DEEP HYBRID NETWORKS BASED RESPONSE SELECTION FOR MULTI-TURN DIALOGUE SYSTEMS

Xishuo Li, Lijun Zhang, Wenge Rong, Baiwen Li, Qi Li
From Beihang University, China

Introduction

Proper response selection is an important challenge for a meaningful **multi-turn dialogue**.

To this end, not only the coherence among the whole dialogue but also the interaction between utterance in adjacent turns need to be properly employed as the context for response selection.

In this paper, we propose a **deep hybrid network (DHN)** to distill such contextual information.

The empirical study on two commonly used public datasets has shown the proposed model's potential.

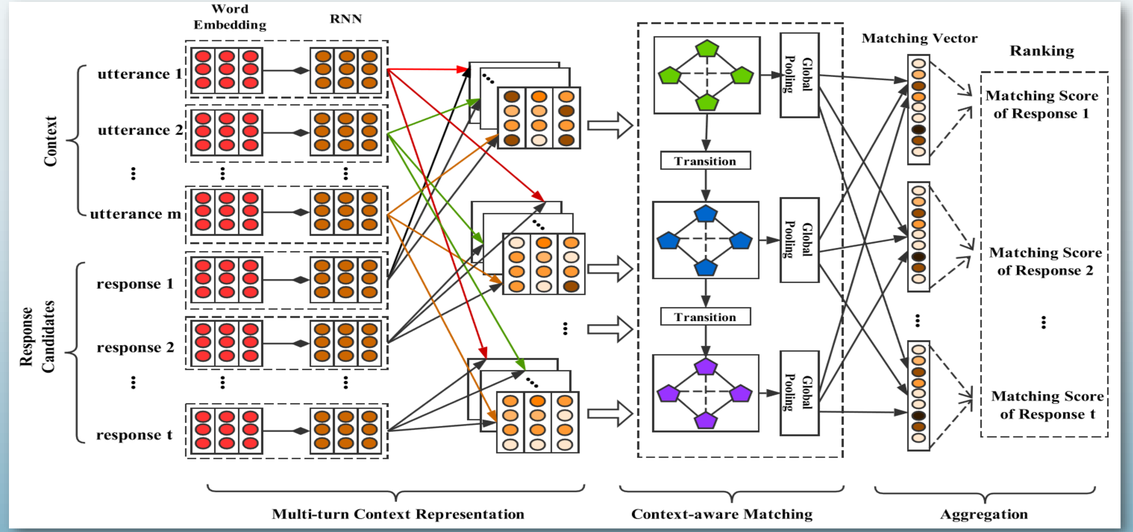
Methodology

The architecture of DHN can be roughly divided into three parts, namely **multi-turn context representation, context-aware matching, and aggregation**.

The multi-turn context representation match each candidate and all previous utterances with RNN.

Next, deep convolutional blocks will extract such contextual representation and output a matching vector.

Finally, the matching vector will be further analyzed to generate a matching score, and all response candidates' machine scores will be sorted.



Experiments

We conduct experiment study on **two public datasets.**

To evaluate the proposed method, employ **Rn@k** (recall at position k in n candidates), **MAP** (Mean Average Precision), **MRR** (Mean Reciprocal Rank) and **P@1** (Precision-at-one) to evaluate our experimental results.

Table 1. Comparison results on two public datasets

Model	Ubuntu Corpus			Douban Conversation Corpus					
	R_{10}@1	R_{10}@2	R_{10}@5	MAP	MRR	P@1	R_{10}@1	R_{10}@2	R_{10}@5
LSTM	0.638	0.784	0.949	0.485	0.537	0.320	0.187	0.343	0.720
biLSTM	0.630	0.780	0.944	0.479	0.514	0.313	0.184	0.330	0.716
MV-LSTM	0.653	0.804	0.946	0.498	0.538	0.348	0.202	0.351	0.710
Match-LSTM	0.653	0.799	0.944	0.500	0.537	0.345	0.202	0.348	0.720
Attentive-LSTM	0.633	0.789	0.943	0.495	0.523	0.331	0.192	0.328	0.718
DL2R	0.626	0.783	0.944	0.488	0.527	0.330	0.193	0.342	0.705
Multi-view	0.662	0.801	0.951	0.505	0.543	0.342	0.202	0.350	0.729
SMN	0.726	0.847	**0.961**	0.529	0.569	0.397	0.233	0.396	0.724
Our proposed	**0.733**	**0.852**	**0.961**	**0.562**	**0.621**	**0.432**	**0.241**	**0.407**	**0.754**

Related work

Modeling multi-turn conversations are under the spotlight to **utilize valuable contextual information** to build an intelligent dialogue system.

Lowe et al. tried to concatenate the context together, and Yan et al. reformulate the context by selecting utterances, while Wu et al. match responses via constructing word-level and segment-level similarity matrixes.