

Learnable Dependency-based Double Graph Structure for Aspect-based Sentiment Analysis

Yinglong Ma and Yunhe Pang

School of Control and Computer Engineering
North China Electric Power University
Beinong Road #2, Beijing 102206, China
{yinglongma, pangyunhe}@ncepu.edu.cn

Abstract

Dependency tree-based methods might be susceptible to the dependency tree due to that they inevitably introduce noisy information and neglect the rich relation information between words. In this paper, we propose a learnable dependency-based double graph (LD2G) model for aspect-based sentiment classification. We use multi-task learning for domain adaptive pretraining, which combines Biaffine Attention and Mask Language Model by incorporating features such as structure, relations and linguistic features in the sentiment text. Then we utilize the dependency enhanced double graph-based MPNN to deeply fuse structure features and relation features that are affected with each other for ASC. Experiment on four benchmark datasets shows that our model is superior to the state-of-the-art approaches.

1 Introduction

Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment analysis task for extracting the emotional orientation of the given one or more aspects in a sentence. ABSA is generally divided into two parts: aspect extraction (AE) and aspect sentiment classification (ASC). We will focus on ASC task in this paper.

Dependency-based approaches have illustrated promising performance by utilizing the spatial structure of the dependency tree (DT). Early studies focus on the syntax feature enhanced aspect-based sentiment analysis (ABSA) models (Brun et al., 2014; Kiritchenko et al., 2014). Dependency-based analysis currently tends to use graph neural networks to learn representation over DT due to that graph neural network can directly learn on

graph data including DTs (Kipf et al., 2016). DT is regarded as an undirected tree structure to produce a symmetry adjacency matrix for Graph Convolutional Networks (GCN) (Sun et al., 2019; Zhang et al., 2019). (Wang et al., 2020) use GNNs to capture implicit relation between aspect words and others and achieved encouraged results. (Wang et al., 2021) utilized reinforcement learning for aspect-based sentiment classification (ASC).

In this paper, we consider how to effectively fuse structure and relation features for ASC. the following two questions: 1) How to utilize relations of edges between words for improving ASC? Most of existing approaches consider the structure information and neglected the rich edge relations between words with exception to the works (Chen et al., 2022; Tian et al., 2021a; Tian et al., 2021b). However, these works consider graph features (i.e., structure, edge relations, tree distance) as different input sources, and do not explore how to effectively fuse various features because these features are affected with each other. 2) How to effectively handle the multi-feature noisy information from DT for improving ASC? Existing approaches directly utilize the DT generated by a tool like spaCy. Direct application of DT is susceptible to noisy information in the dependency tree (Tang et al., 2020), which may achieve lower performance compared to using only the flat structure. DT includes 45 kinds of relations such as *nsubj*, *amod* and *conj*, and so on. Dependency parsing requires effective way to reduce DT's noisy information.

To handle the two problems, we propose a learnable dependency-based double graph (LD2G) model for aspect-based sentiment classification. We use multi-task learning (MTL) for domain adaptive pretraining and reducing DT's noisy information, by combining Biaffine Attention Model (BAM) and Mask Language Model (MLM) and

parsing structure, relations and linguistic features. We utilize the dependency enhanced double graph-based MPNN to deeply fuse structure features and relation features that are affected with each other for ASC. Experiment on four benchmark datasets shows that our model is superior to the state-of-the-art approaches. To the best of our knowledge, LD2G is the first work that jointly considers the structure and relations features in a unified domain adaptive pretraining framework. The main contributions of this paper are as follows.

- We propose a learnable dependency-based double graph MPNN model to learn dependency-based feature for ABSA.
- We present an MTL method for domain adaptive pretraining by combining Biaffine Attention and MLM model.
- Experiments were made over four datasets, and the results show that our model outperforms state-of-the-art approaches.

2 Related work

Generally, sentiment analysis aims to predict the sentiment of a sentence or a document. In contrast ABSA is a fine-grained task, which is to perform sentiment analysis on aspects by first identifying aspect words and then predicting the sentiment of the aspect. Early methods mainly use machine learning methods, but these methods are more dependent on manual feature labeling (Jiang et al., 2011; Vo et al., 2015). Later, neural network approaches were found to have better performance. A common method was to model the sequence of aspect words and context. The popular LSTM model and attention mechanism were commonly used to fuse the context semantics and obtain the important parts from it. In recent years, approaches based on pre-trained model has become the main method for contextual semantic integration, which brings considerable improvement to sentiment analysis.

