0.45 training validation 0.40 test 0.35 0.30

/important-logs/training/short-term-reward-model/36/ALL but almabench.log

