2024 International Conference on Computational Linguistics and Natural Language Processing (CLNLP) | 979-8-3503-9209-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/CLNLP64123.2024.00020

SISK: Integrating Aspect-oriented Semantic Information and Syntactic Knowledge for Aspect-based Sentiment Analysis

Yanjiang Chen, Kai Zhang*, Zirui Liu, Meikai Bao, Xukai Liu, and Qi Liu
State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China, Hefei, China
{yjchen, liuzirui, baomeikai, chthollylxk}@mail.ustc.edu.cn, {kkzhang08, qiliuql}@ustc.edu.cn

***Abstract*—Aspect-based sentiment analysis (ABSA) is a fine-grained classification task that focuses on capturing the relationship between aspect and its context within a sentence. Recent years, many works have effectively utilized dependency trees to leverage syntactic information and achieved significant improvements. However, approaches focusing solely on syntax might face from potential mismatch between the syntactic dependency tree and the semantic nature of sentiment classification. Therefore, there is a need for research that concentrate on semantic understanding while leveraging syntactic knowledge to assist classification. In this paper, we introduce a novel method named Integrating Aspect-oriented Semantic Information and Syntactic Knowledge (SISK) for ABSA. Initially, we employ self-attention mechanisms to explore the semantics within the sentence and re-weight semantic features based on position of words. Subsequently, we leverage dependency trees and graph convolutional networks to extract syntactic knowledge. In addition, we utilize aspect features to guide the extraction and obtain aspect-oriented syntactic features. Finally, we interactively integrate semantic and syntactic information to predict sentiment polarities. Extensive experiments on three benchmark datasets show that our approach outperforms many advanced models, proving that combining semantics and syntax effectively enhances ABSA performance.**

***Index Terms*—aspect-based sentiment analysis; attention mechanism; graph convolutional networks; dependency trees; pre-trained language models**

I. Introduction

Aspect-based Sentiment Analysis (ABSA) [1], [2] is a fine-grained classification task which focuses on determining the sentiment polarity of specific aspects. For instance, given the sentence "The service in that restaurant is terrible but the atmosphere is comfortable", the goal of ABSA is to predict the sentiment polarities of two aspects "service" and "atmosphere", which should be negative and positive, respectively.

The primary challenge of ABSA is the precise identification of the relationships between aspects and their corresponding contexts, especially opinion words [3], [4], [5]. Earlier methods predominantly focus on constructing sentiment lexicons [1], [6], [7]. However, these methods rely heavily on meticulously crafted sentiment lexicons and are sensitive to their quality. In recent years, various studies [8], [9], [10], [11], [12], [13], [14], [15] [16] have used attention mechanisms to capture the semantic relationship between an aspect and its context. The core idea of these methods is to assign higher attention scores to words in the sentence, thereby guiding the final classification. Moreover, many methods [17], [18], [19], [20], [21], [22], [23] utilize graph neural networks (GNNs) on dependency trees to capture structural dependencies between an aspect and its sentiment expressions. These models leverage various types of syntactic information through different approaches and achieve appealing results.

Despite substantial advancements by existing methods, they still suffer the interference caused by irrelevant contextual words. On one hand, attention-based models may erroneously attribute high attention scores to words that are not pertinent to the specific aspect when sentences contain multiple aspects or diverse sentiment expressions. For instance, in the sentence "The price is really reasonable though the service is a little poor", the opinion word "reasonable" might obtain higher attention scores than the word "poor" concerning the aspect "service". Therefore, methods that just focus on attention mechanisms struggle to avoid noise interference caused by mismatched opinion terms and contexts. On the other hand, although approaches concentrating on syntax mitigate these issues, they are encumbered by several limitations. Firstly, the outputs from dependency parsers sometimes exhibit inaccuracies, which might introduce more noise instead of offering valid information. Secondly, colloquial or grammatically irregular sentences provide sparse syntactic information. Thirdly, sentiment analysis is a semantics-oriented task with a certain gap from syntactic structures. Therefore, methods that fully comprehend sentence semantics and appropriately utilize syntactic knowledge remain to be further explored.

