A Fine-Grained Sentiment Analysis Model Based on Multi-Task Learning

Xin Fan and Zhonglin Zhang

April 19, 2024

A fine-grained sentiment analysis model based on multi-task learning

Xin Fan and Zhonglin Zhang
(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou,730070)

Abstract

Fine-grained sentiment analysis based on textual data is currently a prominent research topic in the realm of natural language processing. The objective of this analysis is to forecast the various dimensions of sentiment within a sentence. Nevertheless, the majority of existing sentiment analysis models predominantly concentrate solely on aspect extraction or sentiment tendency analysis, catering to single-task patterns. In light of the aforementioned issues, this paper introduces a fine-grained sentiment analysis model called BLAB (BERT Local Context Focus AD-BiReGU), which is founded on multi-task learning principles. In this model, the incorporation of the AD-BiReGU module into the BERT-LCF framework enables it to perform aspect word extraction and fine-grained sentiment analysis concurrently. Initially, the pre-trained BERT model is employed to capture the initial features of both local and global contexts. Within the feature extraction layer, local context features are extracted through the incorporation of the local context focus mechanism with the multi-head attention mechanism, facilitating dynamic context feature masking. Simultaneously, the two-layer BiReGU model, grounded in the attention mechanism, is employed to incorporate context information into the neural network model, thereby capturing long-term dependencies between labels and textual features to extract global features. Subsequently, the fusion of local text information and global information occurs, followed by their input into the nonlinear layer to derive the ultimate sentiment polarity results. Comparative experiments indicate that the incorporation of the AD-BiReGU module yields a discernible enhancement in performance for the aspect word extraction task within multitasking scenarios.

Keywords: Feature extraction; Aspect polarity classification; Aspect level sentiment analysis; Multi task learning; BERT model

1 Introduction

In recent years, the internet has experienced rapid development, resulting in a proliferation of texts reflecting users' emotional inclinations. Accurately and timely analyzing the sentiments expressed in these texts[1] can significantly contribute to comprehending the emotional attitudes of the public towards commercial products, policies, and regulations, thereby offering a dependable foundation for adjusting product marketing strategies or policy regulations. Traditional coarse-grained text sentiment analysis methods are inadequate to meet contemporary requirements; thus, achieving fine-grained text sentiment analysis has become a prominent concern in the field of natural language processing.

Sentiment analysis is categorized into text-level, sentence-level, and aspect-oriented sentiment analysis to cater to diverse research audiences[2]. Presently, the majority of research on textual sentiment analysis focuses on the chapter and sentence levels, with limited attention given to aspect-oriented sentiment analysis. Chapter-level and sentence-level text sentiment analysis constitute coarse-grained sentiment analysis, wherein sentiment polarity is assessed based on the entirety of the text or individual sentences, respectively. Nonetheless, human language is inherently rich in content, with paragraphs or even sentences often addressing multiple aspects. Limiting sentiment analysis solely to the chapter and sentence levels results in significant information loss, yielding conclusions that diverge considerably from reality. Hence, finer-grained aspect-level sentiment analysis[3] is necessary to effectively process such textual data. For instance, consider the statement, "The price of the computer was too high, but the performance was amazing. "In this review, the aspects highlighted are "price" and "performance, " with the associated opinion words being "high" and "amazing" respectively. "high" represents a negative emotion and "amazing" represents a positive emotion. In this case, aspectual sentiment analysis[4] helps to capture the different aspects of the text more comprehensively.

2 Related work

In the early days, aspectual sentiment analysis relied heavily on constructing sentiment lexicons[5] to extract the sentiment values of sentiment words in text. Despite the simplicity of this approach, it suffers from the disadvantage of being highly dependent, causing problems due to the lack of sentiment words. In addition, sentiment words often need to be constructed manually, thus creating a heavy workload.

