Response-act Guided Reinforced Dialogue Generation for Mental Health Counseling

Aseem Srivastava[1], Ishan Pandey[1], Md. Shad Akhtar[1], Tanmoy Chakraborty[2]

[1]IIIT Delhi, India [2]IIT Delhi, India

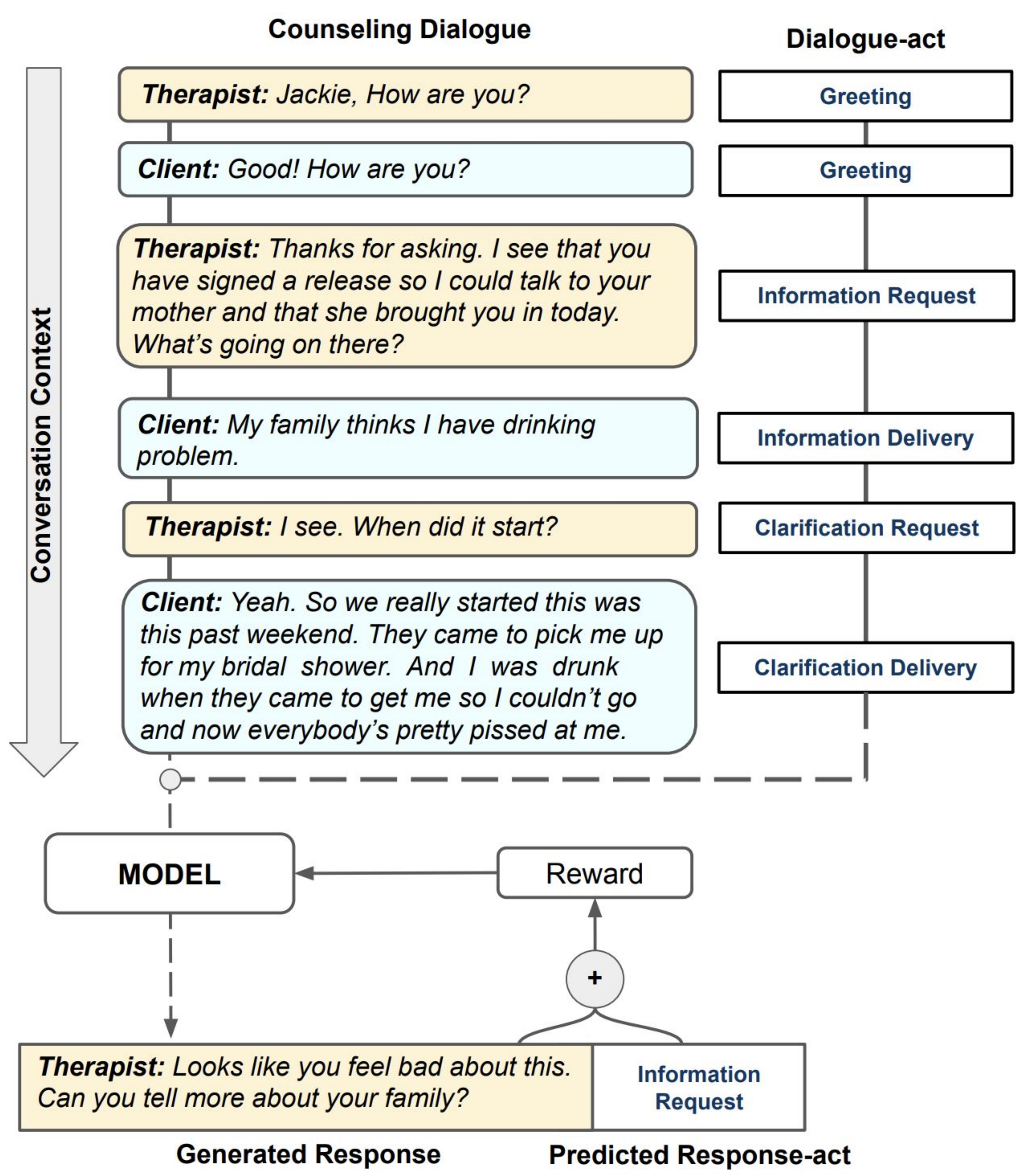
{aseems, ishan20304, shad.akhtar}@iiitd.ac.in; tanchak@iitd.ac.in

