

MUIDIAL: Improving Dialogue Disentanglement with Intent-Based Mutual Learning

基于用户意图编码和互学习的对话解耦技术

江子攸, 石琳*, Celia Chen, 沐方文, 张煜旻, 王青*

The 31th International Joint Conference on Artificial Intelligence (IJCAI-22)

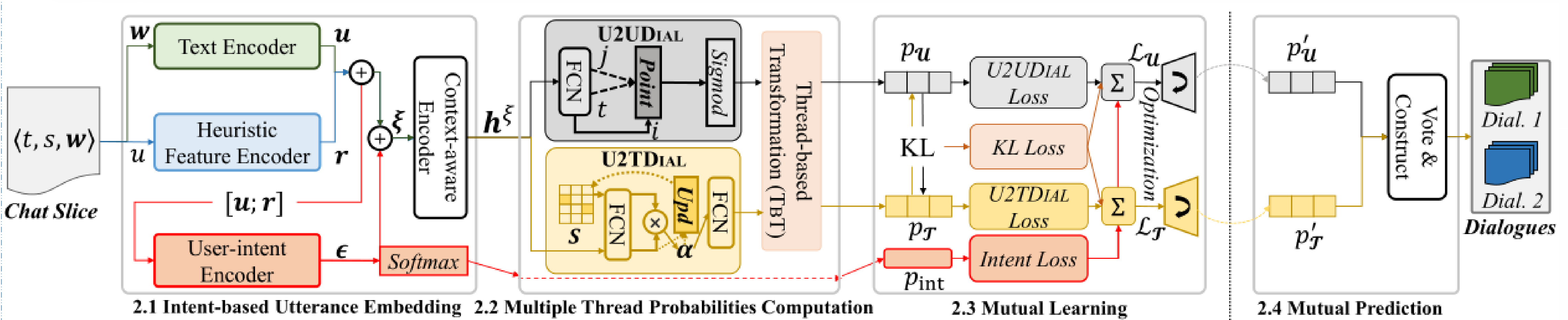
主要联系人: 石琳 (15001193593, shilin@iscas.ac.cn)

Introduction

- Existing models often utilize either an utterance-to-utterance (U2U) prediction to determine whether two utterances that have the “reply-to” relationship belong to one dialogue, or an utterance-to-thread (U2T) prediction to determine which dialogue-thread a given utterance should belong to.
- We propose MUIDIAL, a novel dialogue disentanglement model:
 - We exploit the user intent in embedding each utterance;
 - Inspired by mutual learning, we utilize a mutual learning framework to train the model.