To address the aforementioned issues, in this paper, we propose a novel framework called Integrating Aspect-oriented Semantic Information and Syntactic Knowledge for Aspect-based Sentiment Analysis (SISK). This approach employs attention mechanisms and graph convolutional networks (GCN) to interactively leverage semantics and syntactic knowledge, which effectively capturing essential sentiment elements to the fine-grained classification. Specifically, we first employs self-attention mechanisms to explore the semantics of sentences, which capture implicit sentiment information and extract crucial contexts within the sentence. To mitigate potential biases inherent in the attention scores, we design a position-aware re-weighting module. Subsequently, we then create ad-

* Corresponding Author.

979-8-3503-9209-8/24/$31.00 ©2024 IEEE
DOI 10.1109/CLNLP64123.2024.00020

jacency matrices from dependency trees and use GCN to learn syntactic information, improving relational constraints within the sentences. In order to capture aspect-oriented syntactic information, we utilize aspect features to guide the process of syntactic information extraction. Finally, we interactively integrate semantic and syntactic information for better sentence comprehension, thereby diminishing the impact of noise and improving prediction accuracy. As fine-tuning paradigm of Pre-trained Language Models (PLMs) show exceptional performance in ABSA [24], [25], we utilize BERT [26] as the encoder for the ABSA task. Extensive experiments on three benchmark datasets demonstrate the effectiveness of our proposed model. The main contributions of this paper are as follows:

- We propose a novel ABSA approach that integrates both semantic and syntactic information of the sentence through interactive attention, allowing them to mutually reinforce and constrain each other. This integration effectively mitigates noise from irrelevant contextual information, thereby enhancing the accuracy of sentiment analysis predictions.
- We utilize attention mechanisms to extract contextual semantics, incorporating position information to capture aspect-aware semantic features. In addition, we introduce syntactic information using dependency trees, employing aspect features to guide GCN to extract aspect-oriented syntactic features.
- The experimental results validate the effectiveness of our approach and underscore the feasibility of integrating both semantic and syntactic considerations.

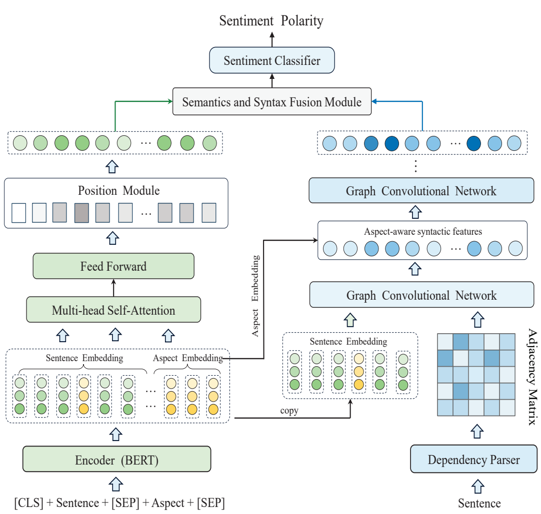

Fig. 1: The overall framework for our proposed SISK.

II. METHODOLOGY

In this section, we describe our SISK in detail. Specifically, we begin with the problem definition, followed by encoder module and the overall architecture of SISK. As illustrated in Figure 1, our proposed framework comprises three main components: (1) Aspect-aware Semantics Extraction Module, (2) Aspect-oriented Syntax Extraction Module, (3) Semantics-Syntax Fusion Module.

Problem Definition. Given a sentence-aspect pair (s, a), where $s = \{s_1, s_2, ..., s_l\}$ represents the sentence, and $a = \{a_1, a_2, ..., a_m\}$ represents the given aspect. ABSA aims to predict the sentiment polarity c (i.e.,positive, neutral, and negative) of the aspect a within the sentence s.

Text Encoder. we utilize BERT as the encoder in our SISK. Specifically, given the sentence s and the aspect a, we construct $[CLS]\ s\ [SEP]\ a\ [SEP]$ as the input, obtaining the feature representations of the sentence $H = \{h_1, h_2, ..., h_n\}$, where $h_i \in R^d$, and d denotes the dimension of the embedding layer. The features of the aspect term a is denoted by $H_a = \{h_1^a, h_2^a, ..., h_m^a\}$.

