Article

Not peer-reviewed version

SentiSyn: Modeling Structure-Enhanced Graph Networks for Aspect-Level Sentiment Analysis

Chewson David , Nicolas Usunier , Fei Li * , Woods Ali

Posted Date: 27 November 2023

doi: 10.20944/preprints202311.1653.v1

Keywords: aspect-level sentiment analysis; syntax-enhanced graph networks; dependency-based sentiment modeling

Article

SentiSyn: Modeling Structure-Enhanced Graph Networks for Aspect-Level Sentiment Analysis

Chewson David [1], Nicolas Usunier [1], Fei Li [2,*] and Woods Ali [1]

[1] University of Charleston
[2] National University of Singapore
* Correspondence: feih@comp.nus.edu.sg

Abstract: In this paper, we delve into the realm of aspect-level sentiment analysis, a sophisticated analytical task focused on pinpointing and interpreting the sentiment directed towards specific elements within a sentence. Traditional methods in this domain, primarily based on neural networks, have often overlooked the critical role of syntactic structures in sentences. To bridge this gap, we have developed the Syntax-Enhanced Sentiment Graph Network (SentiSyn). This pioneering model represents a significant step forward in aspect-level sentiment analysis, bringing to the forefront the utilization of word dependency relationships to enrich sentiment analysis. SentiSyn stands out by its innovative use of a dependency graph, a tool that meticulously maps out the intricate web of syntactic relationships surrounding a target aspect in a sentence. This approach allows SentiSyn to effectively capture and channel sentiment-related characteristics that are deeply rooted in the syntactic context of the aspect target. By doing so, SentiSyn unlocks a deeper understanding of sentiment dynamics in textual content, enabling a more nuanced and accurate sentiment analysis. Our comprehensive experimental evaluation of SentiSyn showcases its remarkable capabilities. When combined with advanced embedding techniques like GloVe and BERT, SentiSyn demonstrates a superior performance edge over several existing sentiment analysis methods. This performance leap is not just incremental; it represents a significant enhancement in the field of sentiment analysis, underscoring the importance of syntactic context in understanding sentiments. Furthermore, our analysis delves into how SentiSyn effectively leverages these embeddings to gain a more profound and contextually rich insight into sentiment dynamics. The results from our tests indicate that SentiSyn, with its unique approach to integrating syntactic structures and advanced embeddings, sets a new benchmark in aspect-level sentiment analysis, offering both enhanced accuracy and deeper sentiment understanding.

Keywords: aspect-level sentiment analysis; syntax-enhanced graph networks; dependency-based sentiment modeling

---

1. Introduction

Aspect-based sentiment analysis is an intricate process that involves discerning specific sentiments – positive, negative, or neutral – associated with distinct aspects within a textual context. This nuanced analysis stands in contrast to the more general sentence-level sentiment analysis, which broadly categorizes the entire sentence's sentiment. Aspect-level analysis excels in situations where sentences contain mixed sentiments about different elements [1–4]. For example, a statement like "The pasta was excellent, but the service left much to be desired" showcases this complexity by simultaneously expressing positive sentiments towards the food and negative sentiments towards the service.

In the early stages of aspect-level sentiment analysis, the primary approach involved the utilization of manually designed features, such as sentiment lexicons and various linguistic indicators. These tools were pivotal in identifying and classifying sentiments related to specific aspects. However, the emergence of advanced neural network techniques, especially those built upon Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) architectures, marked a significant shift in the field [24]. These modern approaches largely focus on processing sentences as sequences of

 doi:10.20944/preprints202311.1653.v1

words, embedding critical aspect-related information into sentence representations through innovative methods like attention mechanisms and gating techniques. Despite their advancements, a common shortfall of these models is their tendency to overlook the syntactic structure of sentences [25–28]. This oversight is notable as syntactic nuances play a vital role in accurately identifying and linking sentiment features directly to their respective aspects [29,30].

To address this gap, our research introduces a groundbreaking framework, the Syntax-Enhanced Sentiment Graph Network (SentiSyn). This model is a significant departure from traditional methods as it conceptualizes sentences as dependency graphs [24,31]. This representation forms direct connections between aspect targets and the words relevant to them, thus preserving the natural syntactic relationships within the sentence [4,9,32–34]. In SentiSyn, a graph attention network, augmented with an LSTM unit, is employed to dynamically and effectively propagate sentiment attributes from syntactically significant neighboring words to the targeted aspect. This method ensures that the inherent syntactic order and connections within the sentence are maintained, providing a distinct advantage over previous models that necessitated alterations to the sentence's syntactic structure.

