Aspect-level sentiment analysis merged with knowledge graph and graph convolutional neural network

Zuhua Dai, Yuanyuan Liu*, Shilong Di and Qi Fan
College of Computer Science and Engineering, Northwest Normal University, Lanzhou, China

*Corresponding author: 2019211755@nwnu.edu.cn

Abstract. Aspect level sentiment analysis belongs to fine-grained sentiment analysis, which has caused extensive research in academic circles in recent years. For this task, the recurrent neural network (RNN) model is usually used for feature extraction, but the model cannot effectively obtain the structural information of the text. Recent studies have begun to use the graph convolutional network (GCN) to model the syntactic dependency tree of the text to solve this problem. For short text data, the text information is not enough to accurately determine the emotional polarity of the aspect words, and the knowledge graph is not effectively used as external knowledge that can enrich the semantic information. In order to solve the above problems, this paper proposes a graph convolutional neural network (GCN) model that can process syntactic information, knowledge graphs and text semantic information. The model works on the "syntax-knowledge" graph to extract syntactic information and common sense information at the same time. Compared with the latest model, the model in this paper can effectively improve the accuracy of aspect-level sentiment classification on two datasets.

1. Introduction
Text sentiment analysis is the basic field of natural language processing (NLP). It refers to the semantic analysis of text with subjective sentiment to identify the sentiment polarity expressed by the text. According to the different granularity of the processed text, sentiment analysis is divided into three levels: chapter level, sentence level and aspect level [1]. Different from text-level and sentence-level sentiment analysis tasks, Aspect Based Sentiment Analysis (ABSA) is to explore different aspects of an entity's sentiment tendencies. In the field of e-commerce, aspect-level sentiment analysis can obtain consumer groups' preference for various aspects of products based on the analysis of the review text of users' purchases, and then more accurately understand consumers' product preferences. Guide the positive emotional aspects, realize intelligent marketing, improve product quality for negative emotional aspects, reduce the risk of user complaints, and serve consumers more effectively.

For the ABSA task, the method based on deep learning [2-6] is to first represent the semantic feature vector of the text, and then build a neural network model to extract the features of the target and predict the emotional polarity. Tang et al. [2] used two LSTM structures to model the contextual semantics of the text before the target aspect word and the text content after the target aspect word. With the effective application of the attention mechanism-based deep learning network in machine translation tasks, it has also been applied to ABSA tasks. For example, Wang et al. [3] proposed the ATAE-LSTM model to
merge contextual semantic vectors extracted from LSTM. Representation is spliced with the embedding vector of text aspect words, and then combined with the attention mechanism to capture the semantic information of the context, and perform aspect-level emotional tendency discrimination; Ma et al. [4] proposed an interactive attention network model, IAN, which uses interactive attention mechanism to fuse the contextual semantics with the target semantics information. This type of model automatically extracts the feature representation of the text and the target through a multi-layer neural network, which can well extract the contextual semantic features of the text, but cannot extract the syntactic dependency between the context word and the target word. To solve this problem, recent research work [7] [8] [9] began to introduce graph neural networks into the ABSA task.

Zhang et al. [7] used graph convolutional networks (GCN) to model text syntactic dependency graphs, and constructed an aspect-level sentiment classification model SGCN by fusing syntactic information and contextual semantic relations. Zhao et al. [8] proposed an aspect-level sentiment classification framework SDGCN based on the attention map convolution, which can be effective To capture the emotional dependence relationship between multiple aspects in a sentence; Xiao et al. [9] proposes a graph convolutional network model AEGCN based on attention coding.

Based on the above analysis, this article proposes a model that combines syntactic information and emotional common sense information for the ABSA analysis task to solve the following problems: (1) This article proposes a joint modelling syntactic knowledge and common sense knowledge based on the construction of a "syntax-knowledge" graph The network model extracts contextual syntactic information and emotional common sense information to better target terms. (2) For phrase-style aspect words, this paper proposes an interactive attention mechanism. After extracting the contextual semantics between multiple words of the target aspect word, the interactive attention mechanism interacts with the syntactic information to better obtain the semantic information and grammatical information of the target word.

2. Related work

In recent years, graph neural networks have been widely used in sentiment analysis tasks due to their powerful heterogeneous data computing capabilities, which can integrate graph structure information in feature representation learning.

