Encoder-Decoder Network with Cross-Match Mechanism for Answer Selection

Zhengwen Xie, Xiao Yuan, Jiawei Wang and Shenggen Ju*

College of Computer Science, Sichuan University, Chengdu 610065, China

Abstract

Answer selection (AS) is an important subtask of question answering(QA) that aims to choose the most suitable answer from a list of candidate answers. Existing AS models usually explored the single-scale sentence matching, whereas a sentence might contain semantic information at different scales, e.g.Word-level, Phrase-level, or the whole sentence. In addition, these models typically use fixed-size feature vectors to represent questions and answers, which may cause information loss when questions or answers are too long. To address these issues, we propose an Encoder-Decoder Network with Cross-Match Mechanism (EDCMN) where questions and answers that represented by feature vectors with fixed-size and dynamic-size are applied for multiple-perspective matching. In this model, Encoder layer is based on the "Siamese" network and Decoder ayer is based on the "matching-aggregation" network. We evaluate our model on two tasks: Answer Selection and Textual Entailment. Experimental results show the effectiveness of our model, which achieves the state-of-the-art performance on WikiQA dataset

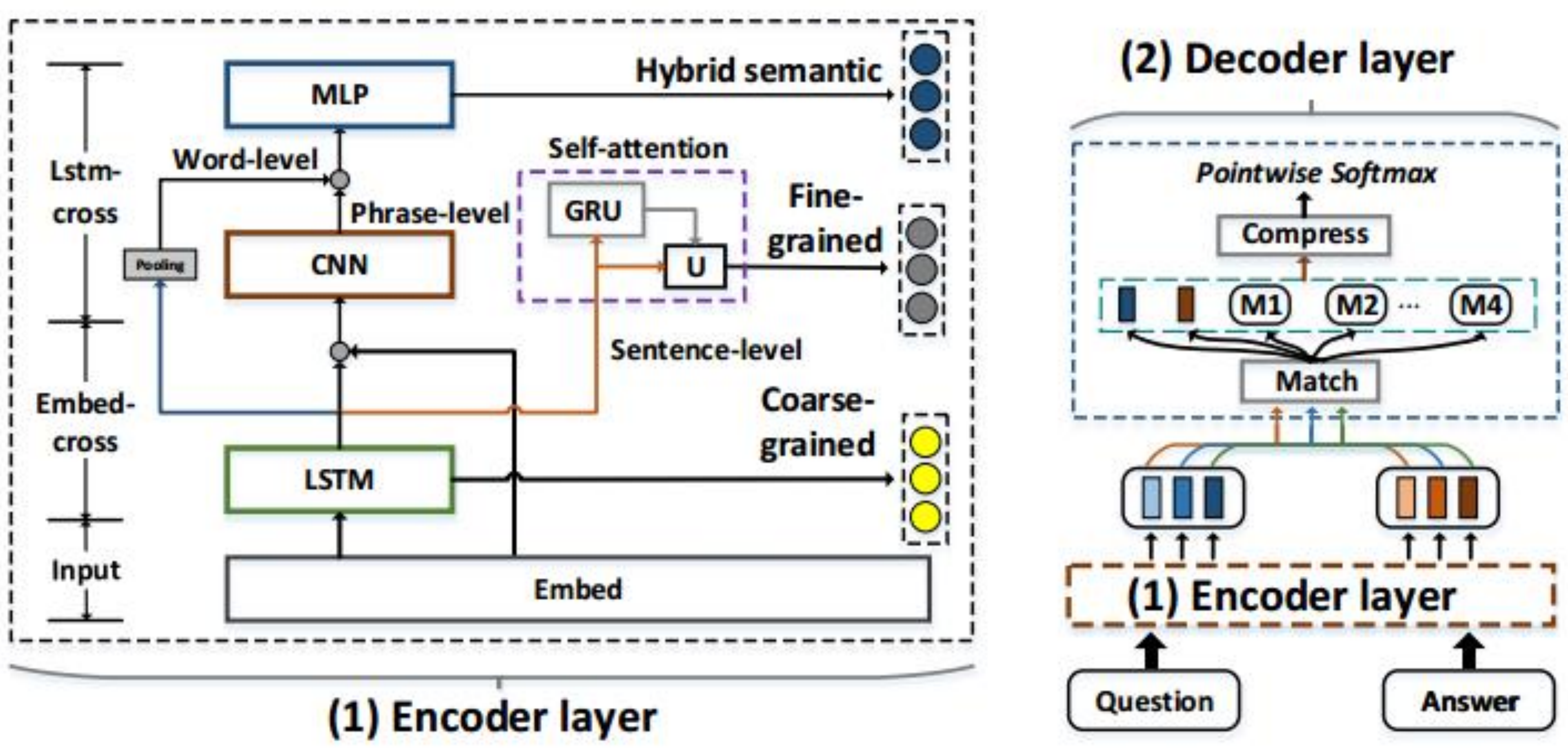

Fig. 1. The Architecture for Encoder-Decoder Network with Cross-Match Mechanism (EDCMN), The left-hand side shows the details about the Encoder layer.

