Improving Role-Oriented Dialogue Summarization with Interaction-Aware Contrastive Learning

Weihong Guan, Shi Feng, Daling Wang, Faliang Huang, Yifei Zhang, Yuan Cui

Northeastern University, Shenyang, China

LREC-COLING 2024, **Torino**

1-Introduction

- In many real life dialogue scenarios, each speaker has an official role (e.g. user or agent) acting for corresponding responsibility.
- Each role has its own view and goal. It is possible and important to summarize the main content of each role.
- Lin et al. (2021) proposed the role-oriented dialogue summarization task and construct a related-dataset CSDS based on Chinese customer service dialogues.

1-Introduction

Dialogue
0 A: 有什么可以帮助您的吗。(Is there anything I can help with?)
1 Q: 我购买的商品可以更换地址吗? (Can I change the address of the goods I purchased?)
2 A: 下单之后是改不了的呢。(It cannot be changed after placing the order.)
3 Q: 那我取消吧。(Then I'll cancel.)
2 A: 是哪个订单的呢。(Which order is it?)
3 Q: [订单编号]。([Order Number])
6 Q: 我已经取消了。(I have canceled.)
7 A: 全部取消吗? 您只取消了一部分呢(Cancel all? You only canceled a part of it.)
8 Q: 请全部取消，其他的显示关联订单无法取消。(Please cancel all of it. Other parts cannot be canceled because of associated orders.)
8 A: 好的，您稍等。这个订单是三方商家的，需要商家审核一下。(Okay, just a moment. This order is from a third-party merchant and needs to be reviewed by the merchant.)
9 Q: 好的 。(Ok.)
10 A: 请问还有其他需要帮忙的吗。(Is there anything else you need help with?)
10 Q: 没有了,谢谢。(That is all, thanks.)
11 A: 感谢您的支持，再见。(Thank you for your support, goodbye.)

User Summary	用户询问购买的商品能否更换地址。用户希望客服帮助取消全部订单。(The user asked whether the address of the purchased product could be changed. The user wants the agent to help cancel all orders.)
Agent Summary	客服回答用户修改地址即可。客服回答如果取消订单用的券可以退回。(The agent replied that the user can modify the address. The agent replied that coupons used can be returned if the order is canceled.)
Final Summary	用户询问购买的商品能否更换地址。客服表示下单后不可以。用户希望客服帮助取消全部订单。客服表示需要商家审核。(The user asks whether the address of the purchased product could be changed. The agent says it was not possible after placing the order. The user wants the agent to help cancel all orders. The agent says it needs to be reviewed by the merchant.)

- Interaction between different roles is crucial for the role-oriented dialogue summarization.
  - Interaction can help **track the key information** scattered across different roles
  - Interaction can help **understand the main content of each role**

1-Introduction

- Existing methods do not build mechanisms for the encoder to understand **interaction patterns** between different roles.
- The encoder can not fully comprehend the **role-level interaction** from the flat-concatenated utterances, thus the summarization model are prone to ignore **vital interaction clues** in dialogue and generate inaccurate summaries .
- We propose a **C**ontrastive Learning based **I**nteraction-**A**ware **M**odel (CIAM) for the role-oriented dialogue summarization task.

1-Introduction

Our Contributions：

- We propose an interaction aware contrastive learning method, which could help the model capture interaction patterns between different roles
- We employ decoder start tokens to control what kind of summary to generate which could generate different summaries with a unified model and can utilize the relatedness between different role's summaries.
- Our method can be applied to different seq2seq models and can outperform previous SOTA models and powerful LLMs on two public datasets CSDS and MC.