Using graph neural networks to model syntactic dependencies can enhance the representation of word vectors, improve the model’s ability to model complex emotional semantics, and solve the problem of inaccurate and imprecise expression of contextual emotional semantics of word vectors. In order to effectively use the syntactic dependency graph and extract the structural information of the text, most of the current research work uses graph convolutional neural networks to model the syntactic information of the text [7-11]. Zhang et al. [7] used graph convolutional networks to model syntactic dependency graphs to extract syntactic information, and proposed an aspect-level sentiment classification framework (ASGCN), which uses Bi-LSTM to extract contextual information about texts. In order to extract the features of aspect words, the output of Bi-LSTM is used as the input of the multi-layer graph convolution structure to obtain a feature table fused with syntactic information, and then the mask mechanism is used to filter out the words of non-face words and retain the features of the aspect words; Xiao et al.[9] pointed out that some existing models use average pooling methods to extract features for aspect words, which cannot fully capture the semantic information of the context, and present the problems that the representation easily leads to semantic errors and missing information. In response to these problems, they proposed a graph convolutional network model based on attention coding (AEGCN). Sun et al. [10], proposed using graph convolutional network model to model syntactic information, and then splicing it with the target aspect word representation after the average pooling operation.

3. Network model

The aspect-level sentiment analysis task for this article can be defined as: Given a sentence–aspect pair \((S,A)\), \(A = \{w_{start}, w_{start+1}, \ldots, w_{end-1}, w_{end}\}\), and the "start" and the "end" represent the start
index and end index of the target aspect word in the sentence. \textit{Aspect} is a subsequence of the sentence \( S = \{w_1, w_2, ..., w_{n-1}, w_n\} \). The purpose of aspect-level sentiment classification is to predict the sentiment polarity \( C \in \{0, 1, 2\} \) of the target aspect words in the sentence, where 0, 1, and 2 represent "negative", "neutral" and "positive" emotional polarity.

3.1. Graph construction
In this section, we introduce in detail how to construct a syntactic graph based on the syntactic dependency tree and construct a "syntax-knowledge" graph based on the emotional knowledge base.

3.1.1. Syntax diagram. A syntactic graph \( G_s = \{V_s, E_s, A_s\} \) is constructed based on the syntactic dependency relationship. The edge set \( E_s \) contains the adjacent node pairs that have a dependency relationship in the syntactic dependency tree of the text. The node set \( V_s \) is a set of words contained in the text, and each word represents a node. In this paper, the Standford syntax analyzer is used to obtain the syntactic dependency tree of the sentence. The weight between the nodes \( V_i \) and \( V_j \) is defined as Equation (1):

\[
A_{ij} = \begin{cases} 
1, & \text{if } i, j \text{ are words, } e_{ij} \in E_s \\
1, & \text{if } i = j \\
0, & \text{otherwise}
\end{cases}
\] (1)

For example, the syntactic dependency graph and adjacency relationship of the text "the mouse buttons are hard to push" are shown in Figure 1 and Figure 2:

![Figure 1. Syntax dependency graph.](image1)

![Figure 2. Adjacency relationship graph.](image2)
Syntax graph node feature representation: the words in the text are the nodes of the syntactic graph, and the text words are represented by the BERT word embedding vector as the input of the Bi-GRU network model, and the representation of the fusion text context semantic information $H^{\text{GRU}}$ is used as the feature representation of the syntactic graph node. Therefore, $H^{\text{GRU}}$ is the feature matrix of $G_{s}^i$, $H^{i} \in \mathbb{R}^{n \times d_s}$, where each row of $H^{i}$ is the feature representation of the word node $v_i$, $n$ is the number of words in the text, and $d_s$ is the output dimension of the Bi-GRU network.

3.1.2. "Syntax-Knowledge" graph. The knowledge base introduced in this section is SenticNet emotional knowledge base, which provides related semantic representation, emotion and polarity information for 100,000 natural language concepts. Semantics refers to the most semantically related concepts of the input concept, and emotion refers to the four emotional dimensions (Pleasantness, Attention, Sensitivity, and Aptitude), and the emotional value is the emotional polarity value between -1 and +1.

