Position-aware Hybrid Attention Network for Aspect-Level Sentiment Analysis

Yongqiang Zheng, Xia Li*, Guixin Su, Junteng Ma, Chaolin Ning

School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou, China

1. Motivation

Aspect-level sentiment analysis is a fine-grained task in sentiment analysis, which aims to predict sentiment polarity (i.e., positive, neutral, or negative) of a specific target of a given sentence. Most of the previous methods focus on capturing the context information of words across the sentence related to the target, ignoring the importance of the independent relationship between the opinion words and the target.

We argue that the opinion words are more important in supervising the polarity of the sentence for the given target, that is to say, we can independently consider the importance of the relationship between the target and opinion words.

To address this limitation, we proposed a position-aware hybrid attention network based model which consists of two components, namely opinion attention network and context attention network. The context attention network is used to capture context information between words across sentence with the target, and the opinion attention network is used to incorporate independent relationship between opinion words and the target.

2. Contribution

The main contributions are as follows:

(1) We propose a hybrid attention network to capture the context information between the words across sentence with the target, as well as the independent relationship between the opinion words and the target to obtain more precisely sentiment information of the given target in the sentence.

(2) We conduct several experiments and ablation tests on public laptop and restaurant datasets to validate our model. We will show that our model achieves a stable and effective performance compared with the baseline models.

