

HuanTong Large Language Models by Training Adaptation and Diffusing to Full Parameters

laohur@gmail.com

Abstract

When training a large model, if the cost of training all parameters is high, updating only a small number of parameters can achieve similar results to updating all parameters, with lower cost. This is suitable for small-scale fine-tuning. Based on experience, the volume of data should be ten to a hundred times the size of the model. When fine-tuning with a small number of parameters, only a small number of parameters are updated. After these parameters have been sufficiently trained, continuing to train them does not yield significant benefits. Therefore, the new parameters should be diffused into the old parameters, allowing all parameters of the model to participate in the training. Train using the LoRA method. After the auxiliary parameters have been fully trained, merge them into the base model and continue LoRA Hu et al. (2022) training. This achieves the effect of updating only a small number of parameters in each training iteration while updating all parameters in the overall training process. Training with a small number of parameters to achieve the effect of training with the full set of parameters. And replacing PPO-PTX with DPO Rafailov et al. (2023), suitable for typical simple scenarios. Possibly the easiest full-process RLHF (Reinforcement Learning with Human Feedback) large model process. The application is available at <https://github.com/laohur/HuanTong>.

References

- Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. [LoRA: Low-rank adaptation of large language models](#). In *International Conference on Learning Representations*.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. 2023. [Direct Preference Optimization: Your Language Model is Secretly a Reward Model](#). *arXiv e-prints*, page arXiv:2305.18290.