To empirically validate the effectiveness of SentiSyn, we conducted extensive experiments using datasets from SemEval 2014 [49], focusing on laptop and restaurant reviews. These tests were designed to compare the performance of SentiSyn against several established baselines, particularly in scenarios involving GloVe embeddings [50]. Furthermore, we explored the impact of integrating BERT representations [51] into our model, which resulted in a notable enhancement of SentiSyn's performance. Our comprehensive analysis reveals that SentiSyn is not only highly effective in its sentiment analysis capabilities but also exhibits efficiency in terms of computational resource utilization and processing speed. This efficiency makes SentiSyn a compelling alternative to the direct fine-tuning of large models like BERT, especially in resource-constrained environments.

In summary, the Syntax-Enhanced Sentiment Graph Network represents a significant advancement in the field of aspect-based sentiment analysis. By seamlessly integrating syntactic structures into the sentiment analysis process, SentiSyn opens up new possibilities for more accurate and nuanced sentiment analysis, particularly in complex textual scenarios where multiple sentiments coexist.

2. Related Work

Aspect-level sentiment classification, a specialized subfield of sentiment analysis, aims to discern the sentiment polarity associated with specific aspect targets within contextual sentences [52]. This task necessitates a nuanced understanding of the interplay between language elements and sentiment expressions [9,13,53–59]. Initial approaches in this field heavily relied on transforming a broad array of features, such as sentiment lexicons and parsing contexts, into feature vectors. These vectors were then utilized to train classifiers, typically Support Vector Machines (SVMs). Pioneering works like those of Wagner *et al.* [31] integrated sentiment lexicons with aspect proximity and dependency path distances to enhance SVM classifier training. Kiritchenko *et al.* [60] extended this approach, demonstrating that incorporating parsing context features could significantly increase predictive accuracy.

The advent of neural network methodologies marked a paradigm shift in aspect-level sentiment analysis. LSTM neural networks became prevalent, modeling word sequences within sentences to capture the nuanced sentiment dynamics. Tang *et al.* [9] innovatively employed dual LSTMs to process the context surrounding an aspect target, leveraging the final hidden states as features for classification. Building on this, Wang *et al.* [30] introduced an attention mechanism, inspired by [61], to prioritize aspect-relevant words in sentences. This methodology was further refined by Huang *et al.* [62], who used dual LSTM networks to jointly model sentence and aspect interactions, extracting critical words from the resulting sentence-aspect correlation matrix. Li *et al.* [58] advanced these attention-based models by integrating positional information, enhancing the precision of sentiment analysis.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 November 2023 doi:10.20944/preprints202311.1653.v1

Beyond LSTM-based models, the literature also records the use of deep memory networks, as proposed by Tang *et al.* [29]. These networks feature multiple computational layers, each generating an attention vector over an external memory, showcasing an alternative neural approach. Additionally, some researchers have explored using Convolutional Neural Networks (CNN) for this task [63,64]. In these models, features from the aspect influence the information flow within the CNN processing the sentence [62]. The integration of BERT representations, benefiting from extensive linguistic knowledge obtained through large-scale language modeling [51], has shown considerable progress in this domain [65]. Xu *et al.* [66] achieved notable results by post-training BERT on domain-specific datasets and fine-tuning it, although this required substantial computational resources and time.

Our approach diverges from these neural network-based methods by explicitly leveraging the syntactic structure within sentences. This method propagates sentiment features towards the aspect target along a dependency graph, rather than following the original word sequence. Previous attempts, such as those by Dong *et al.* [4] and Nguyen and Shirai [67], also focused on syntax but required transforming the dependency tree into a binary format and positioning the aspect target at the root. This often led to the displacement of sentiment-modifying words away from the aspect target. In contrast, our methodology retains the original syntactic order, ensuring a more accurate and contextually relevant sentiment analysis.

3. Method

3.1. Text Representation

In our approach, a sentence $s = [w_1, w_2, ..., w_i, ..., w_n]$ of length n, containing an aspect target w_i, is transformed into a vector representation. Each word w_i in the sentence is mapped to a vector $x_i \in \mathbb{R}^d$, where d denotes the dimensionality of the embedding space. We employ a standard dependency parser [68] to convert the sentence into a dependency graph, where nodes correspond to words linked by syntactic dependencies. This graph structure enables us to propagate features from the neighborhood of an aspect target. For instance, the sentence “delivery was early too” is depicted as a dependency graph, highlighting the connectivity and feature propagation pathways around the aspect “delivery”. In cases where an aspect target comprises multiple words, we substitute the entire sequence with a placeholder “__target__” before parsing. This process results in a single node in the dependency graph, representing the entire aspect target. The feature vector for this node is computed as the average of the embedding vectors of the constituent words of the aspect target.

3.2. Graph Attention Network

We leverage a Graph Attention Network (GAT) [69], a variant of the graph neural network [70], as a core component of our model, SentiSyn. This network propagates features from a node’s syntactic context to the node representing the aspect target in the dependency graph. For a graph with N nodes, each associated with an embedding vector x, a GAT layer aggregates information from the hidden states of neighboring nodes. An L-layer GAT enables the propagation of features from nodes up to L hops away to the aspect target node.

