Domain Adapters to Recover Practical Sizes

- Bapna and Firat (2019) propose tiny tunable "adapter" layers.
  1. Pretrain on a large mixed-language corpus.
  2. Inject adapter layers.
  3. Finetune adapter layers for each of the target tasks.