With the continuous development of related research, Researchers in the field of deep learning are very concerned with aspect-level sentiment analysis. For example, Man R[6] used BERT combined with convolutional neural networks to effectively extract local features of text, and the experiments on public datasets performed well. Although these methods effectively enhance the context-dynamic representation of words, the shallow features obtained by relying only on simple segmentation lack semantic information and ignore the complex semantic relationships between words. Scholars like Han et al.[7] employed a multi-channel modeling strategy, integrating CNN and Bi-LSTM, to significantly enhance the efficacy of microblog sentiment analysis; while researchers such as Li Hui et al. [8] utilized a technique involving the fusion of dual-channel features with GRU and CNN to conduct comparative analyses across three datasets, revealing that their proposed dual-channel model outperformed single-channel approaches. Despite utilizing multiple channels for in-depth analysis of textual expressions, the method inadequately addresses the issues of information loss and insensitivity to location information stemming from the CNN pooling process during local feature extraction. Nonetheless, BERT-PT (BERT post-training), BERT-SPC, and BERT-LCF (BERT local context focus), all developed based on BERT, enhance the BERT model’s focus on the correlation of attribute word sequences by appending these words at the end of input sequences, thereby mitigating the risk of information loss.

The multi-head attention mechanism, proposed by the Google team[9], performs attention operation computations within the input text sequence to discover connections among sequence elements and automatically learns the weight distribution. Additionally, the ReGU (Residual Gated Unit) proposed by Luo et al[10] is an extra RNN unit with a structure comprising two gates that regulate the flow of input and hidden state information. Like skip connections, ReGU can utilize one gate to transmit input to the output, aiding in deeper training and the extraction of valuable text features. Similar to skip connections, ReGU can use one gate to pass the input to the output, which helps to train more deeply and obtain useful text features.

In summary, this paper introduces fine-grained sentiment analysis techniques and proposes a fine-grained sentiment analysis model BLAB (BERT local context focus AD-BiReGU) for comment text. The model utilises a multi-task learning framework and incorporates the AD-BiReGU module on the basis of BERT-LCF, which uses the dual embedding mechanism and ReGU as auxiliary information of the model on the basis of traditional BiLSTM. The use of the attention mechanism and the BiLSTM model takes full account of the importance of different words in the text sequence and text features to better encode the output sequence and capture the long-term dependencies between labels. The simultaneous extraction of emotionally aspectual words and corresponding emotional tendencies is achieved.

3 Model methods

In this section, we outline the proposed BLAB method, which consists of four components: a BERT pre-training layer, a text feature extraction layer, a feature fusion layer, and a multi-task output layer. The overall architecture of BLAB is shown in Fig. 1.

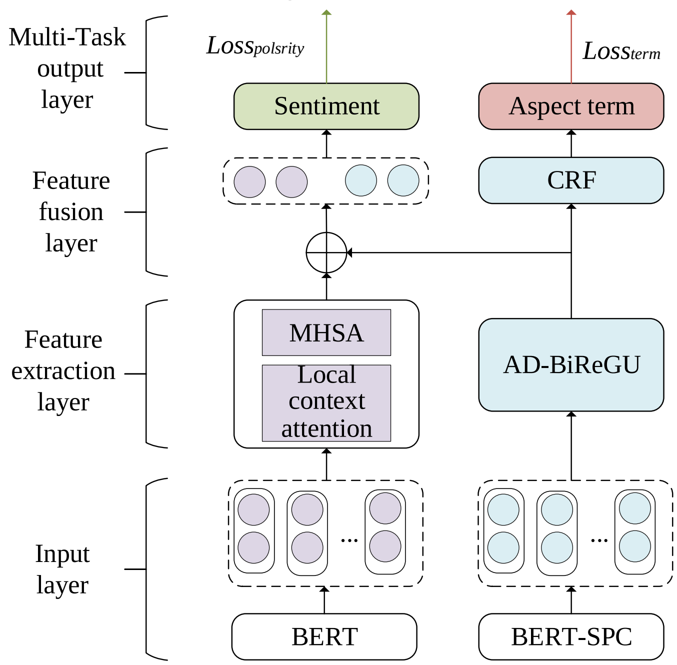

Figure 1: BLAB Multi Task Model Framework

3.1 BERT pre-trained model layer

BERT (Bidirectional encoder representation from transformers) is a language representation model built on the basis of large sample pre-training, which can effectively express the rich semantic information contained in the text. Two separate BERT pre-trained models were used in the article to improve the model's ability to extract features, BERT-BASE and BERT-SPC, are introduced into the text task to deeply obtain information about the characteristics of the text content. They are integrated into the text task to comprehensively capture information regarding the characteristics of the textual content. In this paper, two distinct BERT pre-trained models are utilized. Words within sentences are represented as X^l and X^g to capture global and local contextual information, respectively.

$$O_{\text{BERT}}^{l} = \text{BERT}^{l}(X^{l}) \tag{1}$$

$$O_{\text{BERT}}^{\text{g}} = \text{BERT}^{\text{g}}(X^{\text{g}}) \tag{2}$$

where O_{BERT}^{l} and $O_{\text{BERT}}^{\text{g}}$ denote the output representations of the local context and global context processors, while BERT^{l} and BERT^{g} represent the BERT pre-training model counterparts modeling global semantic information and local semantic information, respectively.