Cross-Match Mechanism

What is cross-matching mechanism？

the cross-matching mechanism is mainly composed of an encoder layer and a decoder layer. The encoder layer obtains sentence representations at **Word-leve**l, **Phrase-level** and **Sentence-level** by LSTM, CNN and Self-attention components respectively. The decoder layer first applies Match functions to get six interaction representations about QA pairs, and then merge them to the final vector with the Compress function

Why cross-matching mechanism works？

The biggest benefit of the Cross-Match Mechanism is that it realizes short-circuit connection through concatenate feature, which makes some of the features extracted from earlier layers **may still be used directly by deeper layers**, meanwhile, it can match vectors in multiple ways

Experimental Results and Ablation Analysis

In the Encoder layer, we remove Word-level feature, The MAP drops by 1.5% and MRR drop by 1.9%. It shows that **Word-level information is a useful supplement** to the model. we get rid of the Conv1D segment and find that the influence is huge, the MAP and MRR drop by 3.1% and 4.1% respectively which confirms the **effect-tiveness of Phrase-level feature**.

In the Decoder layer, We can observe that **Mat feature** reduction had caused the most enormous influence (6%) in these match methods, indicating matching feature vectors with dynamic-size successfully acquires many rich characteristics of sentences

Table 4. Ablation analysis for Answer Selection On WikiQA test set.

Model structure	MAP	MRR
Full Model	0.740	0.752
(1) w/o Embed-cross fea	0.731	0.739
(2) w/o Word-level fea	0.725	0.733
(3) w/o Phrase-level fea	0.709	0.711
(4) w/o Sentence-level fea	0.724	0.735
(5) w/o Sub fea (fixed-size)	0.720	0.733
(6) w/o Mul fea (fixed-size)	0.722	0.732
(7) w/o Mat-fea (dyna-size)	0.681	0.695
(8) w/o Match	0.672	0.675

Table 2. Performance for answer sentence selection on WikiQA and TREC-QA test set.

Model	WikiQA		TREC-QA (clean)	
	MAP	MRR	MAP	MRR
AP-BiLSTM [12]	0.671	0.684	0.713	0.803
ABCNN [27]	0.692	0.711	0.777	0.836
PWIM [5]	0.709	0.723	0.738	0.827
LDC [24]	0.706	0.723	0.771	0.845
IARNN [20]	0.734	0.742	-	-
BiMPM [23]	0.718	0.731	0.802	0.875
IWAN [14]	0.733	0.750	0.822	0.889
MCAN-SM [18]	-	-	0.827	0.880
MAN [19]	0.722	0.738	0.813	0.893
EDCMN (proposed)	**0.740**	**0.752**	**0.811**	**0.896**

Conclusion

In this paper, we propose an Encoder-Decoder Network with Cross-Match Mechanism,where the encoder layer is based on the "Siamese" network and decoder layer is based on the "matching-aggregation" network. The Cross-Match Mechanism which is the core innovation captures information of sentences at different scales including Sentence-level, Phrase-level, and Word-level. In addition, it explores sentence matching both on vectors with dynamic-size and fixed-size and it is more suitable for identifying complex relations between questions and the answers. In the experiments, we show that proposed model achieves state-of-the-art performance on the WikiQA dataset. In future work, we will incorporate external knowledge bases into our model to improve its performance. Furthermore, unlabeled data is much easier to obtain than labeled data. We will explore unsupervised methods for answer selection.

References

1. Tran, N.K., Niedereée, C.: Multihop Attention Networks for Question Answer Matching.325–334 (2018). https://doi.org/10.1145/3209978.3210009.

2. Wang, B. et al.: Inner Attention based Recurrent Neural Networks for Answer Selection.1288–1297 (2016). https://doi.org/10.18653/v1/p16-1122.

3. Khot, T. et al.: SCITAIL: A Textual Entailment Dataset from Science Question Answering.AAAI Conf. Artif. Intell. (2018).

Acknowledgements

This research was partially supported by the Sichuan Science and Technology Program under Grant Nos. 2018GZ0182, 2018GZ0093 and 2018GZDZX0039.