Highlights

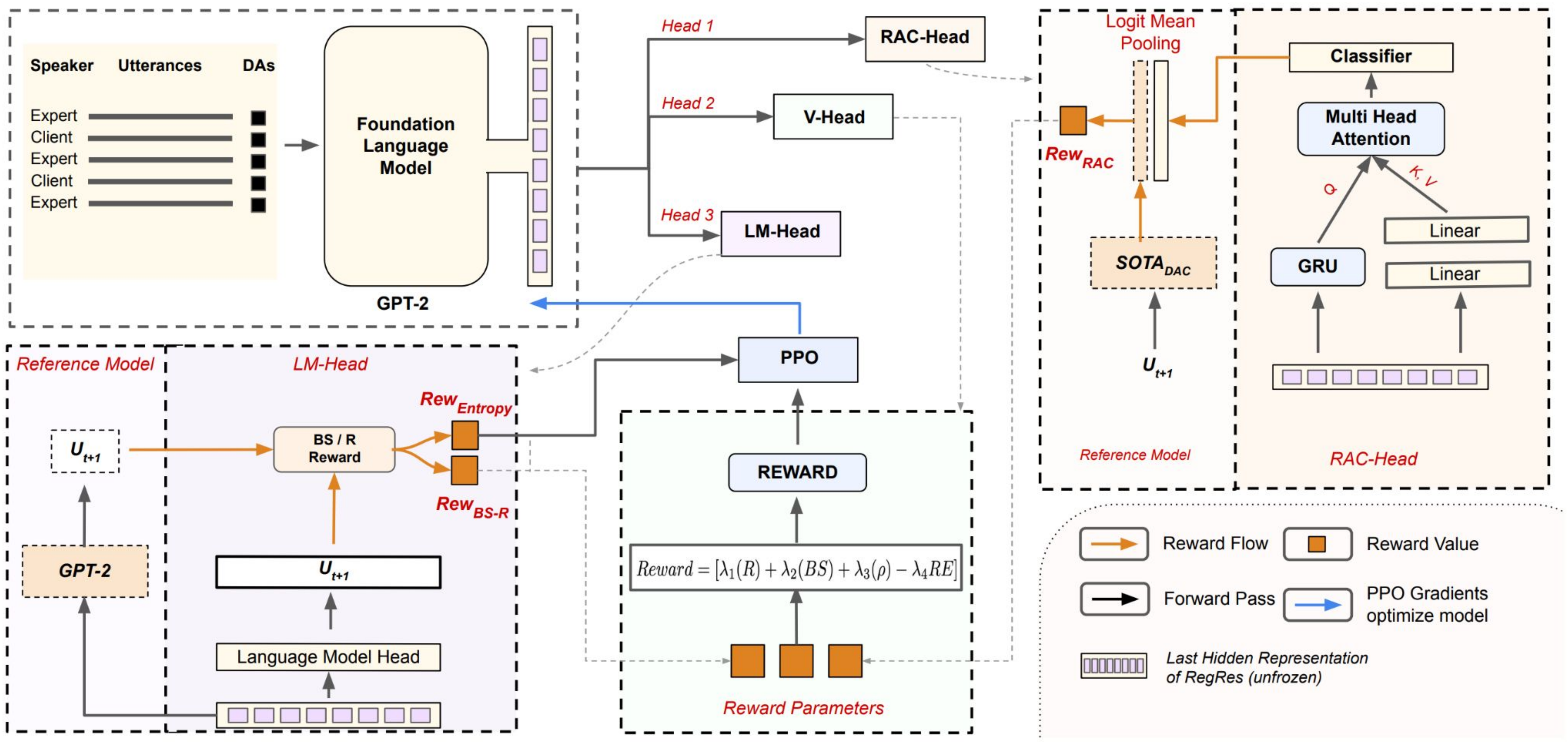
- We exploit future dialogue-acts (aka response-acts) in guiding the response generation model to generate the intended (controlled) response for virtual mental health assistants.
- We propose a novel transformer-reinforcement-learning (TRL) driven response-act guided model, READER, to generate response.
- Our evaluation on the HOPE dataset shows significant improvements in the performance of response generation over several competing baselines.
- We conduct a thorough and qualitative human evaluation on the generated responses and establish that the proposed approach is qualitatively efficient as well.