The GAT updates the hidden state of node i at layer $l+1$ using multi-head attentions [71], formulated as follows:

$$h_{l+1}^{i} = \mathop{\Big\Vert}_{k=1}^{K} \sigma\left(\sum_{j\in n[i]} \alpha_{lk}^{ij} W_{lk} h_l^j\right) \tag{1}$$

$$\alpha_{lk}^{ij} = \frac{\exp\left(f\left(a_{lk}^{T}\left[W_{lk}h_l^i \parallel W_{lk}h_l^j\right]\right)\right)}{\sum_{u\in n[i]}\exp\left(f\left(a_{lk}^{T}\left[W_{lk}h_l^i \parallel W_{lk}h_l^u\right]\right)\right)} \tag{2}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 November 2023 doi:10.20944/preprints202311.1653.v1

Here, $\|$ is the vector concatenation operation, α_{lk}^{ij} denotes the attention coefficient from node i to its neighbor j in the k-th attention head at layer l, and $W_{lk} \in \mathbb{R}^{\frac{D}{K} \times D}$ is a linear transformation matrix for input states. D represents the dimension of hidden states, σ is the sigmoid function, and $f(\cdot)$ is the LeakyReLU activation function [72]. The attention context vector $a_{lk} \in \mathbb{R}^{\frac{2D}{K}}$ is learned during training.

We simplify this feature propagation as:

$$H_{l+1} = \mathrm{GAT}(H_l, A; \Theta_l) \tag{3}$$

where $H_l \in \mathbb{R}^{N \times D}$ is the stacked states for all nodes at layer l, $A \in \mathbb{R}^{N \times N}$ is the adjacency matrix of the graph, and Θ_l is the parameter set of the GAT at layer l.

3.3. Target-Dependent Graph Attention Network (SentiSyn)

Our model, SentiSyn, incorporates an LSTM to model the target-specific dependencies across layers, which is essential for filtering out noise in the graph [73**?**]. The approach ensures that at layer 0, the hidden state of a target node h_0^t solely depends on its local features, and at each subsequent layer l, information from the l-hop neighborhood relevant to the target is incrementally integrated into the hidden state via the LSTM unit. The hidden and cell states of the LSTM for a target node t are updated as follows, starting from the temporary hidden state $\hat{h}_l^t$:

$$i_l = \sigma(W_i \hat{h}_l^t + U_i h_{l-1} + b_i) \tag{4}$$

$$f_l = \sigma(W_f \hat{h}_l^t + U_f h_{l-1} + b_f) \tag{5}$$

$$o_l = \sigma(W_o \hat{h}_l^t + U_o h_{l-1} + b_o) \tag{6}$$

$$\hat{c}_l = \tanh(W_c \hat{h}_l^t + U_c h_{l-1} + b_c) \tag{7}$$

$$c_l = f_l \circ c_{l-1} + i_l \circ \hat{c}_l \tag{8}$$

$$h_l = o_l \circ \tanh(c_l) \tag{9}$$

where $\sigma(\cdot)$ is the sigmoid function, $\tanh(\cdot)$ is the hyperbolic tangent function, and $W_i, U_i, W_f, U_f, W_o, U_o, W_c, U_c$ are parameter matrices, with b_i, b_f, b_o, b_c as bias vectors. $\circ$ represents element-wise multiplication, and i_l, f_l, o_l are the input, forget, and output gates, respectively.

The feed-forward process of SentiSyn is summarized as:

$$\begin{aligned} H_{l+1}, C_{l+1} &= \mathrm{LSTM}(\mathrm{GAT}(H_l, A; \Theta_l), (H_l, C_l)) \\ H_0, C_0 &= \mathrm{LSTM}(XW_p + [b_p]_N, (0, 0)) \end{aligned}$$

where C_l are the stacked cell states of the LSTM at layer l. The initial hidden state and cell state of the LSTM are set to 0. $W_p \in \mathbb{R}^{d \times D}$ projects the stacked embedding vectors X into the hidden state dimension, and $[b_p]_N$ denotes the stacking of the bias vector b_p N times.

3.4. Final Classification

After processing through L layers of SentiSyn, we extract the final representation h_L^t for the aspect target from the node representations H_L. This representation is then linearly transformed for classification:

$$P(y = c) = \frac{\exp(W h_L^t + b)_c}{\sum_{i \in C} \exp(W h_L^t + b)_i} \tag{10}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 November 2023 doi:10.20944/preprints202311.1653.v1

Here, W and b are the weight matrix and bias for the linear transformation, and C is the set of sentiment classes. The model predicts the sentiment polarity of the aspect target as the class with the highest probability.

We optimize our model using cross-entropy loss with L_2 regularization:

$$\text{loss} = -\sum_{c \in C} I(y = c) \cdot \log(P(y = c)) + \lambda \|\Theta\|^2$$

where $I(\cdot)$ is an indicator function, λ is the L_2 regularization coefficient, and Θ represents all model parameters.