Another kind of approach takes advantage of the spatial structure of syntax features by deep learning methods. Early work has studied the syntax feature enhanced ABSA models (Brun et al., 2014, Kiritchenko et al., 2014). Dependency tree and constituent tree can represent the syntactic information between the words, but these approaches based on them has not become more popular until deep learning method was applied.

After the improvements brought by deep learning, syntax features can effectively improve the performance of ABSA task. In recent years, syntax approaches tend to use graph neural networks to represent the dependency tree, because GNNs can directly model non-Euclidean data which fits dependency tree well. (Sun et al., 2019; Zhang et al., 2019) combined Bi-LSTM and GCN (Kipf et al., 2016) to extract feature from both linear and spatial word representations, in which dependency tree is seemed as an undirected dependency tree to produce a symmetry adjacency matrix for GCN. (Dai et al., 2021) generated an induced tree structure based on the pre-trained model instead of dependency tree, and the experiment shows that pre-trained models can optimize the induced tree in the following ABSA tasks. (Zhou et al., 2021) shortened the distance between aspect word and emotional words by learning a new aspect-centric tree. (Tang et al., 2020) combined a dependency graph and an attention-based graph to learn representation based on GNN. There are some approaches such as (Chen et al., 2022; Tian et al., 2021a; Tian et al., 2021b; Wang et al., 2021), further explored the use of specific relation on the dependency trees. These works utilize the original features of the dependency trees.

Dependency-based methods add effective syntactic features to sentiment analysis tasks, but at the same time, they inevitably inherit wrong dependency tree information, which brings adverse influence on downstream task. Therefore, it has become a crucial problem how to utilize relations of edges between words and handle the multi-feature noisy information effectively from DT for improving ASC

3 LD2G model

The LD2G model overall is shown in Figure 1. We use a two-stage approach to implement aspect sentiment analysis: 1) Domain adapted pretraining, and 2) Double graph MPNN-based aspect sentiment classification. The whole model is divided into three modules: BERT module, biaffine parser and message passing neural network. We first apply MTL on BERT and biaffine parser module for domain adaptive pretraining by combining BAM and MLM, then obtain the sparse score matrix of structure graph and the feature matrix of edge relations from BAM. Both them are further fused into

a dependency enhanced double graph matrix. The LD2G model can learn from the dependency graphs and optimize the dependency graph structure during the training process.

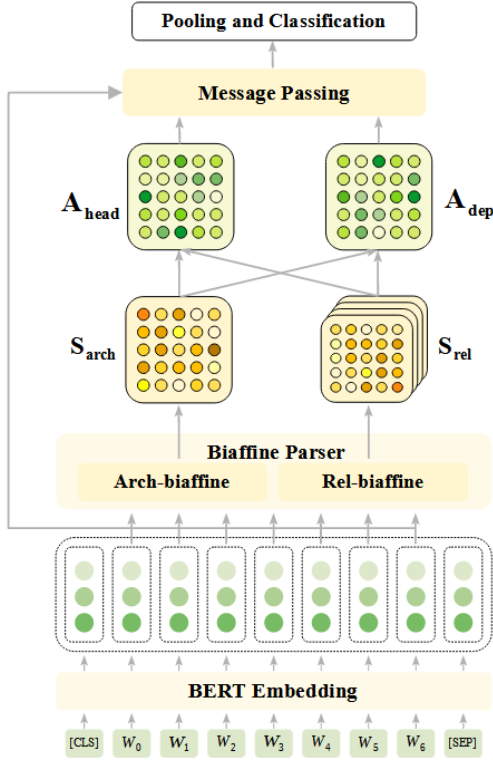


Fig 1. The architecture of LD2G model.