2-Method

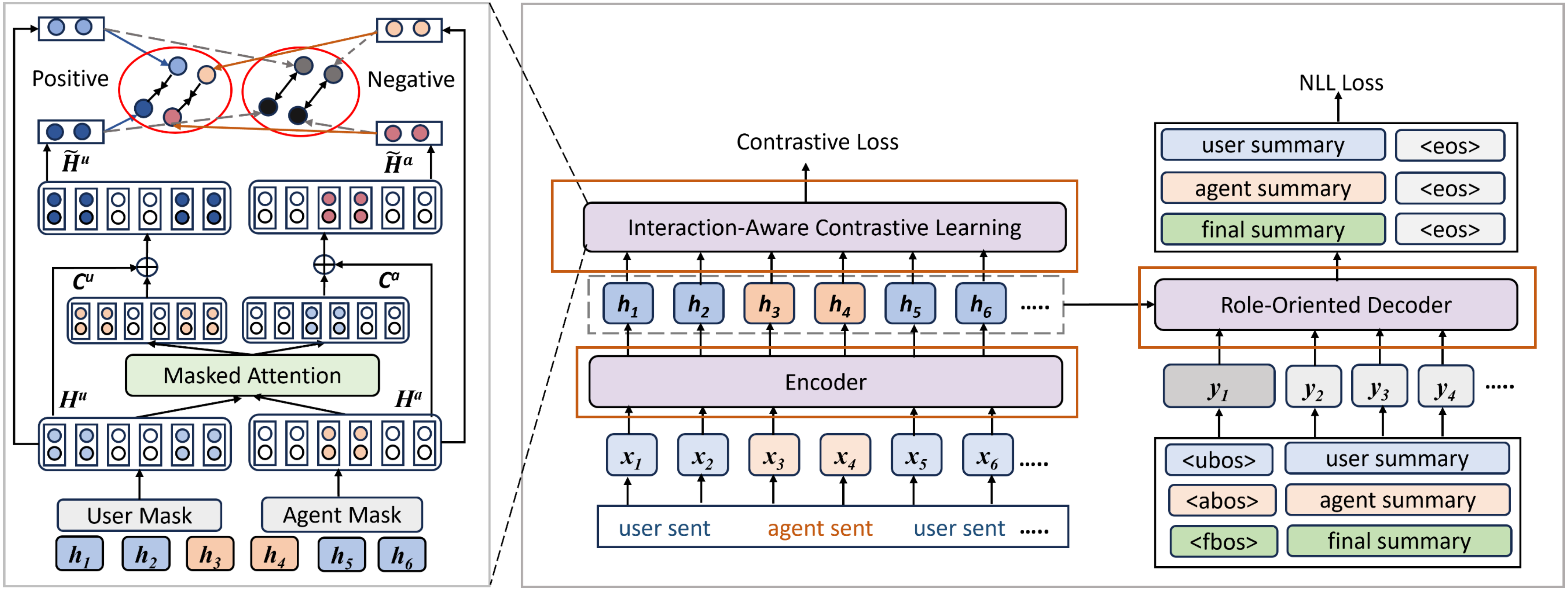
Task Formulation :

Source: A dialogue D containing m utterances $\{u_1, \dots, u_m\}$ and p roles $R = \{r_1, \dots, r_p\}$. By concatenating all the utterances and related speaker roles, we obtain the final input $\{x_1, \dots, x_n\}$.

Target: A summary for each role: y^{user}, y^{agent} and a summary for the whole dialogue y^{final}

All public datasets for this task have two roles, one asking questions (user, patient) and one answering questions (agent, doctor).

2-Method



Encoder:

$\{h_1, \dots, h_n\} = \text{Encoder}\ (\{x_1, \dots, x_n\})$

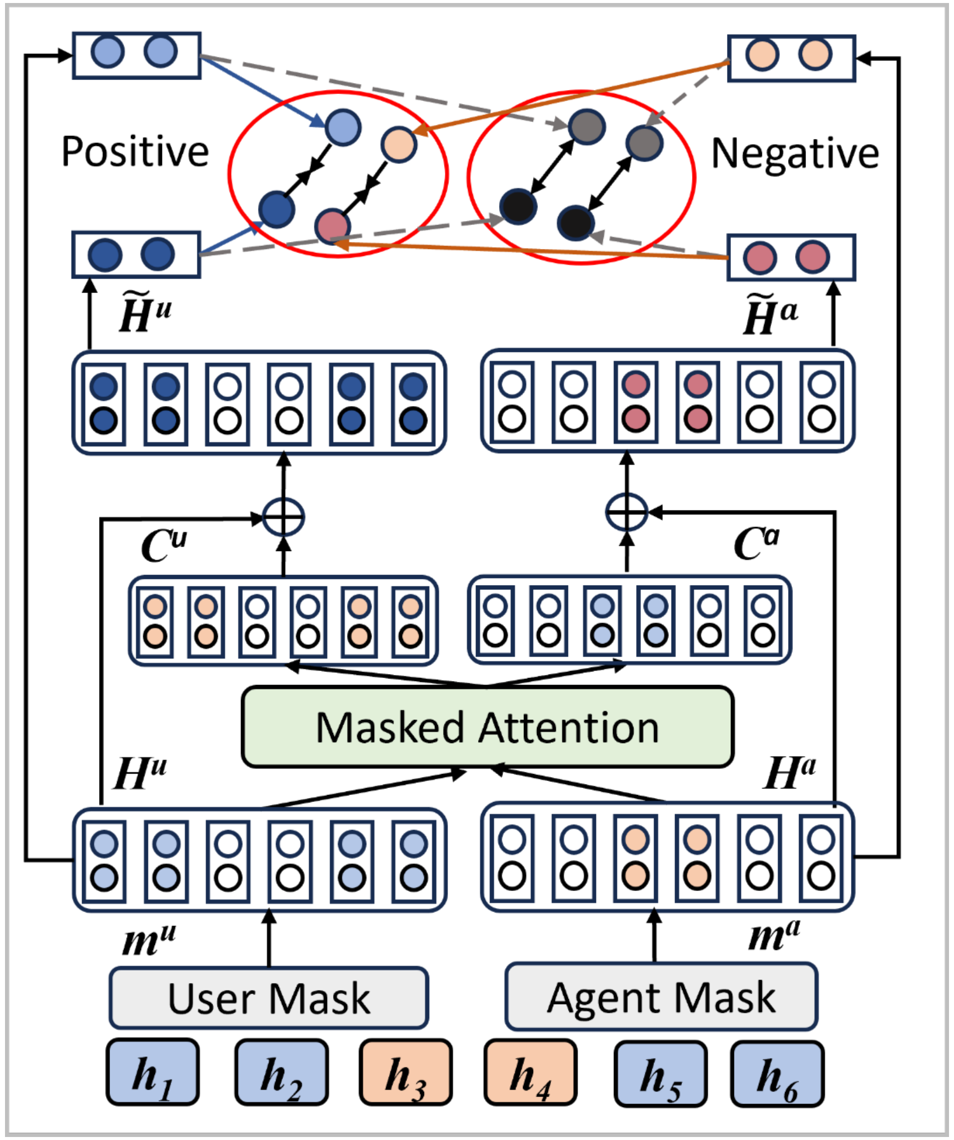
Role-Oriented Decoder:

employ different decoder start tokens to control what kind of summary to generate

$$-\sum_{i=1}^{T} \log P(y_i \mid y_{<i}, bos, X)$$

Interaction-Aware Contrastive Learning

2-Method



Self-Representations

$$H^u = H \odot m^u,$$
$$H^a = H \odot m^a,$$

Cross-Representations

$$A_{u2a} = \text{softmax}\left(\frac{H^u (H^a)^T}{\sqrt{d}} + M_{u2a}\right)$$
$$A_{a2u} = \text{softmax}\left(\frac{H^a (H^u)^T}{\sqrt{d}} + M_{a2u}\right)$$