A. Aspect-aware Semantics Extraction Module

Multi-head Self-attention Mechanisms. To model the interaction between the aspect and the entire sentence, we employ self-attention modules to capture hidden information, thereby extracting key content and enhancing the model's ability to identify long-range dependencies. Specifically, after obtaining the sentence embedding representation H through the encoder, we employ multi-head self-attention mechanisms (MultiHead) with n heads to capture the sentence's overall semantics, which can be formulated as:

$$H_m = MultiHead(HW_h^q, HW_h^k, HW_h^v) \tag{1}$$

where W_h^q, W_h^k, W_h^v are the learnable parameters. $H_m = \{h_1^m, h_2^m, ..., h_n^m\}$ represents the sentence features with internal introspection and contextual interaction.

Feed-Forward Network. Given that the MultiHead is essentially a series of linear transformations, we further employ Feed-Forward Network (FFN) to achieve nonlinear transformations of the features. Specifically, the FFN includes two linear transformation layers with a ReLU activation function between them, as denoted by the following equations:

$$H_f = ReLU(H^m W_1 + b_1) * W_2 + b_2 \tag{2}$$

where W_1, b_1, W_2, b_2 are the learnable parameters. Consequently, we transform the initial sentence representation H into a feature vector $H_f = \{h_1^f, h_2^f, ..., h_n^f\}$.

Aspect-aware Semantic Features. Previous studies [10], [27] indicate that words near to the aspect are often more relevant, thus emphasizing the importance of nearby contexts. Inspired by this, we perform linear weighting on H_f according to the positional information of the aspect to obtain aspect-aware semantic features H^{sem}. This strategy aims to mitigate the natural noise and bias caused by the self-attention analysis process. We design a function P to compute this:

$$p_i = \begin{cases} 1 - \frac{(k+m-i)}{C} & 1 \le i < k+m \\ 1 - \frac{i-k}{C} & k+m \le i \le n \end{cases} \tag{3}$$

$$h_i^{sem} = P(h_i^f) = p_i h_i^f \tag{4}$$

where p_i represents the positional weight for the i-th word, k is the index of the starting position of the aspect, m is the aspect's length, and C is a predefined constant. The final output of this module is $H^{sem} = \{h_1^{sem}, h_2^{sem}, h_3^{sem}, ..., h_n^{sem}\}$.

B. Aspect-oriented Syntax Extraction Module.

Graph Convolutional Networks (GCNs) are uniquely suited to operate on graph-structured data. In graph data structures, GCNs efficiently encode local neighborhood information by performing convolution operations on adjacent nodes. By employing multi-layer GCNs, each node is capable of aggregating information from its neighboring nodes, thus facilitating the learning of more comprehensive global representations.

Syntax-based GCN. In this module, we utilize spaCy, an advanced natural language processing toolkit, to construct the dependency tree of a given sentence. Dependency trees can not only capture the syntactic relationships between words, but also reveal the structure and semantics of the sentence through the edges. By constructing dependency trees, we transform the sentence into a graph structure. More concretely, we build the adjacency matrix A based on the tree, where A_{ij} represents whether node i is connected to node j.

$$\boldsymbol{A_{ij}} = \begin{cases} 1 & \text{if } \mathrm{link}(i,j) = True \text{ or } i = j, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where $link(i,\ j)$ represents whether i-th and j-th token have a dependency relationship. Then, we input the sentence embedding H into the GCN to extract the syntactic features, which can be calculated as follows:

$$\tilde{h}_i^l = \sum_{j=1}^{n} A_{ij} W^l h_j^{l-1} \tag{6}$$

$$h_i^l = ReLU\left(\tilde{h}_i^l/(d_i+1) + b^l\right) \tag{7}$$

where A_{ij} is our defined adjacency matrix, W^l and b^l are learnable parameters, h_i^l represents the final output of the i-th word through the l-th GCN layer, and h_j^{l-1} represents the output of the j-th word from the previous GCN module. $d_i = \sum_{j=1}^{n} \mathbf{A}_{ij}$ represents the degree of the i-th word in the dependency graph. Specifically, the initial input is the sentence embedding H from the encoder, and the output of the l-th GCN layer is denoted as:

$$H^l = \{h_1^l, h_2^l, h_3^l, ..., h_n^l\} \tag{8}$$

Aspect-oriented Syntactic Information. Unlike sentence-level sentiment classification, ABSA focuses on determining the sentiment of specific aspects. Therefore, we design aspect-oriented syntactic information extraction, which integrates the aspect embedding H^a with the syntactic features H^l derived from the l-th GCN layer. Specifically, we use the aspect features to guide GCN to extract aspect-oriented syntactic information H^l, which is formalized as follows:

$$H^l = W_s H + (W_g H^l + W_a \bar{H}_a) \otimes w \tag{9}$$

where $\bar{H}_a$ represents aspect embedding, when the aspect consists of multiple words, their average value is used as the representation. H represents the original sentence embedding. W_s, W_g, W_a and w are trainable parameters. The output H^l is the aspect-oriented syntactic feature, which is also the input of next GCN layer. Finally, the output of the final GCN layer is $H^{syn} = \{h_1^{syn}, h_2^{syn}, h_3^{syn}, ..., h_n^{syn}\}$.

C. Semantics-Syntax Fusion Module

After extracting the aspect-aware semantic features $H^{sem} \in R^{n \times d_h}$ and aspect-oriented syntactic features $H^{syn} \in R^{n \times d_h}$, we jointly fuse these features to obtain the final sentiment representation. Specifically, we design a method for the interaction between semantic and syntactic features. The core idea of this design is to retrieve syntactic features that are significantly related to semantics during the interaction between semantics and syntax, and to assign attention weights to context words using these syntactic features. Initially, we perform a dot product operation to facilitate the interaction between the semantic and syntactic features of each word. Then, we quantify the significance of each node's syntactic relevance relative to semantics by summing the results. Finally, we assign syntax-based attention weights to each word according to this significance. This design allows syntactic knowledge to assist in the analysis of semantic features. The detailed computation is as follows:

$$\beta_t = \sum_{i=1}^{n} h_t^{sem\top} h_i^{syn} \tag{10}$$

$$\alpha_t = \frac{\exp(\beta_t)}{\sum_{i=1}^{n} \exp(\beta_i)} \tag{11}$$

$$e = \sum_{t=1}^{n} \alpha_t h_t^{sem} \tag{12}$$

After obtaining the final sentiment feature e, we pass it through a fully connected layer and a softmax normalization layer to map it to the probability distribution over the three sentiment polarities:

$$\hat{y} = softmax(W_e e + b_e) \tag{13}$$

TABLE I: The statistics of three benchmark datasets.

Datasets	Positive		Neutral		Negative	
	Train	Test	Train	Test	Train	Test
Restaurant	2164	728	637	196	807	196
Laptop	994	341	464	169	870	128
Twitter	1561	173	3127	346	1560	173

D. Model Training

During training process, we employ the Label Smoothing Regularization (LSR) method to alleviate overfitting problem. Specifically, LSR mitigates this issue by adjusting the weights of the true labels in the loss function, redefining the assignment weights for each label. For a sample x, we let $q(k|x)$ denote

TABLE II: Experiment Results (%) on Laptop, Restaurant, and Twitter Datasets.

Models	**Laptop**		**Restaurant**		**Twitter**	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
IAN [9]	72.10	-	78.60	-	-	-
RAM [10]	74.49	71.35	80.23	70.80	69.36	67.30
ASGCN [17]	75.55	71.05	80.77	72.02	72.15	70.40
BERT-SPC [28]	78.99	75.03	84.46	76.98	74.13	72.73
RGAT-BERT [18]	78.21	74.07	86.60	81.35	76.15	74.88
HGCN [29]	79.59	76.24	86.45	80.60	76.52	75.37
T-GCN [30]	80.88	77.03	86.16	79.95	76.45	75.25
ACLT [21]	79.68	75.83	85.71	78.44	75.48	74.51
AG-VSR [31]	79.92	75.85	86.34	80.88	76.45	75.04
SSEGCN [22]	81.01	77.96	87.31	81.09	77.40	76.02
CLEAN [32]	80.45	77.25	87.05	81.40	-	-
SISK	**81.50**	**78.36**	**87.32**	**82.23**	**77.60**	**76.61**

