

Resources

2024-02-20

Publicly available training LLM/VLM logbooks

Logbooks and chronicles of training LLM/VLM are one of the best sources to learn from about dealing with training instabilities and choosing good hyper parameters.

If you know of a public LLM/VLM training logbook that is not on this list please kindly let me know or add it via a PR. Thank you!

The listing is in no particular order other than being grouped by the year.

2021

- BigScience pre-BLOOM 108B training experiments (2021): [chronicles](#) | [the full spec and discussions](#) (backup: [1](#) | [2](#))

2022

- BigScience BLOOM-176B (2022): [chronicles-prequel](#) | [chronicles](#) | [the full spec and discussions](#) (backup: [1](#) | [2](#) | [3](#))
- Meta OPT-175B (2022): [logbook](#) | [Video](#) (backup: [1](#))
- THUDM GLM-130B (2022): [en logbook](#) | [Mandarin version](#) (backup: [1](#) | [2](#))

2023

- HuggingFace IDEFICS-80B multimodal (Flamingo repro) (2023): [Learning log](#) | [Training Chronicles](#) (backup: [1](#) | [2](#))
- BloombergGPT 50B LLM - section C in [BloombergGPT: A Large Language Model for Finance](#)