4. Experiments

4.1. Datasets

For evaluating the performance of SentiSyn, our novel approach, we utilized two domain-specific datasets from SemEval 2014 Task 4 [49]. These datasets encompass reviews from laptops and restaurants, with each data point comprising a sentence and an associated aspect term labeled with sentiment polarity by expert annotators. Following the methodology of [63,74], we initially allocated 500 training instances as a development set[1] for model optimization, subsequently merging this with the training dataset for final model training. The composition of these datasets is detailed in Table 1.

Table 1. Dataset distribution across sentiment categories.

Dataset	Positive	Neutral	Negative
Laptop-Training	767	373	673
Laptop-Development	220	87	193
Laptop-Testing	341	169	128
Restaurant-Training	1886	531	685
Restaurant-Development	278	102	120
Restaurant-Testing	728	196	196

4.2. Implementation Details

Our dependency graphs are generated using the Stanford neural parser [68]. We explore two embedding methods: 300-dimensional GloVe embeddings [50] and BERT representations [51], using the large uncased English model implemented in PyTorch[2]. BERT's input format consists of a sentence-aspect pair, and we extract sentence representations for aspect-level sentiment analysis. Due to differences in tokenization between the parser and BERT, we average BERT's sub-word unit representations to obtain embeddings for dependency graph tokens.

For hidden state dimensions, we use 300, and BERT representations are mapped to this dimensionality through linear projection. SentiSyn employs 6 attention heads and is trained with a batch size of 32, applying l_2 regularization (term λ 10^{-4}) and dropout [75] at a rate of 0.7 on input embeddings. We initially utilize the Adam [76] optimizer with a learning rate of 10^{-3}, followed by stochastic gradient descent for fine-tuning and model stabilization.

SentiSyn is implemented using PyTorch Geometric [77] on a Linux setup with Titan XP GPUs.

[1] Development set splits available at https://github.com/vanzytay/ABSA_DevSplits.
[2] PyTorch BERT implementation: https://github.com/huggingface/pytorch-pretrained-BERT

4.3. Baseline Comparisons

SentiSyn's performance is compared against various established methods:

SVM with Feature Engineering employs n-gram, parsing, and lexicon features for aspect-level sentiment analysis [60].

Context-LSTM (TD-LSTM) models context around the aspect using two LSTM networks. In contrast, SentiSyn uses GAT to incorporate syntax context. The sentiment prediction leverages the concatenated final hidden states of the LSTMs [9].

Attention-LSTM (AT-LSTM) employs a LSTM model for sentence representation and combines this with aspect embeddings to generate an attention vector, using the weighted sum of hidden states for the final representation [30].

Memory Network (MemNet) applies repeated attention to word embeddings, with the last attention output used for prediction [29].

Interactive Attention Network (IAN) models both sentence and aspect using LSTM networks, generating mutual attention vectors for target and sentence representations [11].

Parse-Gated CNN (PG-CNN) uses aspect features as gates in a CNN for sentence feature extraction [63].

AOA-LSTM introduces an attention-over-attention network for joint modeling of aspects and sentences [62].

BERT-AVG and BERT-CLS respectively use average sentence representations and the "[CLS]" token representation from BERT for training and fine-tuning.

Table 2 illustrates that SentiSyn, with both GloVe and BERT embeddings, outperforms existing methods. Feature-based SVM's strong performance underscores the significance of feature engineering and syntax understanding. SentiSyn's superior performance compared to Context-LSTM validates the importance of syntactic context. BERT-AVG and BERT-CLS, particularly after fine-tuning, show remarkable results, though the fine-tuning process can be unstable. SentiSyn enhances the predictive power of BERT representations, achieving accuracy rates around 80% and 83% for laptops and restaurants, respectively.

Table 2. Performance comparison of SentiSyn and other methods on laptop and restaurant datasets. Layer numbers in parentheses indicate SentiSyn's configuration.

	Laptop	Restaurant
Feature+SVM	70.5	80.2
Context-LSTM	68.1	75.6
Attention-LSTM	68.9	76.2
MemNet	72.4	80.3
IAN	72.1	78.6
PG-CNN	69.1	78.9
AOA-LSTM	72.6	79.7
SentiSyn-GloVe (3)	73.7	81.1
SentiSyn-GloVe (4)	**74.0**	80.6
SentiSyn-GloVe (5)	73.4	**81.2**
BERT-AVG	76.5	78.7
BERT-CLS	77.1	81.2
SentiSyn-BERT (3)	79.3	82.9
SentiSyn-BERT (4)	79.8	**83.0**
SentiSyn-BERT (5)	**80.1**	82.8

4.4. Target Information Impact

Our ablation study evaluates the influence of explicitly capturing target information in SentiSyn. Removing the LSTM unit in SentiSyn, denoted as GAT here, disables explicit target information utilization. Results in Table 3 demonstrate that explicit target capturing consistently boosts SentiSyn's performance over the GAT model. On average, accuracy improvements of 1.2 and 0.95 percentage points are observed for GloVe and BERT variants, respectively.