Use senticnet5 as external knowledge and combine the syntactic dependencies of the text to construct a "syntax-knowledge" graph $G_{zs} = \{V_{zs}, E_{zs}, A_{zs}\}$, containing syntactic information and external knowledge information. Its nodes $V_{zs}$ includes text word nodes and knowledge nodes extracted from text words, and the edge sets $E_{zs}$ including the syntactic dependence of text words, the emotional dependence of text word nodes and knowledge nodes, and the dependence between knowledge nodes. The weight between nodes $v_i$ and $v_j$ is defined as Equation (2):

$$A_{yz} = \begin{cases} 
1, & i \text{ and } j \text{ are words, } e_{yz} \in E_{zs} \\
1, & i \text{ is word and } j \text{ is knowledge node, } e_{yz} \in E_{zs} \\
1, & i \text{ and } j \text{ are knowledge node, } e_{yz} \in E_{zs} \\
1, & i = j \\
0, & \text{otherwise} 
\end{cases}$$

(2)

Node feature representation: The representation of a knowledge node needs to be able to contain the structural information of the node in the knowledge graph. This article uses the existing affective-space emotional space embedding representation affective-space. This vector representation maps the concept of SenticNet to continuous low-dimensional embedding matrix $E^{\text{eff}}$. Without losing the semantic and emotional relevance of the original space. The vector representation $H^{\text{eff}}$ of the knowledge node can be obtained by looking up the embedding matrix $E^{\text{eff}}$. The feature matrix $H^{zs}$ of the graph $G^{zs}$ is the concatenation of $H^{\text{GRU}}$ and $H^{\text{eff}}$, where each row of $H^{zs}$ is the feature vector of the text word or knowledge node $V_i$. The "Syntax-Knowledge" graph of the text "the mouse buttons are hard to push" is shown in Figure 3.
3.2. Introduction to the network model
This paper uses senticnet5 [11] as external knowledge to construct an emotional knowledge graph, and combines the syntactic dependency of the text to construct a "syntax-knowledge" graph, which is used as the input of the graph convolutional neural network model to extract the syntax related to the target aspect word information and external knowledge information. Then, the multi-head position attention mechanism is used to embed the position information for the feature representation of the target aspect words. For the target aspect composed of multiple words, the model uses the Bi-GRU network to extract the contextual semantics of the target aspect, and the extracted features interact with the syntactic information of the text through the interactive attention mechanism. The network model proposed in this paper is mainly composed of the following parts: (1) BERT: Input the preprocessed datasets into BERT to generate word embedding vectors. (2) Bi-GRU: Extract the contextual semantic information of the review text and the contextual semantic information of the aspect words. (3) Graph construction: extract the syntactic dependency graph of the text, and integrate senticnet5 as external knowledge to construct a "syntax-knowledge" graph on this basis. (4) GCN: Extract the feature representation of the word node in the target aspect based on the "syntax-knowledge" graph. (5) Multi-head position attention mechanism: integrate position information into the multi-head attention mechanism, and integrate position information for the semantic representation of the text. (6) Multi-head interactive attention mechanism: fusion of syntactic information for the contextual semantic representation of words in the target aspect. The model framework proposed in this paper is shown in Figure 4:

![Figure 3. Syntax-Knowledge graph.](image-url)
3.2.1. Pre-trained language model. This article uses the BERT [13] pre-training model. The BERT model is a language model based on two-way Transformer proposed by Google. For a text $S = \{w_1, w_2, ..., w_n\}$, express the input as a sequence $\text{seq} = [(CLS), w_1, w_2, ..., w_n, (SEP)]$. Text embedding result $X = \text{BERT}(\text{seq}) \in \mathbb{R}^{d_{\text{word}}}$ where $d_{\text{word}}$ is the embedding dimension. The vector representation $x_i \in \mathbb{R}^{d_{\text{word}}}$ of the word $w_i$.

3.2.2. Context semantic extraction module. In this paper, the Bi-GRU network is used to extract the contextual semantics of the text and the contextual semantics of the target terms. The Bi-GRU network is composed of a basic GRU structure. For each word vector of the input text, it is input to a forward GRU unit and a backward GRU unit, and then the outputs of the two are spliced to obtain a bidirectional GRU output.
3.2.3. Graph Convolutional Neural Network. After the construction of the "syntax-knowledge" graph is completed, the feature representation of its nodes and the adjacency matrix are used as the input of the GCN model, and GCN is used to model the "syntax-knowledge" graph, which combines the textual syntactic dependency information and external knowledge for the target aspect word Node information. The calculation formula is shown in (3):

\[ H_{zs}^{j+1} = \sigma(\tilde{A}_{zs}H_{zs}^{j}W_{zs}^{j}) \] (3)

\[ \tilde{A}_{zs}^{j+1} = D_{zs}^{-\frac{1}{2}}A_{zs}D_{zs}^{-\frac{1}{2}} \] is the normalized symmetric adjacency matrix, \( W_{zs}^{j} \) is \( j \)-th layer Weight matrix of GCN network. \( D_{zs} \) is the degree matrix of \( A_{zs} \), \( D_{zs} = \sum_{j} A_{zs}^{(x)w}_{ji} \)