$$C^u = (A_{u2a} \odot m^u) H^a$$
$$C^a = (A_{a2u} \odot m^a) H^u$$

Interaction-Aware Representations

$$\tilde{H}^u = (H^u \oplus C^u) W^u,$$
$$\tilde{H}^a = (H^a \oplus C^a) W^a,$$

Positive Sample: The original dialogue D

Negative Samples: destroying interactions between different roles in dialogue

- Keep all utterances of the user and mask all utterances of the agent.
- Then randomly select a different dialogue $\tilde{D}$ and sample a consecutive of utterances of the agent in $\tilde{D}$ to fill the masked utterances of D. Repeat the process until all the masked utterances in D has been filled.
- Repeat this operation multiple times and generate K negative samples.
- Keep all utterances of the agent and replace all utterances of the user, obtaining another K negative samples. Finally, obtain 2K+1 samples, the first sample is the positive sample
- We generate H_i^u , H_i^a , $\tilde{H}_i^u$, $\tilde{H}_i^a$ for the i-th sample.

InfoNCE Loss

$$\mathcal{L}_{user} = -\log \frac{e^{sim\left(H_1^u, \tilde{H}_1^u\right)/\tau}}{\sum_{j=1}^{(2K+1)} e^{sim\left(H_j^u, \tilde{H}_j^u\right)/\tau}},$$

$$\mathcal{L}_{agent} = -\log \frac{e^{sim\left(H_1^a, \tilde{H}_1^a\right)/\tau}}{\sum_{j=1}^{(2K+1)} e^{sim\left(H_j^a, \tilde{H}_j^a\right)/\tau}},$$

$$\mathcal{L}_{con} = \mathcal{L}_{user} + \mathcal{L}_{agent}.$$

Total Loss

$$\mathcal{L} = \mathcal{L}_{nll} + \gamma\mathcal{L}_{con}$$

4-Experiments

Dataset:

CSDS: Customer Service
MC: Medical Inquiry

Baselines:

- **BERTAbs** and **BART** (Lewis et al., 2020)
- **BERT-CIAM** and **BART-CIAM**
- **BERT-both** and **BART-both**: assign each role a specific decoder and employ two attentions to model interaction between roles
- **BERT-RAC** and **BART-RAC**: employ a discrete role prompt to control model for different summaries, and employ a centrality model to capture salient utterances
- **BERT-GLC** and **BART-GLC**: employ global-to-local centrality scores to capture sub topics on the basis of RAC

4-Experiments

CSDS	ROUGE-1	ROUGE-2	ROUGE-L	BERTScore
BERT	55.41/52.71/49.61	39.42/36.39/33.88	53.41/50.45/46.88	78.52/79.23/76.39
BERT-Both	57.24/57.36/51.92	40.12/40.70/36.37	54.87/55.17/49.52	79.85/80.70/77.23
BERT-RAC	57.35/57.75/52.23	40.34/41.05/36.75	55.12/55.53/49.89	79.89/80.69/77.27
BERT-GLC	57.59/58.14/52.34	**41.28**/41.84/36.48	**55.74**/55.86/50.16	79.89/80.71/77.28
BERT-CIAM	**57.66**/**58.73**/**52.55**	41.12/**42.01**/**36.92**	55.51/**56.72**/**50.20**	**79.90**/**81.25**/**77.39**
w/o IACL	56.24/57.20/51.21	40.28/40.92/36.09	54.38/55.29/49.16	79.61/80.89/76.79
BART	58.66/60.35/54.13	43.35/45.09/39.37	56.60/58.13/51.18	79.54/81.14/77.31
BART-Both	59.21/60.53/54.22	43.88/45.39/39.96	57.32/58.28/51.90	79.74/81.37/77.41
BART-RAC	59.86/61.67/54.83	44.42/46.14/40.29	57.86/59.45/52.43	79.97/81.92/77.60
BART-GLC	60.07/61.72/54.82	44.55/46.21/40.11	58.06/59.51/52.46	**80.10**/81.90/77.61
BART-CIAM	**60.27**/**62.21**/**55.04**	**44.63**/**46.35**/**40.46**	**58.20**/**59.88**/**52.69**	80.01/**82.03**/**77.63**
w/o IACL	59.39/61.69/54.68	43.85/46.13/40.12	57.34/59.37/52.39	79.77/81.88/77.58

