Performer fine-tuning

Example authors: @TevenLeScao, @Patrickvonplaten

Paper authors: Krzysztof Choromanski, Valerii Likhosherstov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Davis, Afroz Mohiuddin, Lukasz Kaiser, David Belanger, Lucy Colwell, Adrian Weller

Requirements

`datasets` , `flax` and `jax` . `wandb` integration is built-in if you want to use it.

Examples

`sanity_script.sh` will launch performer fine-tuning from the bert-base-cased checkpoint on the Simple Wikipedia dataset (a small, easy-language English Wikipedia) from `datasets` . `full_script.sh` will launch performer fine-tuning from the bert-large-cased checkpoint on the English Wikipedia dataset from `datasets` .

Here are a few key arguments:

- Remove the `--performer` argument to use a standard Bert model.

- Add `--reinitialize` to start from a blank model rather than a Bert checkpoint.

- You may change the Bert size by passing a different checkpoint to the `--model_name_or_path` argument.

- Passing your user name to the `--wandb_user_name` argument will trigger weights and biases logging.

- You can choose a dataset with `--dataset_name` and `--dataset_config` . Our viewer will help you find what you need.