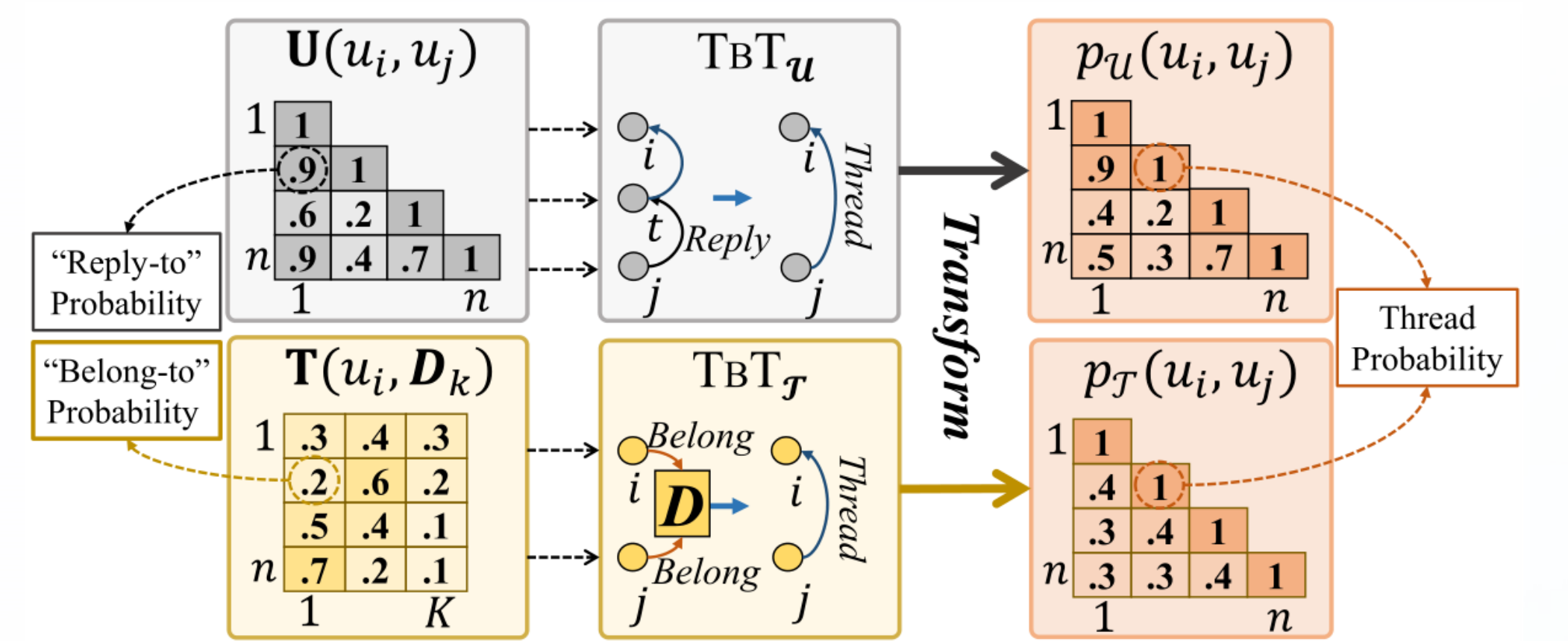
Utter.	Timestamp	Speaker	Textual Message	Intent
1	[11:31]	S1	Hello everyone, morning to Gitter!	Greeting
2	[11:31]	S1	Hello? Can anyone help me on bundling Angular 2 app into a 'bundle.js' file, put onto Heroku.	Original Question
3	[11:33]	S2	Why I cannot run tomcat on PyCharm. Err. I am a beginner.	Original Question
4	[11:34]	S3	Hello@S1, welcome!	Greeting
5	[11:35]	S3	Screenshot plz. Give me your screenshot on your Angular2 APP.	Information Request
6	[11:36]	S3	Whenever you've meet such this problem, exactly, you can try REBOOTING your APP IDE.	Information Giving
7	[11:39]	S4	Reconnect server. Reconnect the simplest way to solve breakdown problems :).	Information Giving
8	[11:40]	S2	Thanks, I'll give a try.	Feedback
9	[11:42]	S5	The CLI makes that pretty easy. An Angular seeds and their build steps might be a good to start.	Information Giving
10	[11:43]	S1	Thank you very much, nicely done! That works.	Feedback

Methodology



Code	Label	Description of labels
OQ	Original Question	Speaker proposes the first question to initialize a dialogue.
FQ	Follow-Up Question	Speaker raises follow-up questions about the related issues.
IG	Information Giving	Speaker provides some information to other speakers.
IS	Information Seeking	Speaker seeks more information from other speakers.
FB	Feedback	Speaker provides reactions/information to solutions posted by other speakers.
OT	Others	Greetings, junk messages or other uncategorized utterances.

The categories of user-intents.



The detail of TBT.

$$U2UDIAL: u(u_i, u_j) = \begin{cases} 1.0, & \text{if } 1 \leq j = i \leq n \\ \text{Sigmod}(h_i^\xi W_u h_j^{\xi T}), & \text{if } 1 \leq j < i \leq n \end{cases}$$