Ablation Study:

w/o IACL: without interaction-aware contrastive learning

4-Experiments

MC	ROUGE-1	ROUGE-2	ROUGE-L	BERTScore
BERT	84.07/95.13/81.66	79.90/94.50/76.73	83.04/95.08/80.42	92.68/97.86/91.71
BERT-Both	84.69/95.19/82.11	80.76/94.63/77.49	83.68/95.14/80.92	93.02/97.90/91.91
BERT-RAC	85.12/95.50/82.62	81.30/94.80/77.91	84.07/95.72/81.36	93.11/97.89/92.29
BERT-GLC	85.64/95.49/82.87	81.44/94.97/78.05	84.16/96.10/81.57	93.15/97.92/**92.36**
BERT-CIAM	**85.87/95.96/83.04**	**81.82/94.98/78.53**	**84.62/96.11/81.69**	**93.39/98.16**/92.14
w/o IACL	84.83/95.20/82.25	80.93/94.59/77.63	84.02/95.21/80.96	93.05/97.88/91.95
BART	88.37/95.42/86.33	84.75/94.99/82.33	87.38/95.37/85.30	93.65/97.94/92.63
BART-Both	88.52/95.63/87.06	85.22/95.42/82.89	87.55/95.96/85.79	93.72/97.89/92.67
BART-RAC	89.43/96.78/88.21	86.29/95.86/84.58	88.47/96.12/86.56	94.01/98.13/92.84
BART-GLC	89.55/96.84/88.47	86.47/96.14/**84.62**	88.56/96.23/86.77	94.17/98.25/**92.96**
BART-CIAM	**89.85/96.86/88.73**	**86.93/96.31**/84.56	**88.83/96.74/86.84**	**94.26/98.55**/92.90
w/o IACL	88.87/96.23/87.95	85.78/95.84/83.71	87.96/96.10/86.14	93.89/98.06/92.78

4-Experiment

Comparasion with LLMs

➢ **ChatGPT** (gpt-3.5-turbo-0613) and **GPT4** (gpt-4-0613)

CSDS	ROUGE-1	ROUGE-2	ROUGE-L	BERTScore
BART-CIAM	**60.27/62.21/55.04**	**44.63/46.35/40.46**	**58.20/59.88/52.69**	**80.01/82.03/77.63**
ChatGPT	51.96/48.85/46.21	33.89/31.76/28.56	48.25/44.98/42.25	75.46/76.74/73.90
GPT4	53.04/49.78/47.97	35.10/32.98/30.97	49.72/46.32/43.98	77.12/77.15/75.10

给定一段中文客服对话，请分别生成用户视角的摘要，客服视角的摘要和整段对话的总体摘要。(Given a Chinese customer service dialogue, please generate a summary for the user, a summary for the agent, and an overall summary for the whole dialogue.)

our examples:)

{ User Summary }

客服摘要：(Agent Summary:)

{ Agent Summary }

总体摘要：(Overall Summary:)

{ Final Summary }

.....

请为下面的测试对话生成用户摘要，客服摘要和总体摘要。(Please generate the user summary, the agent summary, the overall summary for the following test dialogue:)

对话：(Dialogue:)

{ Dialogue }

用户摘要：(User Summary:)

客服摘要：(Agent Summary:)

总体摘要：(Overall Summary:)

4-Experiment

Human Evaluation

Score (0,1,2):

- **Informativeness**: whether the generated summary could correctly contain the key information
- **Conciseness**: whether the generated summary could avoid redundant and unnecessary information

Human Preference:

Ask evaluators to select a best summary or several best summaries from the generated summaries of different models

CSDS	Informativeness	Conciseness	Human Preference
GPT4	1.45/1.38/1.34	1.29/1.16/1.19	0.40/0.30/0.29
BART	1.52/1.48/1.42	1.10/1.18/1.16	0.19/0.38/0.31
BART-GLC	1.48/1.44/1.45	1.43/1.46/1.53	0.45/0.45/0.40
BART-CIAM	**1.59/1.51/1.53**	**1.54/1.58/1.56**	**0.57/0.56/0.53**

Thanks for watching!