3.2 Feature extraction layer

At the feature extraction level, this study primarily utilizes the local context focusing mechanism and integrates the multi-head attention mechanism to capture local contextual information. For global contextual features, this study employs a two-layer BiReGU network model built upon the attention mechanism. Building upon the BiLSTM framework, the ReGU structure is introduced to facilitate deeper training and the extraction of pertinent textual features. The features extracted by the initial BiReGU layer undergo processing through the attention mechanism, aimed at discerning significant information within the text and autonomously learning the relatively salient words in the input text sequence. Utilizing the output from the word attention computation layer as the input for the subsequent BiReGU layer enables the method to efficiently capture global features of textual vectors.

The local context focusing mechanism employs Semantic Relative Distance (SRD) to discern the proximity of the context to the target aspect. Through employing this mechanism, models can capture local semantic features. To enhance the semantic characterization of local contexts, a multi-head attention mechanism is employed subsequent to the dynamic masking layer for context features.

Multi-Head Self-Attention (MHSA) mechanism uses multiple attention heads, each of which is an independent self-attention mechanism from which a different set of weights and representations are learnt. In Multi-Head Attention mechanism, the input sequence is first mapped to several different query, key and value spaces by linear transformations. Each attention head independently computes attention on the mapped queries, keys, and values to derive a representation for every position. Subsequently, the representations from each attention head in the model are combined through linear transformations and splicing operations to yield the final output representation.

For global contextual features, each layer of the two-layer BiReGU network model utilises a bi-directional ReGU structure to process both the forward and backward contextual information, so as to fully mine the contextual information of the input sequence. Initially, the word vectors are fed into the first BiReGU layer to acquire the hidden state output of both the forward and backward ReGU. Subsequently, the hidden states from both directions are concatenated to produce the final output. Then, the output of the first BiReGU layer undergoes attention computation to assess the significance of individual words. Recognizing the challenge of acquiring comprehensive global feature information with a single BiReGU layer, a two-layer BiReGU structure is employed. Here, the output of the word attention computation layer serves as input to the second BiReGU layer to capture richer global feature information.

3.3 Feature fusion layer

The feature fusion layer fuses local context information and global context information to enhance the textual representation of specific attribute words and outputs the feature X_{polarity}, calculated as

$$X_{\text{polarity}} = W_f \bullet \left(\text{concat}\left[O_{\text{l}}^{\text{CDW}};\quad O_{\text{g}}^{\text{MHSA}}\right]\right) + b_f \tag{3}$$

Among them, W_f and b_f are learnable weight matrices and bias terms.

3.4 Multi task output layer

The multi-task fine-grained sentiment analysis model presented in this paper includes two output layers that utilize the Softmax activation function for exponential normalization to predict the probability distribution of attribute word sentiment polarity. The sentiment polarity prediction task employs the cross-entropy loss function for calculating its loss. To generate the final aspect term labels, Conditional Random Field (CRF) is employed as the final layer instead of a softmax classifier. This substitution is intended to enhance performance by mitigating the high dependency between annotations.

4 Experiment

4.1 Dataset

The two datasets Restaurant14 and Laptop14 from SemEval-2014Task, as well as the Twitter dataset, were used in the experimental aspect to validate the model. In the comment messages in these three public datasets, the three different emotional polarities correspond one to one with the aspect words in each sentence, and a description of the dataset is given in Table 1.

Dataset	positive		negative		neutral	
	Training set	Test set	Training set	Test set	Training set	Test set
Restaurant	2164	728	807	196	637	196
Laptop	994	341	870	128	464	169
Twitter	1561	173	1560	173	3127	346

Table 1 Dataset Distribution

4.2 Analysis of experimental results

To assess the effectiveness of the proposed multi-task model for attribute extraction and fine-grained sentiment prediction on the e-commerce review text dataset, it was chosen as the baseline to visualize the model's impact more clearly. The results are presented in Table 2.