Table 3. Ablation study highlighting the benefits of explicit target information in SentiSyn.

Dataset	Laptop			Restaurant		
layer	3	4	5	3	4	5
GAT-GloVe	73.0	72.1	72.4	79.6	80.0	79.7
SentiSyn-GloVe	73.7	74.0	73.4	81.1	80.6	81.2
GAT-BERT	78.1	78.5	78.5	82.6	82.2	82.3
SentiSyn-BERT	79.3	79.8	80.1	82.9	83.0	82.8

4.5. Exploring Model Layer Impact

This section delves into the influence of the number of layers in our SentiSyn model, experimenting with depths ranging from 1 to 6 layers. And we observe that a single-layer SentiSyn model, when integrated with GloVe embeddings, underperforms. This suggests that the crucial sentiment words related to the aspect targets are typically more than one hop away. Enhancing the model depth to three layers markedly ameliorates SentiSyn's performance with GloVe embeddings. In contrast, SentiSyn augmented with BERT representations displays greater depth robustness. Even a solitary layer in SentiSyn, when combined with BERT, yields satisfactory outcomes on both datasets. This can be attributed to BERT's inherent ability to embed contextual information into its representations. Nevertheless, further depth augmentation continues to refine performance, with optimal results achieved when the model depth exceeds three layers.

4.6. Comparative Analysis of Model Sizes

In Table 4, we compare the model size of our SentiSyn framework with several baseline models and the BERT model. For baselines, we utilize a publicly available PyTorch implementation for size assessment. Our SentiSyn model, when integrated with GloVe embeddings, demonstrates a smaller footprint compared to the LSTM-based counterparts. MemNet (3) exhibits the smallest model size. The incorporation of BERT representations in SentiSyn leads to a marginal size increase, primarily due to the additional linear projection layer for input word representation adaptation. Notably, the shift from GloVe to BERT embeddings only slightly increases the training duration per epoch for a three-layer SentiSyn model on the restaurant dataset, from 1.12 seconds to 1.15 seconds per epoch. In stark contrast, fine-tuning the full BERT model necessitates substantially more time, approximately 226.50 seconds per epoch, underscoring the efficiency of the SentiSyn model in computational resource utilization and training time.

 doi:10.20944/preprints202311.1653.v1

Table 4. Model size comparison of SentiSyn with various configurations and baseline models.

Models	Model size ($\times 10^6$)
TD-LSTM	1.45
MemNet (3)	**0.36**
IAN	2.17
AOA-LSTM	2.89
SentiSyn-GloVe (3)	1.00
SentiSyn-GloVe (4)	1.09
SentiSyn-GloVe (5)	1.18
BERT-CLS	335.14
SentiSyn-BERT (3)	**1.30**
SentiSyn-BERT (4)	1.39
SentiSyn-BERT (5)	1.49

5. Conclusion

In our research, we introduced a pioneering graph attention network, SentiSyn, specifically tailored for aspect-level sentiment analysis. This innovative approach harnesses the intricate syntactic dependencies within sentences, focusing on the contextual syntax around aspect targets for more precise sentiment classification. Unlike conventional models that process word sequences linearly, SentiSyn brings sentiment-modifying words into closer association with their relevant aspect targets, adeptly navigating through possible syntactic complexities. Our extensive evaluations, conducted on laptop and restaurant review datasets from SemEval 2014, have showcased SentiSyn's superior capabilities. When integrating GloVe embeddings, SentiSyn notably surpassed various existing models in performance. Upon adopting BERT representations, SentiSyn's efficiency was further amplified, delivering enhanced outcomes. Remarkably, SentiSyn achieves these results with a leaner architecture, demanding less computational power and training time compared to the extensive fine-tuning required for the original BERT model.

This work is arguably the first to directly utilize an unaltered dependency graph in aspect-level sentiment analysis, opening new avenues in this research area. However, there is ample scope for refinement. Future iterations could explore the incorporation of an attention mechanism specifically to weigh the significance of individual words within an aspect. Additionally, this study's focus on dependency graphs alone leaves room for integrating various relation types present in these graphs. Incorporating elements such as part-of-speech tags could provide a more nuanced analysis. Finally, amalgamating our graph-based approach with sequence-based models could offer a comprehensive solution, potentially mitigating inaccuracies originating from dependency parsing errors, thereby further enhancing the robustness and accuracy of sentiment analysis.