the probability that the sample x belongs to class k. LSR method updates $q(k|x)$ as follows:

$$q(k|x) = (1 - \epsilon)q(k|x) + \epsilon * u \tag{14}$$

where ϵ is the smoothing factor and u is a manually introduced fixed distribution. We train our SISK to minimize the following loss function:

$$\mathcal{L} = -\sum_{i=1}^{M}\sum_{j=1}^{C} y_i^j \log\left(\hat{y}_i^j\right) + \lambda\|\Theta\|^2 \tag{15}$$

where y_i^j represents the true sentiment probability, C is the number of label classes, M is the number of training samples, Θ denotes all trainable parameters, and λ is the coefficient for L2 regularization.

III. EXPERIMENTS

A. Experimental Setup

Datasets. We conducted experiments on three benchmark datasets in ABSA: Restaurant and Laptop in SemEval 2014 [2], and the Twitter dataset [33]. Each dataset has three sentiment categories: positive, neutral, and negative. Each entry includes a sentence, an aspect term, and the sentiment label of the aspect. We use the official data splits as provided in the original papers. Table I provides detailed information about these datasets.

Implementation Details. We build our SISK based on the *bert-base-uncased*, setting both MultiHead and GCN hidden layer dimensions to 300, with 8 heads in MultiHead. The number of MultiHead and GCN layers is set to 2. The learning rate is adjusted within the range of [1e-5, 2e-5], and the batch size is set within the range of [16, 32]. The dropout rate is set to 0.2. The L2 regularization coefficient λ is set to 1e-5. The Adam optimizer is used for parameter optimization, and performance is evaluated based on Accuracy and macro-F1 scores.

Baselines. To comprehensively assess our SISK thoroughly, we compare its performance on accuracy and macro-F1 value with several advanced models, including: IAN [9], RAM [10], ASGCN [17], BERT-SPC [28], RGAT-BERT [18], HGCN [29], T-GCN [30], ACLT [21], AG-VSR [31], SSEGCN [22], CLEAN [32].

B. Experimental Results

Table II presents the results of our SISK. It can be observed that compared with previous methods, our proposed SISK demonstrates superior performance on all datasets. Specifically, on the Laptop dataset, our method improves accuracy by 3.29%, 0.62%, 0.49%, and 1.05% over RGAT, T-GCN, SSEGCN, and CLEAN, respectively. Comparable enhancements are also evident in the Restaurant and Twitter datasets, underscoring the robustness of our approach.

In contrast to previous approaches that just concentrate on single type knowledge, our approach integrates the analysis of semantic information with syntactic knowledge. Specifically, our SISK harnesses the self-attention mechanism to elucidate aspect-aware semantic features within the sentence, while also employing GCNs to extract aspect-oriented syntactic information with dependency trees. Our model achieves remarkable performance improvements, which illustrate that the extraction of aspect-insight semantic and syntactic information is beneficial for improving model performance in fine-grained sentiment analysis. Furthermore, to verify whether our model effectively combines semantic information and syntactic knowledge to enhance performance, we conduct case studies on several sample sentences in Table III.

C. Ablation Study

In this subsection, we perform ablation studies on three datasets to analyze the importance of each component. The results, presented in Table IV, reveal that each element plays significant influence in model's performance. From the results of "w/o Self-Attention", we can observe that utilizing

TABLE III: We compare our proposed SISK with IAN and ASGCN on several samples. The aspects within the sentences are highlighted in bold. The results are annotated with ✓ to indicate correct predictions and × to indicate incorrect predictions.

Sentence	IAN	ASGCN	SISK
1) The **price** is really reasonable though the **service** is a little poor.	(✓, ×)	(✓, ✓)	(✓, ✓)
2) From the speed to the multi touch gestures, I think this **operating system** beats **Windows** easily .	(✓, ×)	(✓, ×)	(✓, ✓)
3) We upgraded the **memory** to four gigabytes in order to take advantage of the performace increase in **speed**.	(×, ✓)	(×, ✓)	(✓, ✓)
4) Usually the **waiters** are kind enough to split the **dish** in half.	(✓, ×)	(✓, ×)	(✓, ✓)

TABLE IV: Ablation experiment results (%) on Laptop, Restaurant, and Twitter Datasets.