（unit：%）

model	Restaurant	Laptop	Twitter
	$F1_{ate}$/(Acc_{apc}/ $F1_{apc}$)	$F1_{ate}$/(Acc_{apc}/ $F1_{apc}$)	$F1_{ate}$/(Acc_{apc}/ $F1_{apc}$)
BERT-BASE	89.01/(83.28/73.89)	82.59/(80.22/75.38)	95.27/(80.22/75.38)
LCF-ATEPC	88.51/(86.38/80.72)	82.54/(80.37/76.95)	96.49/(78.28/76.30)
BLAB	89.75/(88.54/82.89)	84.57/(81.45/78.02)	97.85/(80.69/78.82)

Table 2 Model Accuracy and F1. $F1_{ate}$, Acc_{apc} , and $F1_{apc}$ are the macro-F1 score of ATE subtask, accuracy and macro-F1 score of APC subtask.

The comparison in Table 2 highlights the performance of BLAB, the joint multi-task attribute extraction and sentiment analysis model proposed in this paper, against other multi-task models. All the multi-task comparison models in this paper are based on BERT. BERT-BASE exhibits poor performance in fine-grained sentiment analysis due to the absence of attribute word information, resulting in an $F1_{apc}$ of only 73.89%. LCF-ATEPC enhances feature extraction for fine-grained sentiment analysis compared to BERT-BASE, but performs similarly to other BERT-based models in attribute extraction, achieving an $F1_{ate}$ value of 88.51%. Upon comprehensive analysis of the experimental results, the model proposed in this paper demonstrates optimal performance in both attribute extraction and fine-grained sentiment analysis tasks, affirming its effectiveness.

5 Conclusion

For fine-grained sentiment analysis, this study introduces the BLAB model, a fine-grained sentiment analysis framework developed through multi-task learning, originating from aspect-oriented sentiment analysis. Initially, two BERT pre-training models are employed to extract both local and global contextual features. Local features are captured utilizing the local focus mechanism and the multi-attention mechanism, while global features are acquired through the application of the AD-BiReGU model. Ultimately, the local and global information is integrated and inputted into the nonlinear layer to derive sentiment analysis outcomes. To enhance performance and address the challenge posed by high dependency between annotations, the paper employs CRF instead of a softmax classifier as the final layer to generate aspect term labels. Experimental evaluations conducted across three publicly available datasets robustly indicate the superior performance of the proposed BLAB model compared to the baseline. Suggestions for further enhancement will be explored subsequently.

* Project supported by the National Natural Science Foundation of China (Grant Nos.61662043) and Phased Research Results of Gansu Philosophy and Social Sciences Planning Project (20YB056)

References

[1] Mayukh S, Ilanthenral K,W.B. V .Emotion quantification and classification using the neutrosophic approach to deep learning[J].Applied Soft Computing,2023,148

[2] Yuehua Z ,Linyi Z ,Chenxi Z , et al. Construction of an aspect-level sentiment analysis model for online medical reviews[J].Information Processing and Management,2023,60(6):

[3] Zohair A ,Jianxin W .A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments[J].Alexandria Engineering Journal,2023,65731-747.

[4] Martin S, Suzanne E ,R. I H , et al. The power of emotions: Leveraging user generated content for customer experience management[J].Journal of Business Research,2022,144997-1006.

[5] Yabing W, Guimin H, Maolin L , et al. Automatically Constructing a Fine-Grained Sentiment Lexicon for Sentiment Analysis[J].Cognitive Computation,2022,15(1):254-271.

[6] Man R, Lin K. Sentiment Analysis Algorithm Based on BERT and Convolutional Neural Network[C]//Proceedings of the 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers. IEEE, 2021: 769-772.

[7] Han Pu, Zhang Wei, Zhang Zhanpeng, et al. Sentiment Analysis of Weibo Posts on Public Health Emergency with Feature Fusion and Multi-Channel[J]. Data Analysis and Knowledge Discovery, 2021,5(11):68-79.

[8] Li Hui, Huang Yujie, Li Jinqiu. Text Sentiment Classification Based on HAN and Two-channel Composite Model[J]. Transducer and Microsystem Technologies,2021,40(08):121-125.DOI:10.13873/J.1000-9787(2021)08-0121-05.)

[9] VASWANI A，SHAZEER N，PARMAR N，et al. Attention is all you need[C]//Advances in Neural Information Processing Systems，2017：5998-6008

[10] LUO H，LI T，LIU Bet al. DOER: dual cross-shared RNN for aspect term -polarity co-extraction[C]//57th Conference of the Association for Computational Linguistics，2019：591-601.