Problem Definition

Counseling Response Generation:
Our work utilizes the next dialogue-act (or response-act) to control the response generation pipeline by bifurcating the core problem into two sub-problems – (i) classifying response-act and then (ii) rewarding the model to generate responses aligning with response-act.



Methodology

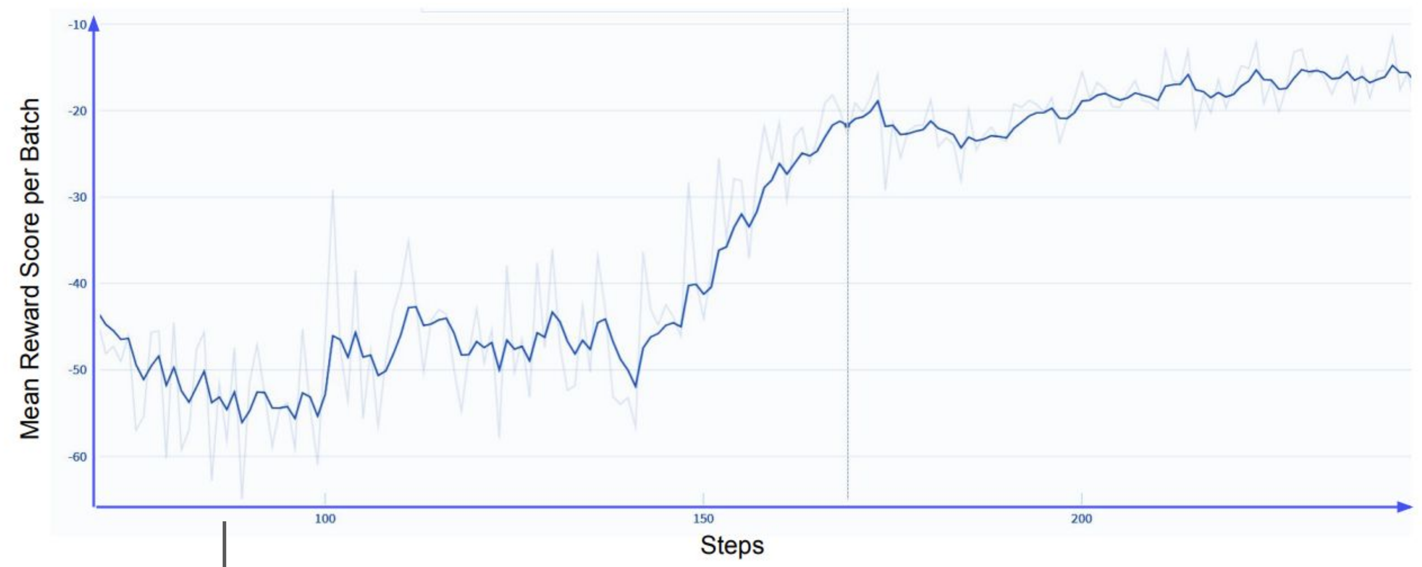


Dataset

HOPE: A mental health counseling conversation dataset. It contains 12.8K utterances from 212 dyadic counseling sessions between therapists and clients.