Models	Laptop		Restaurant		Twitter	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
SISK	**81.50**	**78.36**	**87.32**	**82.23**	**77.60**	**76.61**
w/o Self-Attention	79.47	75.11	85.80	77.99	75.58	74.77
w/o GCN	79.94	75.94	86.16	79.95	75.72	75.56
w/o position	80.88	77.81	86.79	80.05	76.73	76.53
w/o aspect	81.03	77.57	86.70	80.43	76.45	75.40
w/o fusion	80.72	77.50	86.88	80.28	76.59	75.72

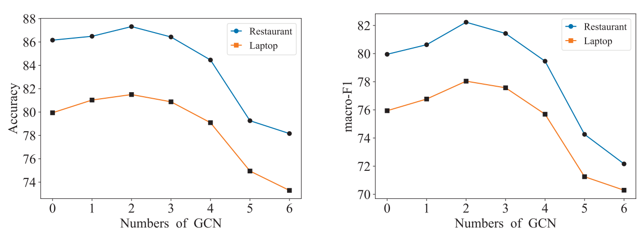

Fig. 2: Accuracy (%) and F1 value (%) on Restaurant and Laptop datasets with different GCN layers.

self-attention to exploring semantics significantly enhances model performance. The results observed in the "w/o GCN" underscore the critical importance of syntactic knowledge and word dependencies in ABSA. Similarly, the performance declines in the "w/o position" and "w/o aspect" illustrate the efficacy of re-weighting features based on aspect position and employing aspect features to guide the extraction of syntactic information in mitigating noise. Finally, the results in "w/o fusion" demonstrate that combining semantic and syntactic information significantly enhances the model's performance.

D. Impact of GCN Layers

To evaluate the impact of GCN layer count on model performance, we test the model on the Restaurant and Laptop datasets with varying GCN layers. As shown in Figures 2, the model's performance initially improves as the number of layers increases, reaching an optimal point before subsequently declining. The highest performance on both datasets is achieved with two GCN layers. Specifically, with a single GCN layer, the propagation range of node information is constrained, hindering SISK's ability to capture long-distance dependencies and thoroughly extract syntactic knowledge. Conversely, an excessive number of GCN layers leads to overly smoothed node representations, resulting in gradient vanishing and information redundancy, which ultimately degrades model performance.

IV. Conclusion

In this paper, we propose a novel framework SISK, which integrates both semantics and syntactic information for the fine-grained classification task. Specifically, we first design a semantic extraction module to capture semantic features with internal introspection and contextual interaction. In addition, we re-weight the features based on the position of the aspect to get aspect-aware semantic representations. Then, We introduce syntactic dependency trees and employ GCNs to extract the syntactic knowledge of the sentence. Notably, we guide the feature extraction process with aspect features to obtain aspect-oriented syntactic information. Finally, we interactively combine the semantic and syntactic features for the final sentiment classification. Extensive experiments on three public datasets show that our method outperform a range of advanced models.

V. Acknowledgements

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

REFERENCES

[1] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent Twitter sentiment classification," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, D. Lin, Y. Matsumoto, and R. Mihalcea, Eds. Portland, Oregon, USA: Association for Computational Linguistics, Jun. 2011, pp. 151–160. [Online]. Available: https://aclanthology.org/P11-1016

[2] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "SemEval-2014 task 4: Aspect based sentiment analysis," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, P. Nakov and T. Zesch, Eds. Dublin, Ireland: Association for Computational Linguistics, Aug. 2014, pp. 27–35. [Online]. Available: https://aclanthology.org/S14-2004

[3] S. Tang, H. Chai, Z. Yao, Y. Ding, C. Gao, B. Fang, and Q. Liao, "Affective knowledge enhanced multiple-graph fusion networks for aspect-based sentiment analysis," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 5352–5362. [Online]. Available: https://aclanthology.org/2022.emnlp-main.359

[4] C. Chen, Z. Teng, Z. Wang, and Y. Zhang, "Discrete opinion tree induction for aspect-based sentiment analysis," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, S. Muresan, P. Nakov, and A. Villavicencio, Eds. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 2051–2064. [Online]. Available: https://aclanthology.org/2022.acl-long.145

[5] X. Sun, K. Zhang, Q. Liu, M. Bao, and Y. Chen, "Harnessing domain insights: A prompt knowledge tuning method for aspect-based sentiment analysis," *Knowledge-Based Systems*, p. 111975, 2024.

