nanoT5: A PyTorch Framework for Pre-training and Fine-tuning T5-style Models with Limited Resources

Steps to test new Transformer idea:

< ------- Reproduce the baseline
- T5-base pre-training on C4 + T5-instruct fine-tuning on SNI
- We document software + hardware + configs + training curves

< ------ Modify the code with your idea
- Exposed training loop + minimalistic T5 implementation
- Written and optimised in PyTorch

< ------ Test it
- Training starts within minutes and takes 16 hours on 1 x A100
- Supports multi-GPU training, weights from HF Hub, other datasets

"nanoT5 is the ideal template to start your LLM research"
~Happy User

850+ Stars