training 0.225 validation test

/important-logs/training/short-term-reward-model/29/ALL but almabench.log