[6] I. Titov and R. McDonald, "Modeling online reviews with multi-grain topic models," in *Proceedings of the 17th International Conference on World Wide Web*, ser. WWW '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 111–120. [Online]. Available: https://doi.org/10.1145/1367497.1367513

[7] S. Kiritchenko, X. Zhu, C. Cherry, and S. Mohammad, "NRC-Canada-2014: Detecting aspects and sentiment in customer reviews," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, P. Nakov and T. Zesch, Eds. Dublin, Ireland: Association for Computational Linguistics, Aug. 2014, pp. 437–442. [Online]. Available: https://aclanthology.org/S14-2076

[8] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, J. Su, K. Duh, and X. Carreras, Eds. Austin, Texas: Association for Computational Linguistics, Nov. 2016, pp. 606–615. [Online]. Available: https://aclanthology.org/D16-1058

[9] D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, ser. IJCAI'17. AAAI Press, 2017, p. 4068–4074.

[10] P. Chen, Z. Sun, L. Bing, and W. Yang, "Recurrent attention network on memory for aspect sentiment analysis," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, M. Palmer, R. Hwa, and S. Riedel, Eds. Copenhagen, Denmark: Association for Computational Linguistics, Sep. 2017, pp. 452–461. [Online]. Available: https://aclanthology.org/D17-1047

[11] F. Fan, Y. Feng, and D. Zhao, "Multi-grained attention network for aspect-level sentiment classification," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, E. Riloff, D. Chiang, J. Hockenmaier, and J. Tsujii, Eds. Brussels, Belgium: Association for Computational Linguistics, Oct.-Nov. 2018, pp. 3433–3442. [Online]. Available: https://aclanthology.org/D18-1380

[12] S. Gu, L. Zhang, Y. Hou, and Y. Song, "A position-aware bidirectional attention network for aspect-level sentiment analysis," in *Proceedings of the 27th International Conference on Computational Linguistics*, E. M. Bender, L. Derczynski, and P. Isabelle, Eds. Santa Fe, New Mexico, USA: Association for Computational Linguistics, Aug. 2018, pp. 774–784. [Online]. Available: https://aclanthology.org/C18-1066

[13] C. Du, H. Sun, J. Wang, Q. Qi, J. Liao, T. Xu, and M. Liu, "Capsule network with interactive attention for aspect-level sentiment classification," 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)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 5489–5498. [Online]. Available: https://aclanthology.org/D19-1551

[14] K. Zhang, H. Zhang, Q. Liu, H. Zhao, H. Zhu, and E. Chen, "Interactive attention transfer network for cross-domain sentiment classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 5773–5780.

[15] X. Tan, Y. Cai, and C. Zhu, "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)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 3426–3431. [Online]. Available: https://aclanthology.org/D19-1342

[16] K. Zhang, Q. Liu, H. Qian, B. Xiang, Q. Cui, J. Zhou, and E. Chen, "Eatn: An efficient adaptive transfer network for aspect-level sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 377–389, 2021.

[17] C. Zhang, Q. Li, and D. Song, "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)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 4568–4578. [Online]. Available: https://aclanthology.org/D19-1464

[18] K. Wang, W. Shen, Y. Yang, X. Quan, and R. Wang, "Relational graph attention network for aspect-based sentiment analysis," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds. Online: Association for Computational Linguistics, Jul. 2020, pp. 3229–3238. [Online]. Available: https://aclanthology.org/2020.acl-main.295

[19] C. Chen, Z. Teng, and Y. Zhang, "Inducing target-specific latent structures for aspect sentiment classification," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, B. Webber, T. Cohn, Y. He, and Y. Liu, Eds. Online: Association for Computational Linguistics, Nov. 2020, pp. 5596–5607. [Online]. Available: https://aclanthology.org/2020.emnlp-main.451

