Dually Interactive Matching Network for Personalized Response Selection in Retrieval-Based Chatbots

Jia-Chen Gu[1], Zhen-Hua Ling[1], Xiaodan Zhu[2] and Quan Liu[1,3]

[1]National Engineering Laboratory for Speech and Language Information Processing, University of Science and Technology of China

[2]ECE, Queen's University, [3]State Key Laboratory of Cognitive Intelligence, iFLYTEK Research

Introduction

Personalized response selection is a task to select a correct response from a list of candidates considering:

(1) the context of the conversation

S1: Hello, how are you doing tonight?
S2: I am well an loving this interaction.
S1: I just got back from the club.
S2: This is my favorite season of a year.
...

(2) the persona of the speaker

S1: I like to dance at the club.
I like taking and posting selkies.
...
S2: I love to meet new people.
Autumn is my favorite season.
...

Problem 1: The context is treated at coarse-grained context-level while each utterance may play different roles.
Method: Compute similarities between each context utterance and each profile at fine-grained utterance-level.

Problem 2: The interactions between the persona and each response candidate are ignored.
Method: Add interactions between the persona and each response.

IMN-Based Persona Fusion

The representation vectors of context, context utterances and persona profiles are denoted as $\mathbf{c}$, $\{\mathbf{u}_m\}_{m=1}^{n_c}$ and $\{\mathbf{p}_n\}_{n=1}^{n_p}$ respectively.

Context-level persona fusion:

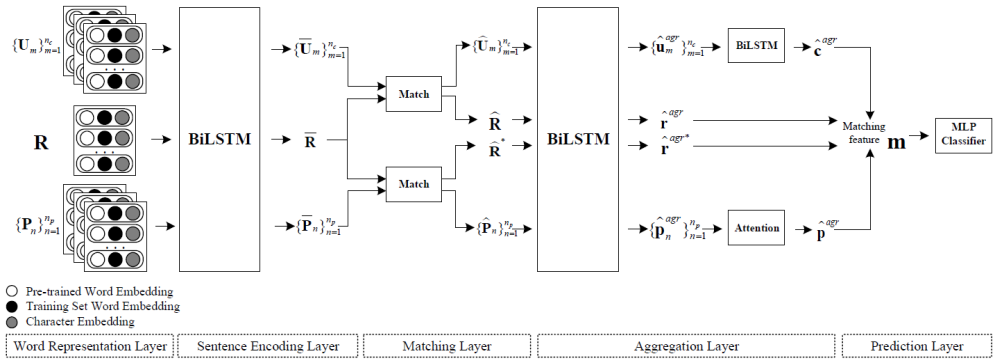
$$\mathbf{c}^+ = \mathbf{c} + \sum_n \mathbf{Softmax}(\mathbf{c} \cdot \mathbf{p}_n)\mathbf{p}_n$$

Utterance-level persona fusion:

$$\mathbf{u}_m^+ = \mathbf{u}_m + \sum_n \mathbf{Softmax}(\mathbf{u}_m \cdot \mathbf{p}_n)\mathbf{p}_n$$

$$\mathbf{c}^+ = \mathbf{Aggregation}(\{\mathbf{u}_m^+\}_{m=1}^{n_c})$$

Dually Interactive Matching Network



To add another interactive matching between the persona and each response, we formulate the task as a dual matching problem.

(1) Word Representation Layer: general, task-specific and character embeddings.
(2) Sentence Encoding Layer: a single-layer BiLSTM to encode each utterance, response and each profile separately.
(3) Matching Layer: a cross-attention module to derive the matching information by the interactive matching between the context and response, and that between the persona and response.
(4) Aggregation Layer: to convert matching matrices into the final matching vector.
(5) Prediction Layer: a MLP classifier to return a score denoting the matching degree.

Experiments

	Self Persona				Their Persona			
	Original		Revised		Original		Revised	
	hits@1	**MRR**	**hits@1**	**MRR**	**hits@1**	**MRR**	**hits@1**	**MRR**
IR baseline	41.0 (+19.6)	-	20.7 (-0.7)	-	18.1 (-3.3)	-	18.1 (-3.3)	-
Starspace	48.1 (+16.3)	-	32.2 (+0.4)	-	24.5 (-7.3)	-	26.1 (-5.7)	-
Profile	47.3 (+15.5)	-	35.4 (+3.6)	-	28.3 (-3.5)	-	29.4 (-2.4)	-
KV Profile	51.1 (+16.2)	-	35.1 (+0.2)	-	29.1 (-5.8)	-	28.9 (-6.0)	-
FT-PC	-	-	60.7 (-)	-	-	-	-	-
IMN_{ctx}	64.3 (+0.5)	76.2 (+0.4)	63.8 (+0.0)	75.8 (+0.0)	63.7 (-0.1)	75.8 (+0.0)	63.5 (-0.3)	75.7 (-0.1)
IMN_{utr}	66.7 (+2.9)	78.1 (+2.3)	64.0 (+0.2)	76.0 (+0.2)	63.9 (+0.1)	75.9 (+0.1)	63.7 (-0.1)	75.7 (-0.1)
DIM	**78.8 (+15.0)**	**86.7 (+10.9)**	**70.7 (+6.9)**	**81.2 (+5.4)**	**64.0 (+0.2)**	**76.1 (+0.3)**	**63.9 (+0.1)**	**76.0 (+0.2)**

Dataset: **PERSONA-CHAT**
Baseline IMN_{ctx}: IMN-based persona fusion at the context-level
Baseline IMN_{utr}: IMN-based persona fusion at the utterance-level
Numbers in parentheses indicate gains or losses after adding the persona conditions.
DIM outperformed its baseline IMN_{ctx} by a margin of **14.5%** and previous models by margins larger than **27.7%** in terms of hits@1 conditioned on original self personas.

Analysis

	hits@1	**MRR**
DIM	78.8	86.7
- persona	63.8	75.8
- context	48.8	60.9

(1) **Ablation tests** of removing either persona-response matching or context-response matching in DIM show both contribute.

Train \ Test	Original	Revised
Original	78.8	66.3
Revised	77.6	70.7

(2) **Transfer tests** were conducted by training and evaluating using mismatched types of personas.

Conclusion

In this paper, we formulate the task of personalized response selection as a dual matching problem. A new model named Dually Interactive Matching Network (DIM) is proposed, which achieves a new state-of-the-art performance on PERSONA-CHAT set.

Reference

Zhang S, Dinan E, Urbanek J, Szlam A, Kiela D, and Weston J. Personalizing dialogue agents: I have a dog, do you have pets too? ACL' 18

Gu J, Ling Z and Liu Q. Interactive matching network for multi-turn response selection in retrieval-based chatbots. CIKM' 19

Contact

Email:
gujc@mail.ustc.edu.cn
Github:
github.com/JasonForJoy/DIM
Homepage:
home.ustc.edu.cn/~gujc/