3.1 Domain adapted pretraining

A multi-task learning (MTL) task is launched for domain adaptive dependency feature fusion, which combines dependency parsing (DP) and MLM (Devlin et al., 2018) by incorporating features such as structure, relations and linguistic features in the sentiment text. For a given sentence $s = \{w_0, w_1, \dots, w_n\}$ as input, we obtain the word embedding $H = \{h_0, h_1, \dots, h_n\}$ with BERT. MLM is applied to BERT model for domain adaptation, using the same mask strategy with (Devlin et al., 2018). Biaffine attention model (Dozat and Manning, 2017) is used based on the output of BERT model for dependency parsing. We use the architecture biaffine and relation biaffine to obtain the score matrices.

$$S_{arch} = Biaffine_{arch}(H) \quad (1)$$

$$S_{rel} = Biaffine_{rel}(H) \quad (2)$$

Where $S_{arch} \in \mathbb{R}^{n \times n}$, and $S_{rel} \in \mathbb{R}^{n \times n \times d}$ for all the edges in S_{arch} . In this paper, $d=45$, which is

the number of dependency relation types. The DP task and MLM task will be combined as a MTL task.

The pretrained dependency graph will be taken as the original input of double graph MPNN model. The pretrained BAM with optimal initialization parameters will be used for MPNN training. After the BAM dependency parsing is trained, we can obtain the intermediate dense score matrix S_{arch} and feature matrix S_{rel} .

3.2 Double graph MPNN-based aspect sentiment classification

The structure information S_{arch} can be regarded as an adjacency matrix A_{arch} , and the relation information S_{rel} as a relation graph matrix A_{rel} by MLP.

$$A_{rel} = MLP(S_{rel}) \quad (3)$$

We use an adjacency matrix and its transpose to distinguish the two directions between head-dependent and dependent-head pairs in dependency graph. Four graphs are respectively represented by adjacency matrices A_{arch} and A_{rel} , and their transposes A_{arch}^T and A_{rel}^T . The structure and relation features can be fused into a dependency enhanced graph A_{head} with transpose A_{dep} .

$$A_{head} = A_{arch} + A_{rel} \quad (4)$$

$$A_{dep} = A_{arch}^T + A_{rel}^T \quad (5)$$

Where all the matrices are $\mathbb{R}^{n \times n}$ matrices.

Row normalization is applied to each matrix. The token embeddings from BERT output H that is straightly fed into the MPNN graph neural networks as node representation H^0 . To learn the node representation over both A_{head} and A_{dep} , we take the bidirectional message passing (Kampffmeyer et al., 2019) as one layer to acquire final hidden states.

$$H^{l+\frac{1}{2}} = LN(ReLu(A_{head}H^l\theta_{head}^l) + H^0) \quad (6)$$

$$H^{l+1} = LN(ReLu(A_{dep}H^{l+\frac{1}{2}}\theta_{dep}^l) + H^0) \quad (7)$$

Where LN denotes layer normalization. After two layers of message passing, we can obtain the final representation. The pooling function is used to the hidden state of aspect tokens for obtaining the aspect representation \mathbf{r} , by a softmax normalization layer to yield a probability distribution \mathbf{p}_c .

$$\mathbf{p}_c = softmax(W_p \mathbf{r} + b_p) \quad (8)$$

Models	Laptop14		Rest14		Rest15		Rest16		
	Acc	F1	Acc	F1	Acc	F1	Acc	F1	
Baselines	ASGCN-BERT	77.90	73.01	83.78	75.02	80.69	62.02	89.99	74.46
	SK-GCN-BERT	79.00	75.57	83.48	75.19	83.20	66.78	87.19	72.02
	BERT + MLP	78.36	74.16	85.35	78.38	82.16	64.96	89.43	74.20
	R-GAT+BERT	78.53	74.63	85.63	78.82	81.61	65.30	90.96	75.26
	ACLT	79.68	75.83	85.71	78.44	84.44	72.08	92.15	78.64
	DGEDT-BERT	79.80	75.60	86.30	80.00	84.00	71.00	91.90	79.00
	TGCN-BERT	80.25	76.92	85.54	78.86	85.07	72.50	91.83	76.86
	KVMN-BERT	79.78	76.14	85.98	77.94	84.14	68.49	90.52	73.68
Ablation study	Pretrain BERT+MLP	79.15	75.15	85.63	78.56	85.24	71.29	92.21	78.02
	LD2G with static biaffine	77.36	73.01	84.46	75.79	84.50	68.34	90.09	72.09
	LD2G w/o MTL pretrain	78.53	74.03	84.73	78.22	83.39	70.39	91.40	78.97
	LD2G w/o rel. graph	79.93	75.94	85.53	78.24	84.87	67.91	91.23	75.13
	LD2G w/o arch. graph	79.62	76.12	85.36	78.24	84.50	65.84	92.21	77.67
	LD2G	80.56	76.61	85.98	79.45	85.79	72.11	92.69	80.48