References

1. Nazir, A.; Rao, Y.; Wu, L.; Sun, L. Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. *IEEE Transactions on Affective Computing* **2020**, *13*, 845–863.
2. Fei, H.; Wu, S.; Li, J.; Li, B.; Li, F.; Qin, L.; Zhang, M.; Zhang, M.; Chua, T.S. LasUIE: Unifying Information Extraction with Latent Adaptive Structure-aware Generative Language Model. Proceedings of the Advances in Neural Information Processing Systems, NeurIPS 2022, 2022, pp. 15460–15475.
3. Kiritchenko, S.; Zhu, X.; Cherry, C.; Mohammad, S. Detecting aspects and sentiment in customer reviews. 8th International Workshop on Semantic Evaluation (SemEval), pp. 437–442.
4. Dong, L.; Wei, F.; Tan, C.; Tang, D.; Zhou, M.; Xu, K. Adaptive recursive neural network for target-dependent twitter sentiment classification. Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short papers), 2014, Vol. 2, pp. 49–54.

5. Fei, H.; Li, J.; Wu, S.; Li, C.; Ji, D.; Li, F. Global Inference with Explicit Syntactic and Discourse Structures for Dialogue-Level Relation Extraction. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI, 2022, pp. 4082–4088.
6. Wu, S.; Fei, H.; Ren, Y.; Ji, D.; Li, J. Learn from Syntax: Improving Pair-wise Aspect and Opinion Terms Extraction with Rich Syntactic Knowledge. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 2021, pp. 3957–3963.
7. Xiang, C.; Zhang, J.; Li, F.; Fei, H.; Ji, D. A semantic and syntactic enhanced neural model for financial sentiment analysis. *Information Processing & Management* **2022**, *59*, 102943.
8. Fei, H.; Zhang, Y.; Ren, Y.; Ji, D. Latent Emotion Memory for Multi-Label Emotion Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 7692–7699.
9. Tang, D.; Qin, B.; Feng, X.; Liu, T. Effective LSTMs for Target-Dependent Sentiment Classification. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, 2016, pp. 3298–3307.
10. Wu, S.; Fei, H.; Li, F.; Zhang, M.; Liu, Y.; Teng, C.; Ji, D. Mastering the Explicit Opinion-Role Interaction: Syntax-Aided Neural Transition System for Unified Opinion Role Labeling. Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence, 2022, pp. 11513–11521.
11. Ma, D.; Li, S.; Zhang, X.; Wang, H. Interactive Attention Networks for Aspect-Level Sentiment Classification. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2017, pp. 4068–4074.
12. Fei, H.; Zhang, M.; Ji, D. Cross-Lingual Semantic Role Labeling with High-Quality Translated Training Corpus. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 7014–7026.
13. Chen, P.; Sun, Z.; Bing, L.; Yang, W. Recurrent attention network on memory for aspect sentiment analysis. Proceedings of the 2017 conference on empirical methods in natural language processing, 2017, pp. 452–461.
14. Fei, H.; Zhang, M.; Li, B.; Ji, D. End-to-end Semantic Role Labeling with Neural Transition-based Model. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, pp. 12803–12811.
15. Zhang, M.; Qian, T. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis. Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), 2020, pp. 3540–3549.
16. Wu, S.; Fei, H.; Ren, Y.; Li, B.; Li, F.; Ji, D. High-Order Pair-Wise Aspect and Opinion Terms Extraction With Edge-Enhanced Syntactic Graph Convolution. *IEEE ACM Trans. Audio Speech Lang. Process.* **2021**, *29*, 2396–2406.
17. Chen, C.; Teng, Z.; Zhang, Y. Inducing target-specific latent structures for aspect sentiment classification. Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), 2020, pp. 5596–5607.
18. Fei, H.; Wu, S.; Ren, Y.; Zhang, M. Matching Structure for Dual Learning. Proceedings of the International Conference on Machine Learning, ICML, 2022, pp. 6373–6391.
19. Huang, L.; Sun, X.; Li, S.; Zhang, L.; Wang, H. Syntax-aware graph attention network for aspect-level sentiment classification. Proceedings of the 28th international conference on computational linguistics, 2020, pp. 799–810.
20. Fei, H.; Li, F.; Li, B.; Ji, D. Encoder-Decoder Based Unified Semantic Role Labeling with Label-Aware Syntax. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, pp. 12794–12802.
21. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D. Nonautoregressive Encoder-Decoder Neural Framework for End-to-End Aspect-Based Sentiment Triplet Extraction. *IEEE Transactions on Neural Networks and Learning Systems* **2023**, *34*, 5544–5556.
22. Hou, X.; Huang, J.; Wang, G.; He, X.; Zhou, B. Selective attention based graph convolutional networks for aspect-level sentiment classification. *arXiv preprint arXiv:1910.10857* **2019**.
23. Fei, H.; Li, J.; Ren, Y.; Zhang, M.; Ji, D. Making Decision like Human: Joint Aspect Category Sentiment Analysis and Rating Prediction with Fine-to-Coarse Reasoning. Proceedings of the ACM Web Conference 2022, WWW, 2022, pp. 3042–3051.
24. Jiang, L.; Yu, M.; Zhou, M.; Liu, X.; Zhao, T. Target-dependent twitter sentiment classification. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. Association for Computational Linguistics, 2011, pp. 151–160.