HOPE	Train	Validation	Test	Total
Dialogue Sessions	149	21	43	212
Client Utterances	4668	595	1119	6382
Therapist Utterances	4751	599	1122	6472
#Total Utterances	9419	1194	2241	12854

Results & Analysis



The incremental flow of reward in the PPO of READER

		R1			R2			RL			BS	METEOR
		P	R	F1	P	R	F1	P	R	F1		
Baselines	DialoGPT [34]	12.34	40.48	15.72	2.92	11.83	4.42	12.23	38.60	15.76	0.7603	0.2021
	GPT2 [21]	12.70	32.63	14.98	3.08	7.92	3.51	13.74	32.05	15.87	0.7445	0.1754
	DialogVED [4]	12.48	31.74	12.8	0.98	2.45	1.22	12.45	31.11	14.46	0.7189	0.2000
	ProphetNet [20]	12.15	34.29	14.48	3.30	10.41	4.17	12.24	33.12	15.18	0.6707	0.1901
	VHCR [17]	11.29	21.33	11.81	2.66	3.49	3.00	10.01	19.72	10.99	0.5953	0.1041
	HRED [24]	11.52	21.51	10.72	1.89	6.42	2.92	12.12	24.36	13.56	0.6259	0.1425
	HRED w/ Sp. Utt. Encoder [35]	11.77	28.63	10.08	1.29	4.19	2.06	12.25	21.27	12.72	0.6171	0.1801
Ours	RagRes w/ DialoGPT	12.41	43.91	16.12	3.70	13.72	4.98	11.92	41.02	16.30	0.7656	0.2098
	READER – RAC-Head	12.64	41.48	15.78	3.60	11.83	4.58	12.3	38.64	15.90	0.7628	0.2039
	READER	**12.82**	**43.93**	**16.15**	**3.77**	**13.67**	**4.93**	**12.51**	**40.82**	**16.32**	**0.7666**	**0.2103**
Ablations	– *Rew*(R)	11.73	38.82	14.65	2.28	8.45	2.96	11.21	35.76	14.53	0.7561	0.1840
	– *Rew*(RAC)	12.36	40.71	15.43	3.13	11.12	4.06	11.91	37.63	15.40	0.7609	0.2000
	– *Rew*(RAC + R)	11.92	38.06	14.70	2.43	8.26	3.11	11.40	34.98	14.58	0.7530	0.1874
	– *Rew*(R + BS)	12.48	41.13	15.57	3.52	11.85	4.47	12.22	38.29	15.77	0.7527	0.2092
	– *Rew*(RAC + BS)	12.01	40.45	15.18	2.72	9.93	3.52	11.46	37.05	14.97	0.7577	0.1908
	$\Delta_{\text{READER}-BEST}(\%)$	↑ 0.94	↑ 8.5	↑ 2.73	↑ 14.24	↑ 15.50	↑ 11.53	↓ 8.90	↑ 5.69	↑ 2.83	↑ 0.82	↑ 4.05

READER beats the best-performing baseline across 10 out of 11 metric with a **significant 22% increase** in R1 Score.

Ablation study supports the motivation behind each contributory module in our method with *Rew(RAC+R)* contributing the most to the model's increased performance.

Paper: dl.acm.org/doi/10.1145/3543507.3583380
Web: www.as3eem.github.io
Contact: aseems@iiitd.ac.in

References

[1] Xiuyi Chen, Jiaming Xu, and Bo Xu. 2019. A Working Memory Model for Task Oriented Dialog Response Generation. In Proceedings of the 57th ACL

[2] Prakhar Gupta, Harsh Jhamtani, and Jefrey Bigham. Target-Guided Dialogue Response Generation. In Findings of NAACL 2022.

[3] Ganeshan Malhotra, Abdul Waheed, Aseem Srivastava, Md Shad Akhtar, and Tanmoy Chakraborty. 2022. Speaker and Time-Aware Joint Contextual Learning for Dialogue-Act Classification in Counselling Conversations. In Proceedings of WSDM '22.