

Benchmark Task: arc_challenge

