zhu, Y., Liu, Z., Bao, Z., Wu, Y., Sun, X., Xu, L.: Span-level aspect-based sentiment analysis via table filling. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. pp. 9273–9284 (Jul 2023)