

Benchmark Task: openbookqa

Pearson correlation: 0.31

