BERT (large) GLUE MNLI on 1 GPU

Validation accuracy vs Time elapsed (s)

- - - 4bs_1gpu_1vn (baseline)
— ✕ — 32bs_1gpu_8vn