Table 1 Performance evaluation of ASC over four datasets.

Where W_p and b_p are both trainable parameters. The loss is calculated by cross entropy:

$$\ell = -\sum_{d \in D} \sum_{c \in P} \log p_c \quad (9)$$

Where D denotes the training dataset, and P denotes all the polarities in ASC.

4 Experiment

4.1 Experimental setup and datasets

We use Stanford Dependency (SD) conversion of the English Penn Treebank (PTB 3.3.0) dependency datasets to train the BERT model and dependency parser. For pretraining and aspect sentiment analysis, we made the experiments over four datasets: Laptop14, Rest14, Rest15 and Rest16. Datasets Laptop14 and Rest14 are respectively from SemEval 2014 task 4 (Pontiki et al., 2014); Dataset Rest15 is from SemEval 2015 task 12 (Pontiki et al., 2015) and Rest16 is from SemEval 2016 task 5 (Pontiki et al., 2016), consisting of data from two categories: laptop and restaurant. Each of the datasets has three categories: positive, negative and neutral.

4.2 Parameters settings

We used the similar settings in dependency parsing and MPNN-based ASC. The common AdamW optimizer was used. The learning rate was $2e-5$, warm

up ratio is 0.1 and dropout rate is 0.1. We set the batch size as 16. In dependency parsing task, because the model converges quickly, we only train the model for 3 epochs to reduce the risk of overfitting while we train 10 epochs on ASC task.

4.3 Baselines

We compare the proposed LD2G¹ with the eight baselines and state-of-the-art alternatives: 1) ASGCN+BERT (Zhang et al., 2019); 2) SK-GCN+BERT (Zhou et al., 2020); 3) BERT-BASE+MLP (Dai et al., 2021); 4) R-GAT+BERT (Wang et al., 2020); 5) ACLT (Zhou et al., 2021); 6) DGEDT-BERT (Tang et al., 2020); 7) TGCN-BERT (Tian et al., 2021a); 8) KVMN-BERT (Tian et al., 2021b).

4.4 Overall Results

We use the accuracy and macro-averaged F1-score as evaluation metrics. The experimental results are shown in Table 1.

In Table 1, we observe that our LD2G model outperforms all the baseline methods over almost all datasets when we use simple message passing layers with residual connection. This shows that the utilization of double graph-based dependency enhanced structure fusing structure and relation features indeed contributes to ASC performance improvement. With the exceptions, LD2G has the

¹ <https://github.com/LLLpc123/LDG>

slightly lower ACC and F1 values than DGEDT-BERT over Rest14, and also has the slightly lower F1 values than TGCN-BERT over lap14 and rest15.

In addition, the results show that all the baselines except for “DGEDT-BERT” have lower ASC

the double graph-based MPNN in LD2G can continuously update and optimize BAM for generating the dependency graph with less DT’s noisy information. In contrast, LD2G with static biaffine cannot dynamically optimize dependency parsing.