$$U2TDIAL: T(u_i) = W_T' \alpha = W_T' \text{Softmax}(S_i W_T h_j^{\xi T})$$

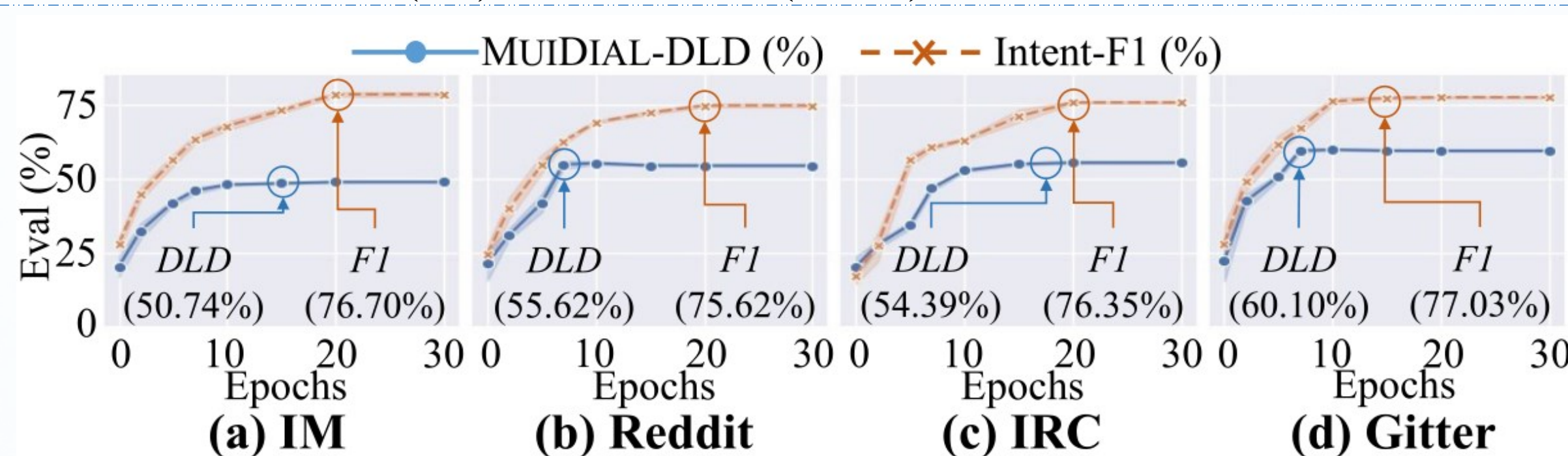
$$TBT_U: p_U(u_i, u_j) = \begin{cases} 1.0, & \text{if } 1 \leq j = i \leq n \\ \frac{1}{i-j} \sum_{j \leq t < i} U(u_i, u_t) p_U(u_t, u_j), & \text{if } 1 \leq j < i \leq n \end{cases}$$

$$TBT_T: p_T(u_i, u_j) = \begin{cases} 1.0, & \text{if } 1 \leq j = i \leq n \\ T(u_i) T(u_j)^T = (TT^T)_{ij}, & \text{if } 1 \leq j < i \leq n \end{cases}$$

Experiments

Models	Metrics	IM				Reddit				IRC				Gitter				Average			
		NMI	ARI	F1	DLD	NMI	ARI	F1	DLD	NMI	ARI	F1	DLD	NMI	ARI	F1	DLD	NMI	ARI	F1	DLD
Baselines	CISIR [Jiang et al., 2018]	20.47	6.45	12.92	25.01	65.77	32.89	35.46	47.11	46.62	3.37	20.60	27.17	64.33	45.57	40.32	48.95	49.30	22.07	27.33	37.06
	PtrNet [Yu and Joty, 2020]	21.05	8.45	13.74	20.13	68.02	31.59	30.76	45.31	60.53	37.14	44.20	54.22	71.36	51.10	46.92	48.99	55.24	32.07	33.91	42.16
	DialBERT [Li et al., 2020]	25.57	10.97	20.13	40.45	71.65	40.05	38.67	47.56	54.61	8.15	16.49	39.30	15.46	11.37	30.29	21.74	41.82	17.64	26.40	37.26
	SSE2E [Liu et al., 2020]	35.75	25.45	22.13	41.52	73.16	42.80	40.45	49.66	62.61	20.58	18.20	41.52	35.20	25.12	27.88	34.50	51.68	28.49	27.17	41.80
	CATD [Tan et al., 2019]	36.46	24.13	23.04	41.39	74.15	43.21	44.70	50.35	65.85	47.14	30.03	52.30	70.46	51.01	48.57	51.65	61.73	41.37	36.59	48.92
MUIDIAL	DAG-LSTM [Pappadopulo et al., 2021]	34.97	25.16	24.25	43.95	73.87	40.67	44.25	51.65	66.27	45.37	31.32	46.59	76.94	51.64	45.35	50.97	63.01	40.71	36.29	48.29
Variants	Intent	39.64	28.99	32.17	52.42	76.97	44.35	45.62	57.46	72.45	52.31	38.65	57.91	79.25	56.52	49.37	61.25	67.08	45.54	41.45	57.26
	U2UDIAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	U2TDIAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

- MUIDIAL achieves the best performance on all four metrics on average, improving by 4.07% (NMI), 4.17% (ARI), 4.86% (F1), and 8.34% (DLD)

