

Mix

Train Reward

8
6
4
2

0 100 200 300 400 500 600

Training Steps

- code
- math
- ifeval

