**Aspect-based sentiment analysis with alternating coattention networks**

1. **Introduction**

ABSA aims at inferring the sentiment polarity (e.g. positive, negative, neutral) of a sentence expressed toward a target which is the aspect of one specific entity.

The ABSA task involves two subtasks: aspect detection and sentiment classification. In ABSA task, multiple targets could appear in one sentence, and each target has related words to modify it. When judging the sentiment of current target, other targets and related words would become noises. Therefore, we need to fully consider the relationship between target and context words when we design neural networks. Meanwhile, using attention mechanism to learn context feature associated with the target for sentiment analysis has also shown its effectiveness in ABSA task.

It learns the attentions in target and context sequentially that it first generates the attention representation for target based on the context and then generates the attention representation for context based on the learned target representation.

This article has the following research contributions:

To reduce the effect of noise words of targets such as preposition and fully utilize the key words of targets to learn context representation, we replace the conventional attention with a new alternating coattention mechanism which models both the target-level and context-level attention in a more intuitive and effective interaction way for ABSA tasks.

To verify the effectiveness of coattention mechanism, we propose the Coattention-LSTM. It learns the nonlinear representations of context and target simultaneously and can extract more effective sentiment features from coattention mechanism.

In order to overcome the drawback of deep memory network applied in ABSA task which explores key context clues with an average target vector, we propose an interaction memory network based on the coattention mechanism named Coattention-MemNet to learn the key features from the target and context alternately with an iteration mechanism.

In order to take the location information into consideration, we propose a location weighted function for coattention mechanism and apply the function to both Coattention-LSTM and Coattention-MemNet.

1. **Literature review**
   1. *Aspect-based sentiment analysis based on LSTM*

ABSA aims at inferring the sentiment polarity of a sentence expressed toward a target which is one aspect of a specific entity. Therefore, the main challenge of ABSA task is how to effectively model the relationship between

target and context.

Neural networks, especially the LSTM network, can encode sentences without feature engineering and have been applied in many natural language processing tasks.

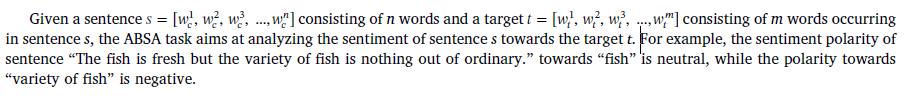
*2.2. Aspect-based sentiment analysis based on memory network*

Memory network encodes the inputs into vectors and stores them in the memory. The computation process of each layer can be divided into two steps. 1) Update the query representation. The query representation of each hop is different. The current layer of query representation is updated with the sum of initial query representation and relevant evidences from the previous layers. 2) Update relevant evidences based on input memory. The attention mechanism is used as computation unit to update the evidences related to the current layer’s query representation. That is, we update the memory vector of context according to the aspect memory vector and historical context memory vector.

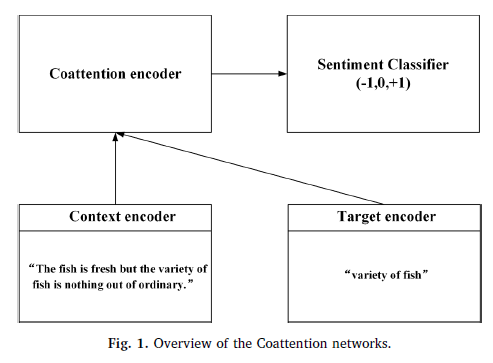
* 1. *Coattention networks*

1. **Methodology**

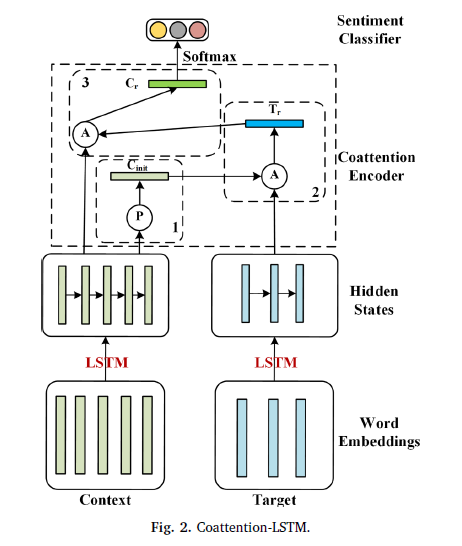
*3.1. Task definition and notation*



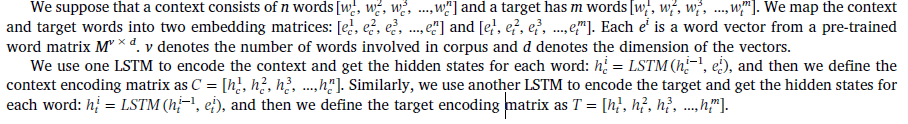
*3.2. An overview of coattention networks*