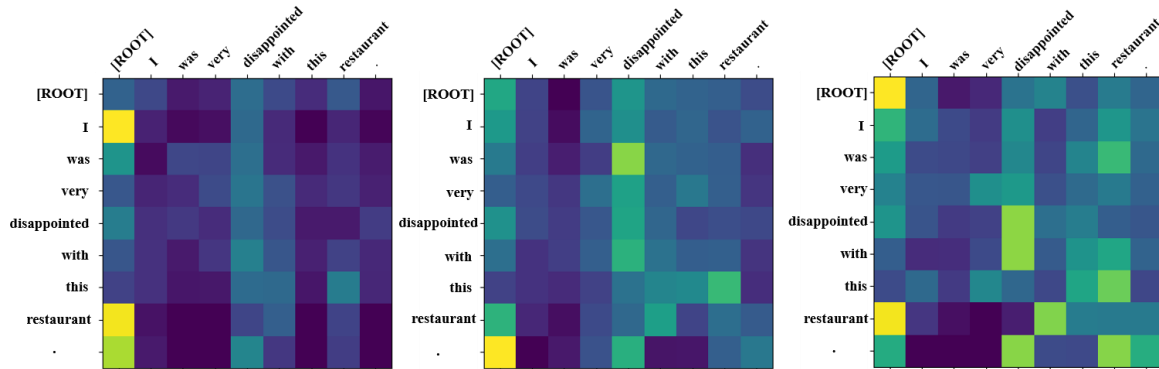


Fig 2. The change process of adjacency matrix. The figures from left to right respectively show the original output of trained dependency parser, the middle state of training process, and the adjacency matrix when the model converges. Light colors indicate that the location has higher weight.

performance than LD2G. A very likely reason is that: 1) BERT-based linguistic pretraining is not enough, and 2) their models do not use MTL pretraining and therefore are difficult to handle and reduce the DT’s noisy information.

4.5 Ablation study

We consider five ablation baselines including: 1) BERT-based pre-train: we replace the MTL pretrain only with BERT’s MLM pretraining on ABSA dataset by a simple BERT+MLP. 2) LD2G with static biaffine. We lock the parameters of biaffine layer to observe the effect of a learnable graph structure. 3) LD2G w/o MTL pretrain. MTL pretrain is removed from LD2G to observe the effect. 4) LD2G w/o rel. graph: Relation graph for all the edges in the dependency tree is removed to observe the effects. And 5) LD2G w/o arch. graph: Architecture graph is removed. The results for ablation study are shown in Table 1.

From Table 1, LD2G with full modules overall outperforms the others with incomplete functions over all the datasets against metrics accuracy and F1 score. First, we observe that LD2G with MTL pretrain performs better than BERT+MLP model with MLM pretraining. This shows that BERT’s MLM pretraining is not enough to reduce the DT’s noisy information, but the MTL with BAM and MLM can effectively capture the various features in dependency parsing and obviously reduce the noise. Second, LD2G with static biaffine has the obviously lower performance than LD2G because

Third, LD2G with MTL pretraining has obviously much higher performance than LD2G without MTL pretraining, which shows that MTL is crucial to effectively capture features and reduce DT’s noisy information. Four, LD2G outperforms each of models LD2G w/o rel. graph and LD2G w/o arch. The combination of both graph structure feature and relation features indeed contributes to performance improvement. This also shows that using learnable dependency-based double graph MPNN indeed contributes to ASC performance.

4.6 Graph visualization-based case study

To illustrate the task-oriented learnable graph optimization of our model, we visualized the adjacency matrix in different phases and epochs. Considering a sample sentence, i.e., “I was very disappointed with this restaurant.” As shown in Figure 2, the generated dependency graph structure is continuously optimized with the training of the model. The column of word ‘disappointed’ is growing lighter, which shows the importance to classification task.

5 Conclusions and future work

In this paper, we proposed an LD2G model to fuse the syntactical structure and relation features for ASC. We use an MTL method for domain adaptive pretraining by combining Biaffine Attention and MLM model. A dependency enhanced double graph-based MPNN is utilized to deeply fuse structure features and relation features that are affected

with each other for ASC. Experiments results show that our model is superior to the state-of-the-art alternatives and the utilization of double graph-based dependency enhanced structure indeed contributes to ASC performance improvement.

In the future, we will analyze combinations of linguistic features and priori knowledge mining for ABSA.

Acknowledgements

This work is partially supported by National key research and development program under Grants (No. 2018YFC0830605, No. 2018YFC0831404).

