Improving Empathetic Response Generation in LLMs Using Direct Preference Optimization

Suhyune Son, Aram So*
Korea University, Human-inspired AI Research
{ssh5131,aram}@korea.ac.kr

Abstract

Empathy is essential for building social bonds in human interactions. While large language models (LLMs) excel in generating fluent dialogue, they often fail to provide emotional support, focusing instead on solutions. This paper presents a method to enhance empathetic response generation in conversational agents, using a dataset aligned with stages of dialectical behavior therapy. A new metric is also proposed to assess the model's ability to generate emotionally supportive responses. Experiments show significant improvements in both empathy and DBT.

Keywords: Empathetic Dialogue, Dialogue Generation

1. Introduction

Empathy is essential for effective social interactions and fosters stronger social connections through caring actions [1, 2]. In counseling and psychotherapy, it builds trust and helps relieve emotional distress [3, 4, 5, 6].

Nonetheless, challenges arise when individuals are hesitant to openly share their emotions [7, 8]. To tackle this issue, virtual therapy and AI chatbots designed to provide empathetic interactions have been developed [9]. Despite advances, LLMs often generate repetitive and emotion-insensitive responses, which can negatively affect user satisfaction [10, 11].

This study introduces a new approach to developing empathetic conversational agents using LLMs. We present DBT, inspired by dialectical behavior therapy [12], which outlines four empathy stages: listening, reflecting accurately, acknowledging emotions, and providing genuine responses. Our model integrates these stages to improve empathetic responses.

We created a dataset with responses categorized as "chosen" or "rejected" based on their adherence to these empathy stages and employed Direct Preference Optimization (DPO)[13] to train the model. Our results indicate that this approach significantly enhances general response quality. We also introduce a straightforward metric to evaluate how well responses adhere to DBT.

*Corresponding author

2. Method

2.1 Supervised Fine Tuning (SFT)

We refine the language model using existing datasets of empathetic dialogues to boost its ability to generate empathetic responses. The dialogues are organized into multi-turn interactions, and the model is trained to produce appropriate responses based on the conversation history. The effectiveness of the model is measured by how well its responses fit the context of the dialogue.

2.2 Direct Preference Optimization (DPO)

To further develop the model's capacity for DBT, we employ Direct Preference Optimization (DPO)[13]. We build a human preference dataset by using ChatGPT(GPT-3.5-TURBO) [14] to generate responses based on conversation histories and DBT principles. Responses are classified as "chosen" or "rejected" depending on how well they adhere to the DBT.

3. Results

We evaluate both the general dialogue generation capabilities and the ability to generate empathetic responses. We compare the impact of applying supervised fine-tuning (SFT)

Table 1. Experimental results for each baseline using the ESConv dataset

	B-4	R-L	Dist	Our Metric
Llama-3-8B	0.11	10.78	13.02	0.2753
+ SFT	0.29	13.40	33.40	0.0255
+ DPO	0.07	10.55	12.85	0.2087
+ SFT + DPO	**0.38**	**13.62**	**33.70**	**0.0205**
Mistral-7B	0.05	8.58	7.45	0.2835
+ SFT	0.11	**14.30**	33.40	0.0672
+ DPO	0.06	8.10	6.75	0.2735
+ SFT + DPO	**0.12**	14.25	**33.65**	**0.0508**
Gemma-7B	0.20	2.98	3.05	0.1885
+ SFT	0.21	2.85	2.90	0.2002
+ DPO	0.34	5.75	12.35	0.1304
+ SFT + DPO	**0.40**	**5.85**	**12.80**	**0.0175**

and direct preference optimization (DPO) to assess the effectiveness of our methods. The results, detailed in Table 1, show that despite the limitations of traditional metrics in evaluating empathy, applying SFT and DPO generally enhances performance across all models. Our proposed metric also shows significant improvement in DBT with the use of SFT and DPO, particularly with DPO providing the most substantial gains. This indicates that our approach using the human preference dataset and DBT theory effectively improves response generation.

4. Conclusion

This study introduces an innovative training methodology to enhance LLMs' ability to generate empathetic responses, based on a framework of four stages of DBT. We present a human preference dataset grounded in DBT theory and demonstrate that DPO improves model performance. We also propose a simple yet effective metric for evaluating empathy in LLM-generated responses. Our experiments confirm that integrating DBT significantly enhances empathetic dialogue capabilities in LLMs. This research lays the groundwork for future developments in empathetic dialogue agents and training methods aligned with human preferences.

Acknowledgements

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (RS-2024-00398115, Research on the reliability and coherence of outcomes produced by Generative AI). This work was supported by the Seoul R&BD Program(CY240127). Following are results of a study on the "Leaders in INdustry-university Cooperation 3.0" Project, supported by the Ministry of Education and National Research Foundation of Korea.

Reference

[1] J. Decety, "The neurodevelopment of empathy in humans," *Developmental neuroscience*, Vol. 32, No. 4, pp. 257–267, 2010.

[2] N. Eisenberg and N. D. Eggum, "Empathic responding: Sympathy and personal distress," *The social neuroscience of empathy*, Vol. 6, No. 2009, pp. 71–830, 2009.

[3] W. A. Anthony, "A methodological investigation of the" minimally facilitative level of interpersonal functioning"." *Journal of Clinical Psychology*, Vol. 27, No. 1, 1971.

[4] K. M. Mitchell and B. G. Berenson, "Differential use of confrontation by high and low facilitative therapists," *The Journal of Nervous and Mental Disease*, Vol. 151, No. 5, pp. 303–309, 1970.

[5] G. M. Gazda *et al.*, "Human relations development: A manual for educators," 1977.

[6] D. Richards, "Features and benefits of online counselling: Trinity college online mental health community," *British Journal of Guidance & Counselling*, Vol. 37, No. 3, pp. 231–242, 2009.

[7] D. Kang, S. Kim, T. Kwon, S. Moon, H. Cho, Y. Yu, D. Lee, and J. Yeo, "Can large language models be good emotional supporter? mitigating preference bias on emotional support conversation," *arXiv preprint arXiv:2402.13211*, 2024.

[8] B. R. Burleson, "Emotional support skills," *Handbook of communication and social interaction skills*, pp. 569–612, 2003.

[9] G. Cameron, D. Cameron, G. Megaw, R. Bond, M. Mulvenna, S. O'Neill, C. Armour, and M. McTear, "Towards a chatbot for digital counselling," *Proceedings of the 31st*

international BCS human computer interaction conference (HCI 2017), 2017.

[10] J. Wei, S. Kim, H. Jung, and Y.-H. Kim, "Leveraging large language models to power chatbots for collecting user self-reported data," *Proceedings of the ACM on Human-Computer Interaction*, Vol. 8, No. CSCW1, pp. 1–35, 2024.

[11] S. Concannon, I. Roberts, and M. Tomalin, "An interactional account of empathy in human-machine communication," *Human-Machine Communication*, Vol. 6, 2023.

[12] M. M. Linehan, "Validation and psychotherapy." 1997.

[13] R. Rafailov, A. Sharma, E. Mitchell, C. D. Manning, S. Ermon, and C. Finn, "Direct preference optimization: Your language model is secretly a reward model," *Advances in Neural Information Processing Systems*, Vol. 36, 2024.

[14] OpenAI-Blog, "Chatgpt: Optimizing language models for dialogue," 2022. [Online]. Available: https://openai.com/blog/chatgpt/