 doi:10.20944/preprints202311.1653.v1

25. Huang, B.; Carley, K.M. Syntax-aware aspect level sentiment classification with graph attention networks. *arXiv preprint arXiv:1909.02606* **2019**.
26. Zhang, C.; Li, Q.; Song, D. Aspect-based sentiment classification with aspect-specific graph convolutional networks. *arXiv preprint arXiv:1909.03477* **2019**.
27. Sun, K.; Zhang, R.; Mensah, S.; Mao, Y.; Liu, X. Aspect-level sentiment analysis via convolution over dependency tree. 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, pp. 5679–5688.
28. Wang, K.; Shen, W.; Yang, Y.; Quan, X.; Wang, R. Relational Graph Attention Network for Aspect-based Sentiment Analysis. *arXiv preprint arXiv:2004.12362* **2020**.
29. Tang, D.; Qin, B.; Liu, T. Aspect level sentiment classification with deep memory network. *arXiv preprint arXiv:1605.08900* **2016**.
30. Wang, Y.; Huang, M.; Zhao, L.; others. Attention-based LSTM for aspect-level sentiment classification. Proceedings of the 2016 conference on empirical methods in natural language processing, 2016, pp. 606–615.
31. Wagner, J.; Arora, P.; Cortes, S.; Barman, U.; Bogdanova, D.; Foster, J.; Tounsi, L. Dcu: Aspect-based polarity classification for semeval task 4. Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014), 2014, pp. 223–229.
32. Li, J.; Fei, H.; Liu, J.; Wu, S.; Zhang, M.; Teng, C.; Ji, D.; Li, F. Unified Named Entity Recognition as Word-Word Relation Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, pp. 10965–10973.
33. Fei, H.; Ren, Y.; Zhang, Y.; Ji, D.; Liang, X. Enriching contextualized language model from knowledge graph for biomedical information extraction. *Briefings in Bioinformatics* **2021**, *22*.
34. Fei, H.; Ren, Y.; Ji, D. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction. *Information Processing & Management* **2020**, *57*, 102311.
35. Fei, H.; Ji, D.; Li, B.; Liu, Y.; Ren, Y.; Li, F. Rethinking Boundaries: End-To-End Recognition of Discontinuous Mentions with Pointer Networks. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, pp. 12785–12793.
36. Mukherjee, R.; Shetty, S.; Chattopadhyay, S.; Maji, S.; Datta, S.; Goyal, P. Reproducibility, replicability and beyond: Assessing production readiness of aspect based sentiment analysis in the wild. Advances in Information Retrieval: 43rd European Conference on IR Research, ECIR 2021, Virtual Event, March 28–April 1, 2021, Proceedings, Part II 43. Springer, 2021, pp. 92–106.
37. Fei, H.; Chua, T.; Li, C.; Ji, D.; Zhang, M.; Ren, Y. On the Robustness of Aspect-based Sentiment Analysis: Rethinking Model, Data, and Training. *ACM Transactions on Information Systems* **2023**, *41*, 50:1–50:32.
38. Zhuang, L.; Fei, H.; Hu, P. Knowledge-enhanced event relation extraction via event ontology prompt. *Inf. Fusion* **2023**, *100*, 101919.
39. Fei, H.; Wu, S.; Ren, Y.; Li, F.; Ji, D. Better Combine Them Together! Integrating Syntactic Constituency and Dependency Representations for Semantic Role Labeling. Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, 2021, pp. 549–559.
40. Chen, X.; Sun, C.; Wang, J.; Li, S.; Si, L.; Zhang, M.; Zhou, G. Aspect sentiment classification with document-level sentiment preference modeling. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 3667–3677.
41. Li, J.; Xu, K.; Li, F.; Fei, H.; Ren, Y.; Ji, D. MRN: A Locally and Globally Mention-Based Reasoning Network for Document-Level Relation Extraction. Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 2021, pp. 1359–1370.
42. Liu, J.; Fei, H.; Li, F.; Li, J.; Li, B.; Zhao, L.; Teng, C.; Ji, D. TKDP: Threefold Knowledge-enriched Deep Prompt Tuning for Few-shot Named Entity Recognition. *CoRR* **2023**, *abs/2306.03974*.
43. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S.; AL-Smadi, M.; Al-Ayyoub, M.; Zhao, Y.; Qin, B.; De Clercq, O.; others. Semeval-2016 task 5: Aspect based sentiment analysis. ProWorkshop on Semantic Evaluation (SemEval-2016). Association for Computational Linguistics, 2016, pp. 19–30.
44. Fei, H.; Li, F.; Li, C.; Wu, S.; Li, J.; Ji, D. Inheriting the Wisdom of Predecessors: A Multiplex Cascade Framework for Unified Aspect-based Sentiment Analysis. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI, 2022, pp. 4096–4103.

