

Speedup of Different Parallelization Strategies on LLaMA-2-70B

