Exact Accuracy (Extended Runs)

- all-200k Evaluation exact accuracy
- hard-1M Evaluation exact accuracy
- all-200k Train exact accuracy
- hard-1M Train exact accuracy