45. Wang, F.; Li, F.; Fei, H.; Li, J.; Wu, S.; Su, F.; Shi, W.; Ji, D.; Cai, B. Entity-centered Cross-document Relation Extraction. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, 2022, pp. 9871–9881.
46. Fei, H.; Ren, Y.; Ji, D. Retrofitting Structure-aware Transformer Language Model for End Tasks. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 2151–2161.
47. Jia, Y.; Wang, Y.; Zan, H.; Xie, Q. Syntactic information and multiple semantic segments for aspect-based sentiment classification. *International Journal of Asian Language Processing* **2021**, *31*, 2250006.
48. Cao, H.; Li, J.; Su, F.; Li, F.; Fei, H.; Wu, S.; Li, B.; Zhao, L.; Ji, D. OneEE: A One-Stage Framework for Fast Overlapping and Nested Event Extraction. Proceedings of the 29th International Conference on Computational Linguistics, 2022, pp. 1953–1964.
49. Pontiki, M.; Galanis, D.; Pavlopoulos, J.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014); Association for Computational Linguistics and Dublin City University: Dublin, Ireland, 2014; pp. 27–35.
50. Pennington, J.; Socher, R.; Manning, C. Glove: Global vectors for word representation. Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), 2014, pp. 1532–1543.
51. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* **2018**.
52. Pang, B.; Lee, L.; others. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* **2008**, *2*, 1–135.
53. Liu, J.; Zhang, Y. Attention modeling for targeted sentiment. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, 2017, pp. 572–577.
54. Wang, S.; Mazumder, S.; Liu, B.; Zhou, M.; Chang, Y. Target-sensitive memory networks for aspect sentiment classification. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 957–967.
55. Fan, F.; Feng, Y.; Zhao, D. Multi-grained attention network for aspect-level sentiment classification. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 3433–3442.
56. Zheng, S.; Xia, R. Left-Center-Right Separated Neural Network for Aspect-based Sentiment Analysis with Rotatory Attention. *arXiv preprint arXiv:1802.00892* **2018**.
57. Wang, B.; Lu, W. Learning latent opinions for aspect-level sentiment classification. Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
58. Li, L.; Liu, Y.; Zhou, A. Hierarchical Attention Based Position-Aware Network for Aspect-Level Sentiment Analysis. Proceedings of the 22nd Conference on Computational Natural Language Learning, 2018, pp. 181–189.
59. Li, X.; Bing, L.; Lam, W.; Shi, B. Transformation networks for target-oriented sentiment classification. *arXiv preprint arXiv:1805.01086* **2018**.
60. Kiritchenko, S.; Zhu, X.; Cherry, C.; Mohammad, S. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 437–442.
61. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473* **2014**.
62. Huang, B.; Ou, Y.; Carley, K.M. Aspect level sentiment classification with attention-over-attention neural networks. International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Springer, 2018, pp. 197–206.
63. Huang, B.; Carley, K. Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 1091–1096.
64. Xue, W.; Li, T. Aspect based sentiment analysis with gated convolutional networks. *arXiv preprint arXiv:1805.07043* **2018**.
65. Sun, C.; Huang, L.; Qiu, X. Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *arXiv preprint arXiv:1903.09588* **2019**.

 doi:10.20944/preprints202311.1653.v1

66. Xu, H.; Liu, B.; Shu, L.; Yu, P.S. BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis. *arXiv preprint arXiv:1904.02232* **2019**.
67. Nguyen, T.H.; Shirai, K. Phrasernn: Phrase recursive neural network for aspect-based sentiment analysis. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 2509–2514.
68. Chen, D.; Manning, C. A fast and accurate dependency parser using neural networks. Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), 2014, pp. 740–750.
69. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv preprint arXiv:1710.10903* **2017**.
70. Scarselli, F.; Gori, M.; Tsoi, A.C.; Hagenbuchner, M.; Monfardini, G. The graph neural network model. *IEEE Transactions on Neural Networks* **2009**, *20*, 61–80.
71. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. Advances in neural information processing systems, 2017, pp. 5998–6008.
72. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. Proc. icml, 2013, Vol. 30, p. 3.
73. Huang, B.; Carley, K.M. Residual or Gate? Towards Deeper Graph Neural Networks for Inductive Graph Representation Learning. *arXiv preprint arXiv:1904.08035* **2019**.
74. Tay, Y.; Luu, A.T.; Hui, S.C. Learning to Attend via Word-Aspect Associative Fusion for Aspect-based Sentiment Analysis. AAAI, 2018.
75. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* **2014**, *15*, 1929–1958.
76. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
77. Fey, M.; Lenssen, J.E. Fast Graph Representation Learning with PyTorch Geometric. *arXiv preprint arXiv:1903.02428* **2019**.