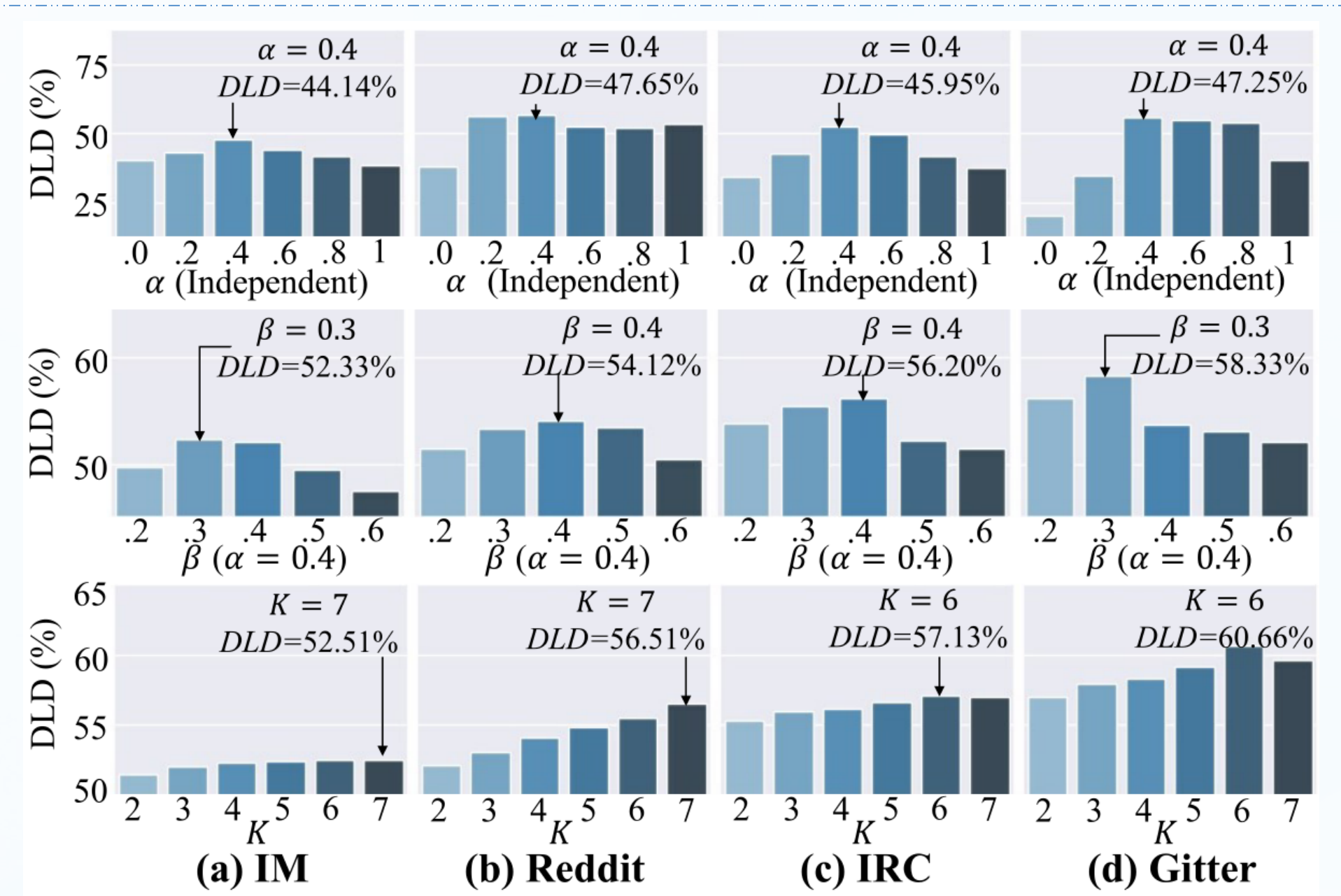


Contribution of User Intent

- The performance of intent classification is satisfactory
- The contribution of user intent to MUIDIAL is positive.

Chosen of Hyperparameters

- Intent Loss Weight α : 0.4
- Mutual Loss Weight β : 0.3/0.4
- Number of States K : 6/7



Conclusion

- We propose a novel intent-based dialogue disentanglement model MUIDIAL
- We propose the dialogue disentanglement model with mutual learning framework, which enriches the utterance embedding with user intents.
- The evaluations on four benchmark datasets show that our model outperforms the baselines by 5% on average.