3.2.4. Multi-head attention mechanism. For the common multi-head attention mechanism (Multi-Head Attention, MHA), define a key sequence \( k = \{k_1, k_2, k_3, ..., k_n\} \) and a query sequence \( q = \{q_1, q_2, q_3, ..., q_n\} \), calculate the attention weight distribution through the key sequence and the query sequence, and weight the weight value to the value sequence, which is usually set key = value in the field of natural language processing. The calculation formula of the attention mechanism is shown in (4), (5):

\[ Attention(k, q) = soft\ max(f_m(k, q))k \] (4)

\[ f_m(k, q) = \text{tanh}([k_i; q_j]W_v) \] (5)

\( f_m \) is the function used to calculate the semantic similarity between \( k_i \) and \( q_j \), and \( W_v \) is the learnable weight matrix. The multi-head attention mechanism can learn different weight matrices, with different parameters. The output of \( n_{head} \) is spliced and projected to a specific hidden dimension \( d_{hid} \) by the Equation (6):

\[ MHA(k, q) = [o_1 \oplus o_2 \oplus ... \oplus o_{n_{head}}] \cdot W_m \] (6)

The multi-head interactive attention mechanism (MHIA) refers to the multi-head attention mechanism of the general form of input \( q \neq k \). A multi-head interactive attention mechanism is carried out for the text representation extracted by the model that combines the syntactic dependency relationship and the context semantic coding \( h^0 \) of the target aspect word, and the syntactic information is merged for the context semantic information of the target aspect word. The calculation formula is as shown in (7):

\[ h^{io} = MHA(h^0, h^0) \] (7)

For aspect-level sentiment analysis tasks, the importance of text words at different positions from the target aspect words is different, the model sets up a multi-head position attention mechanism to achieve this. The coding method is defined as Equation (8):

\[ pos_i = \begin{cases} |i - \text{start}|, & 1 \leq i < \text{start} \\ 0, & \text{start} \leq i \leq \text{end} \\ |i - \text{end}|, & \text{end} < i \leq n \end{cases} \] (8)
The multi-head position attention mechanism (MHPA) formula expressed as Equation (9):

$$h^z = MHA(h^z, x^{pos})$$

(9)

4. Experiments

4.1. Dataset

This paper uses the SemEval2014 [13] data set for comparative experiments. The emotional polarity of the data samples is divided into positive, negative and neutral. Among them, the SemEval2014 data set is a data set for semantic evaluation competition tasks, including user reviews in two areas: laptop and restaurant. Through comparative experiments, it is verified that the model proposed in this article has achieved good sentiment classification performance on different domain datasets. Table 1 shows the statistics of the experimental data used in this article:

| datasets       | positive | negative | neutral | Number of aspect words>=2 |
|----------------|----------|----------|---------|---------------------------|
| Laptop-train   | 994      | 870      | 464     | 38.4%                     |
| Laptop-test    | 341      | 128      | 168     |                           |
| Restaurant-train| 2164     | 807      | 637     | 25.53%                    |
| Restaurant-test| 728      | 196      | 196     |                           |

4.2. Comparative Experiment

Compare the model in this article with the following six network models:

(1) TD-LSTM: Tang et al. [6] used two LSTM structures to model the contextual semantics of the text before the target aspect word and the text content after the target aspect word, and spliced the results together as the final result for prediction.

(2) ATAE-LSTM: The model of Wang et al. [3] combines the vector representation of the fusion contextual semantics extracted from the LSTM with the embedding vector of the text aspect words to perform the aspect level judgment of emotional orientation.

(3) IAN: Ma et al. [8] proposed an interactive attention network model, IAN, which uses LSTM to extract the contextual semantics of text and target words, and then uses an interactive attention mechanism to combine the contextual semantics with target semantics. Perform fusion, and combine the results to determine emotional orientation.