[20] R. Li, H. Chen, F. Feng, Z. Ma, X. Wang, and E. Hovy, "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)*, C. Zong, F. Xia, W. Li, and R. Navigli, Eds. Online: Association for Computational Linguistics, Aug. 2021, pp. 6319–6329. [Online]. Available: https://aclanthology.org/2021.acl-long.494

[21] Y. Zhou, L. Liao, Y. Gao, Z. Jie, and W. Lu, "To be closer: Learning to link up aspects with opinions," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, M.-F. Moens, X. Huang, L. Specia, and S. W.-t. Yih, Eds. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 3899–3909. [Online]. Available: https://aclanthology.org/2021.emnlp-main.317

[22] Z. Zhang, Z. Zhou, and Y. Wang, "SSEGCN: Syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, M. Carpuat, M.-C. de Marneffe, and I. V. Meza Ruiz, Eds. Seattle, United States: Association for Computational Linguistics, Jul. 2022, pp. 4916–4925. [Online]. Available: https://aclanthology.org/2022.naacl-main.362

[23] B. Liang, H. Su, L. Gui, E. Cambria, and R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks," *Knowledge-Based Systems*, vol. 235, p. 107643, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705121009059

[24] K. Zhang, K. Zhang, M. Zhang, H. Zhao, Q. Liu, W. Wu, and E. Chen, "Incorporating dynamic semantics into pre-trained language model for aspect-based sentiment analysis," in *Findings of the Association for Computational Linguistics: ACL 2022*, S. Muresan,

P. Nakov, and A. Villavicencio, Eds. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 3599–3610. [Online]. Available: https://aclanthology.org/2022.findings-acl.285

[25] M. Zhang, Y. Zhu, Z. Liu, Z. Bao, Y. Wu, X. Sun, and L. Xu, "Span-level aspect-based sentiment analysis via table filling," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, A. Rogers, J. Boyd-Graber, and N. Okazaki, Eds. Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 9273–9284. [Online]. Available: https://aclanthology.org/2023.acl-long.515

[26] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "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: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[27] X. Li, L. Bing, W. Lam, and B. Shi, "Transformation networks for target-oriented sentiment classification," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, I. Gurevych and Y. Miyao, Eds. Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 946–956. [Online]. Available: https://aclanthology.org/P18-1087

[28] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, "Attentional encoder network for targeted sentiment classification," in *International Conference on Artificial Neural Networks*, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:67855317

[29] L. Xu, X. Pang, J. Wu, M. Cai, and J. Peng, "Learn from structural scope: Improving aspect-level sentiment analysis with hybrid graph convolutional networks," *Neurocomputing*, vol. 518, pp. 373–383, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231222013522

[30] Y. Tian, G. Chen, and Y. Song, "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*, K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, and Y. Zhou, Eds. Online: Association for Computational Linguistics, Jun. 2021, pp. 2910–2922. [Online]. Available: https://aclanthology.org/2021.naacl-main.231

[31] S. Feng, B. Wang, Z. Yang, and J. Ouyang, "Aspect-based sentiment analysis with attention-assisted graph and variational sentence representation," *Knowledge-Based Systems*, vol. 258, p. 109975, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705122010681

[32] S. Wang, J. Zhou, C. Sun, J. Ye, T. Gui, Q. Zhang, and X. Huang, "Causal intervention improves implicit sentiment analysis," in *Proceedings of the 29th International Conference on Computational Linguistics*, N. Calzolari, C.-R. Huang, H. Kim, J. Pustejovsky, L. Wanner, K.-S. Choi, P.-M. Ryu, H.-H. Chen, L. Donatelli, H. Ji, S. Kurohashi, P. Paggio, N. Xue, S. Kim, Y. Hahm, Z. He, T. K. Lee, E. Santus, F. Bond, and S.-H. Na, Eds. Gyeongju, Republic of Korea: International Committee on Computational Linguistics, Oct. 2022, pp. 6966–6977. [Online]. Available: https://aclanthology.org/2022.coling-1.607

[33] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "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)*, K. Toutanova and H. Wu, Eds. Baltimore, Maryland: Association for Computational Linguistics, Jun. 2014, pp. 49–54. [Online]. Available: https://aclanthology.org/P14-2009