0.40 training validation 0.35 test 0.30

/important-logs/training/short-term-reward-model/39/ALL\_but\_almabench.log