(4) MEMNET: Tang et al. [15] used a deep memory network fused with attention mechanism, and added location feature information to the attention mechanism to realize the classification of aspect-level emotional polarity.

(5) ASGCN: Zhang et al. [7] proposed a model based on graph convolution and syntactic dependency tree to model the contextual semantic and syntactic dependency information of the text, and use the attention mechanism to interact between the two information.

(6) Xiao et al. [9] proposed a graph convolutional network model based on attention coding. The model is mainly composed of a multi-head attention mechanism and a graph convolutional network that models a syntactic dependency graph.

4.3. Analysis of experimental results

The performance results of different models are shown in Table 2. The models are evaluated based on accuracy and F1 index. The final result is the average of three random initialization experiments.
The above table shows the performance of the proposed model and other typical aspect-level sentiment classification models. It can be seen from the above table that compared to the TD-LSTM and ATAE-LSTM models that only use the recurrent neural network and the ASGCN model that only uses the graph convolutional neural network to model the syntactic dependency tree, the fusion of knowledge graph and syntax information proposed in this paper the model achieved better results.

5. Concluding remarks
This paper proposes a model that uses graph convolutional neural networks to jointly model syntactic dependency graphs and common sense knowledge graphs to extract text context semantic information, grammatical information, semantic information of target terms, and common sense information for target terms to enrich the semantics of terms. Experiments have proved that the model that incorporates common sense knowledge graph information has a higher accuracy rate. In the following work, we can further improve the representation of part-of-speech embedding and perform weighting operations for graph convolutional neural networks to improve the extraction of text syntax dependency information.

Acknowledgments
This paper was completed under the careful guidance of teacher Dai Zuhua, and I would like to extend my sincere thanks to the teacher. In addition, the paper was published during the project period of Lanzhou Science and Technology Bureau. Thank you for project funded by Lanzhou Science and Technology Bureau (Project number: 2020-ZD-136).

References
[1] Liu B 2012 Sentiment analysis and opinion mining Synthesis lectures on human language technologies 5(1) 1-167.
[2] Tang D Y, Qin B, Feng X C and Liu T 2016 Effective LSTMs for target-dependent sentiment classification In Proceedings of COLING 3298-3307.
[3] Wang Y Q, Huang M L, Zhao L and Zhu X Y 2016 Attention-based LSTM for aspect-level sentiment classification In Proceedings of EMNLP 606-615.
[4] Ma D H, Li S J, Zhang X D and Wang H F 2017 Interactive attention networks for aspect-level sentiment classification In Proceedings of IJCAI 4068-4074.
[5] Chen P, Sun Z Q, Bing L D and Yang W 2017 Recurrent attention network on memory for aspect sentiment analysis In Proceedings of EMNLP 452-461.
[6] Yin W P, Schütze H and Xiang B 2016 Abcnn: Attention-based convolutional neural network for modeling sentence pairs Transactions of the Association for Computational Linguistics, 4(1) 259-272.
[7] Zhang C, Li Q and Song D 2019 Aspect-based sentiment classification with aspect-specific graph convolutional networks In Proceedings of EMNLP 4560-4570.
[8] Zhao P L, Hou L L and Wu O 2020 Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification Knowledge-Based Systems 193 105443.
[9] Xiao L W, Hu X H and Chen Y N 2020 Targeted sentiment classification based on attentional encoding and graph convolutional networks Applied Sciences 10(3) 957.

[10] Sun K, Zhang R C, Mensah S and Mao Y Y 2019 Aspect-level sentiment analysis via convolution over dependency tree In Proceedings of IJCNLP. 5683-5692.

[11] Kipf T N, Welling M 2017 Semi-supervised classification with graph convolutional networks In Proceedings of ICLR.

[12] Cambria E, Poria S and Hazarika D 2018 SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings In Proceedings of AAAI.

[13] Devlin J, Chang M W, Lee K and Toutanova K 2013 Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding Encyclopedia of Systems Biology.

[14] Pontiki M, Galanis D and Papageorgiou H 2014 SemEval-2014 Task 4: Aspect Based Sentiment Analysis in Proceedings of SemEval 27-35.

[15] Tang D Y, Qin B and Liu T 2016 Aspect level sentiment classification with deep memory network In Proceedings of EMNLP 214-224.