3. Model

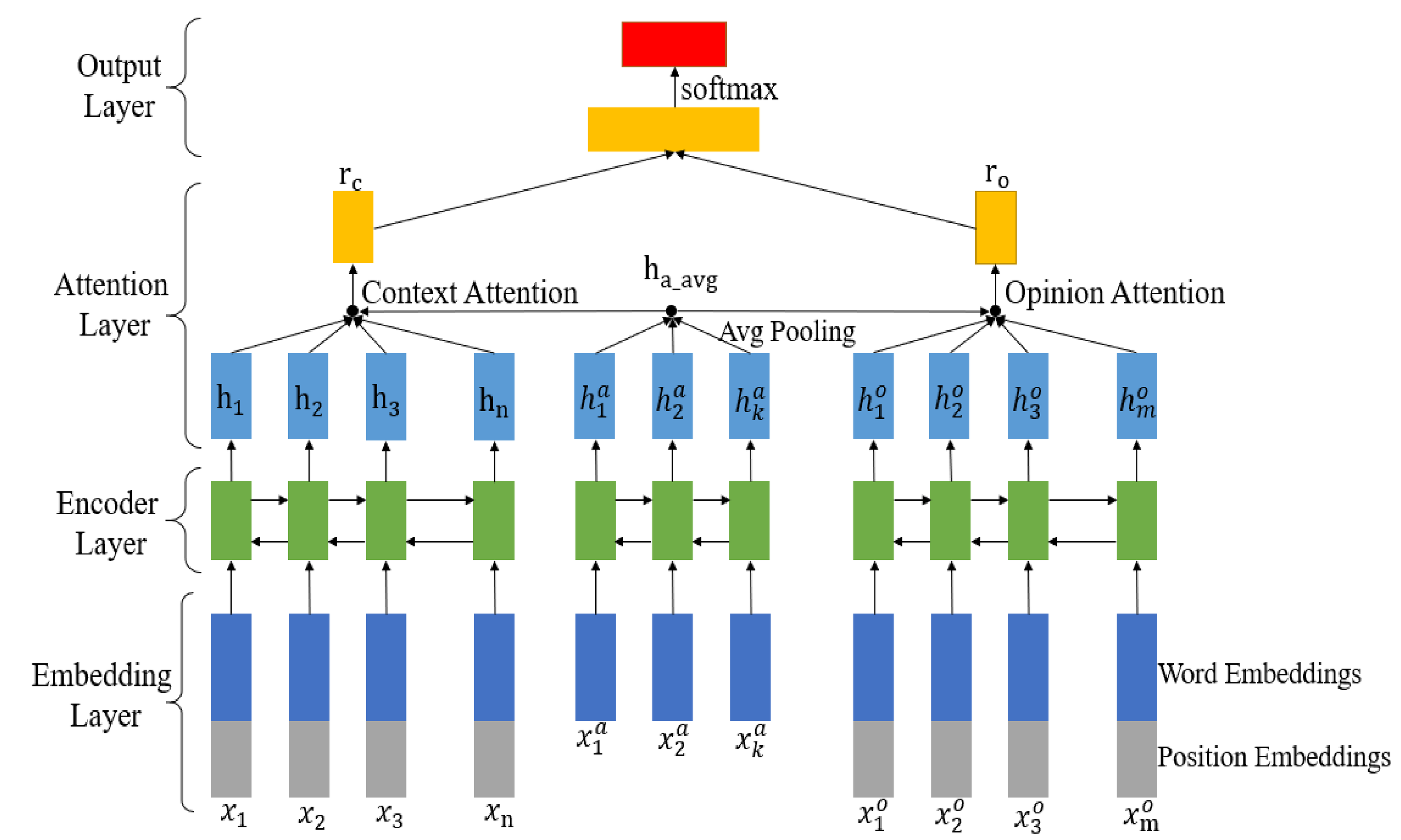

Fig.1.The architecture of our model.

Opinion Attention. Given the target $H_a = \{h_1^a, h_2^a, \dots, h_k^a\}$, and the hidden state of opinion words $H_o = \{h_1^o, h_2^o, \dots, h_m^o\}$, the Opinion Attention score α can be calculated by the following formulas.

$$h_{a\_avg} = \frac{1}{k}\sum_{i=1}^{k} h_i^a \qquad f_o(h_i^o, h_{a\_avg}) = h_i^o W_{att1} h_{a\_avg}{}^{\mathrm{T}}$$

$$\alpha_i = \frac{\exp\left(f_o(h_i^o, h_{a\_avg})\right)}{\sum_{j=1}^{n}\exp\left(f_o(h_j^o, h_{a\_avg})\right)} \qquad r_o = \sum_{i=1}^{n} h_i^o \alpha_i$$

Context Attention. Given the target representation and the hidden state of each words across sentence $H = \{h_1, h_2, \dots, h_n\}$, the contextual attention score β can be calculated by the following formula.

$$f_c(h_i, h_{a\_avg}) = h_i W_{att2} h_{a\_avg}{}^{\mathrm{T}} \qquad \beta_i = \frac{\exp\left(f_c(h_i, h_{a\_avg})\right)}{\sum_{j=1}^{n}\exp\left(f_c(h_j, h_{a\_avg})\right)}$$

$$r_c = \sum_{i=1}^{n} h_i \beta_i$$

4. Experiments

Table 1. Experimental results of different models on the laptop and restaurant datasets.

Model	Laptop	Restaurant
Majority	53.45	65.00
TD-LSTM	68.13	75.63
AE-LSTM	68.90	76.20
ATAE-LSTM	68.70	77.20
MemNet	70.33	79.98
IAN	72.10	78.60
PosATT-LSTM	72.80	79.40
RAM	74.49	80.23
SHAN	74.64	81.02
Ours	**75.71**	**81.43**

Table 2. Experimental results of our model in ablation analysis.

Model	Laptop	Restaurant
Pos-LSTM	72.10	78.04
Pos-Context-ATT	74.45	79.82
Ours	75.71	81.43

5. Conclusions

Our model not only captures the context information of the words related to the target across the sentence, but also obtain the relationship between opinion words and the target. The experimental results carried on the public dataset show that our model is more effective than the compared baseline models. Although hybrid attention proposed in our model achieve good performance, we find that the information of opinion attention is not well used in context attention. In the following research, we will focus on the interaction between the opinion words and the context of the content. We hope that opinion words are helpful to supervise the generation of attention scores in the context, which can make the model focus on context words related to opinion words.

西安电子科技大学
XIDIAN UNIVERSITY