Reference

- Caroline Brun, Diana Nicoleta Popa, and Claude Roux. 2014. XRCE: Hybrid classification for aspect-based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 838–842, Dublin, Ireland. Association for Computational Linguistics.
- Hao Chen, Zepeng Zhai, Fangxiang Feng, Ruifan Li, and Xiaojie Wang. 2022. Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022, pages 2974–2985. Association for Computational Linguistics.
- Y. Chu and T. Liu. 1965. On the shortest arborescence of a directed graph. *Scientia Sinica*, 14: 1396–1400. American Mathematical Society.
- Junqi Dai, Hang Yan, Tianxiang Sun, Pengfei Liu, and Xipeng Qiu. 2021. Does syntax matter? A strong baseline for aspect-based sentiment analysis with roberta. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, pages 1816–1829. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. CoRR, abs/1810.04805.
- Timothy Dozat and Christopher D. Manning. 2017. Deep biaffine attention for neural dependency parsing. In 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings.
- Jason Eisner. 1997. Three new probabilistic models for dependency parsing: An exploration. CoRR, cmlp/9706003.
- Michael Kampffmeyer, Yinbo Chen, Xiaodan Liang, Hao Wang, Yujia Zhang, and Eric P. Xing. 2019. Rethinking knowledge graph propagation for zeroshot learning. In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019, pages 11487–11496. Computer Vision Foundation / IEEE.
- Thomas N. Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. CoRR, abs/1609.02907.1
- Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif Mohammad. 2014. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 437–442, Dublin, Ireland. Association for Computational Linguistics.
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad Al-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia V. Loukachevitch, Evgeniy V. Kotelnikov, Núria Bel, Salud María Jiménez Zafra, and Gülsen Eryigit. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2016, San Diego, CA, USA, June 16-17, 2016, pages 19–30. The Association for Computer Linguistics. 205
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. Semeval-2015 task 12: Aspect based sentiment analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2015, Denver, Colorado, USA, June 4-5, 2015, pages 486–495. The Association for Computer Linguistics.
- Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 27–35, Dublin, Ireland. Association for Computational Linguistics.
- Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao, and Xudong Liu. 2019. Aspect-level sentiment analysis via convolution over dependency tree. 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 2019, Hong Kong, China, November 3-7, 2019, pages 5678–5687. Association for Computational Linguistics.

- Hao Tang, Donghong Ji, Chenliang Li, and Qiji Zhou. 2020. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020, pages 6578–6588. Association for Computational Linguistics.
- Yuanhe Tian, Guimin Chen, and Yan Song. 2021a. Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, pages 2910–2922. Association for Computational Linguistics.
- Yuanhe Tian, Guimin Chen, and Yan Song. 2021b. Enhancing aspect-level sentiment analysis with word dependencies. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, EACL 2021, Online, April 19 - 23, 2021, pages 3726–3739. Association for Computational Linguistics.
- Kai Wang, Weizhou Shen, Yunyi Yang, Xiaojun Quan, and Rui Wang. 2020. Relational graph attention network for aspect-based sentiment analysis. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020, pages 3229–3238. Association for Computational Linguistics.
- Lichen Wang, Bo Zong, Yunyu Liu, Can Qin, Wei Cheng, Wenchao Yu, Xuchao Zhang, Haifeng Chen, and Yun Fu. 2021. Aspect-based sentiment classification via reinforcement learning. In IEEE International Conference on Data Mining, ICDM 2021, Auckland, New Zealand, December 7-10, 2021, pages 1391–1396. IEEE.
- Chen Zhang, Qiuchi Li, and Dawei Song. 2019. Aspect-based sentiment classification with aspect-specific graph convolutional 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 2019, Hong Kong, China, November 3-7, 2019, pages 4567–4577. Association for Computational Linguistics.
- Jie Zhou, Jimmy Xiangji Huang, Qinmin Vivian Hu, and Liang He. 2020. SK-GCN: modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification. *Knowl. Based Syst.*, 205:106292.
- Yuxiang Zhou, Lejian Liao, Yang Gao, Zhanming Jie, and Wei Lu. 2021. To be closer: Learning to link up aspects with opinions. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021, pages 3899–3909. Association for Computational Linguistics.
- Long Jiang, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. 2011. Target-dependent Twitter sentiment classification. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pages 151–160, Portland, Oregon, USA. Association for Computational Linguistics.
- DuyTin Vo and Yue Zhang. 2015. Deep learning for event-driven stock prediction. In Proceedings of IJCAI, Buenos Aires, Argentina.