Fine-tuning ElectraSmall on SQuAD for different optimizers and LR schedules

- AdamW with Constant LR
- AdamW with Cosine Restarts LR Schedule
- Adagrad with Warmup and Linear LR Schedule
- SGD with Cosine Restarts LR Schedule (a.k.a. SGDR)