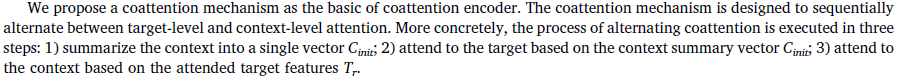
* 1. *Coattention-LSTM.*



*3.3.1. Context and target encoder*

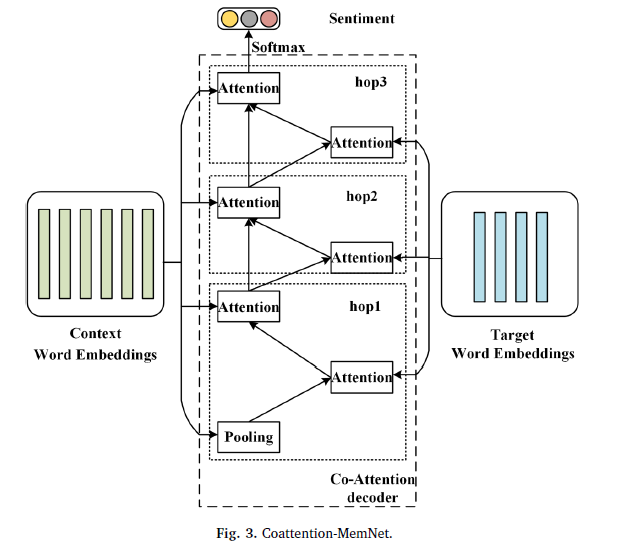


*3.3.2. Coattention encoder*



*3.3.3. Sentiment classifier*

*3.4. Coattention-MemNet*



*3.4.1. Context and target encoder*



*3.4.2. Coattention encoder*

*3.4.3. Sentiment classifier*

*3.5. Location-enhanced coattention*

In the Coattention-LSTM, the location weights can be directly added into the target-level attention and context-level attention. In Coattention-MemNet, we just add the location weights into the target-level attention of 1-hop layer and the context-level attention of last-hop layer, which can not only add the location prior information into the network but also avoid the location weights to limit the iterative learning process of attention weights.

1. **Experiments**

*4.1. Experiment preparation*

*4.1.1. Dataset*

We conduct experiments on SemEval 2014 Task 41 and Twitter (Dong et al., 2014) to validate the effectiveness of our model.

*4.1.2. Evaluation metric and parameters*

*4.2. Compared methods*

*4.3. Result comparisons*

*4.4. Analysis of networks*

*4.5. Analysis of location-enhanced coattention*

**5. Conclusion and future work**

In this paper, we first propose an alternating coattention mechanism which designs an alternating learning process for both targetlevel and context-level feature extraction in a more intuitive and effective way for aspect-based sentiment analysis task. Then we propose Coattention-LSTM network based on coattention mechanism which could learn both attention representations for target and context alternately, and we show that coattention mechanism in Coattention-LSTM could reduce the effect of noise words of targets and fully utilize the key words of targets to learn more effective context representation for sentiment analysis. The results on SemEval 2014 Datasets and Twitter Dataset indicate that Coattention-LSTM could outperform most of existing LSTM-attention methods. Further, we propose Coattention-MemNet which can learn the key features from the target and context alternately and adjust the learned features to make them more effective based on the multiple hops structure. The performance on SemEval 2014 Datasets and Twitter Dataset shows that the Coattention-MemNet achieves better performance than the conventional memory network. Finally, we propose a location weighted function for coattention mechanism based on the relative location information, which can generate different context features for multiple targets in one comment and limit the network to focus on the words around the target. We apply the location weighted function in both Coattention-LSTM and Coattention-MemNet and the results on SemEval 2014 Datasets and Twitter Dataset indicate the effectiveness of this function.

Since our methods use LSTM or word embedding layer to represent the words and attention is only a weighted sum function, it is difficult for them to learn the complex relationship between words such as the negations modifiers and the implicit sentiment phrases. Therefore, we need to improve the LSTM or attention, in order to make the network learn complex relationship between context words or map the results of attention layer to the classification space in a non-linear manner. In future work, we plan to design a context learning function capture the relationship between local context words with a complexity between word embedding and LSTM. We will also consider adding external knowledge into the network to solve the problem of negations modifier modeling and identify unknown sentiment words and phrases.