Continuum Robot PPO report

Model Name	./model/test_judge_1
Time	2025-04-23 11:42
Model_Type	PPO
seed	1
Timesteps	6000000
Control Mode	1
Device	cuda
Network Arch	[1024, 1024, 512]
Average Error	1.2201767950095048
batch	20000
buffer_size	100000
train_freq	4
learning_starts	20000
n_steps	4096
n_epochs	10
learning_rate	0.0003
n_env	24
best_reward	-0.00048039122323137955

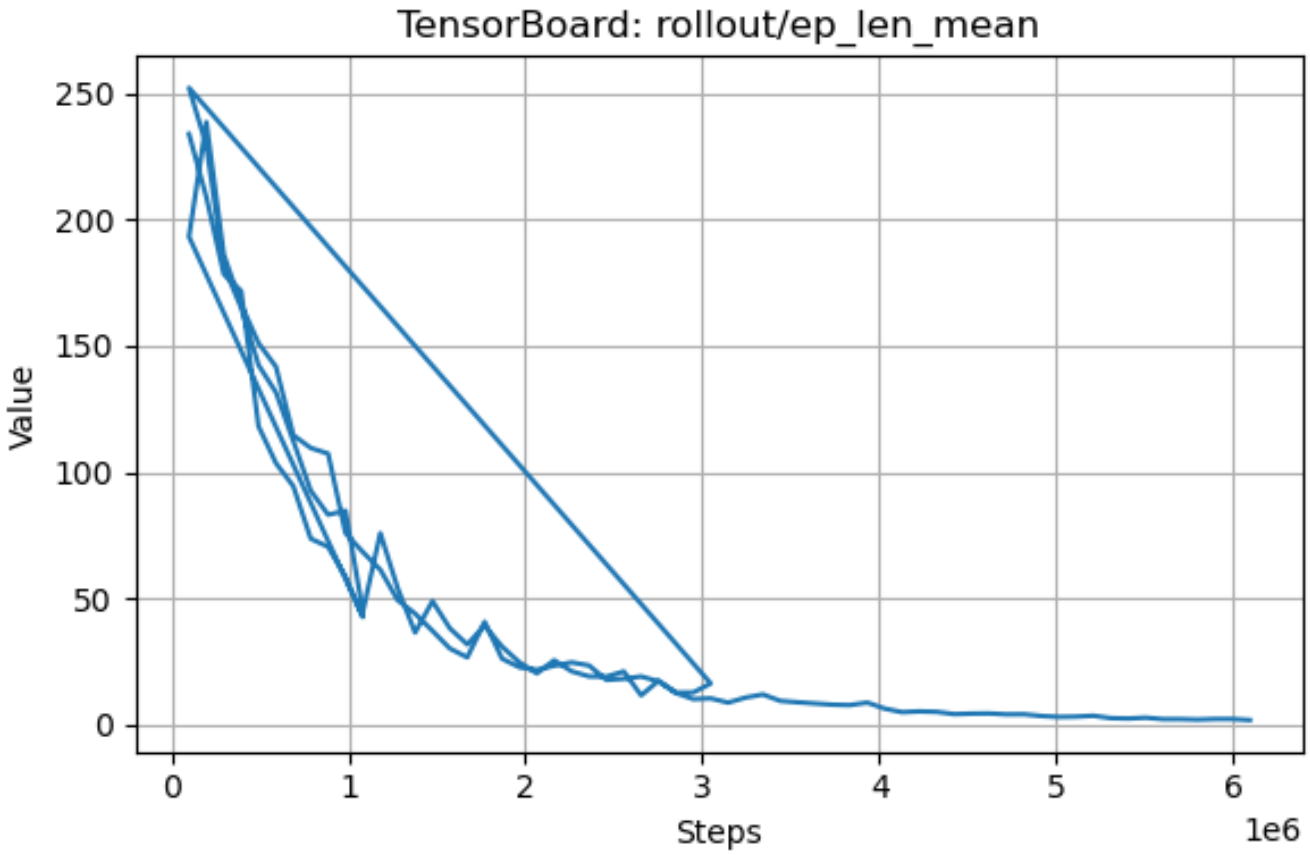
Reward Function Source

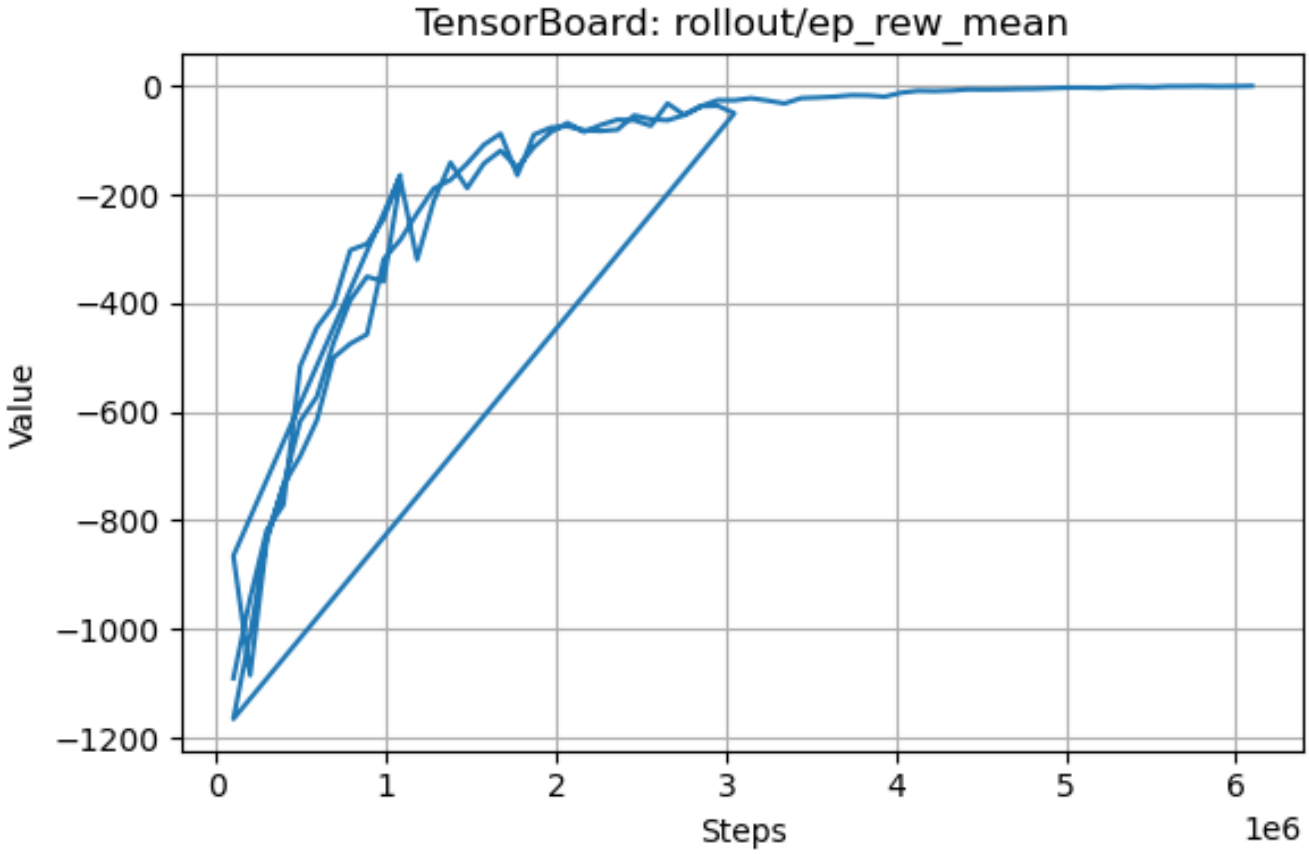
```
def my_custom_reward(distance):
    return -distance + np.exp(-distance**2) * 5
```

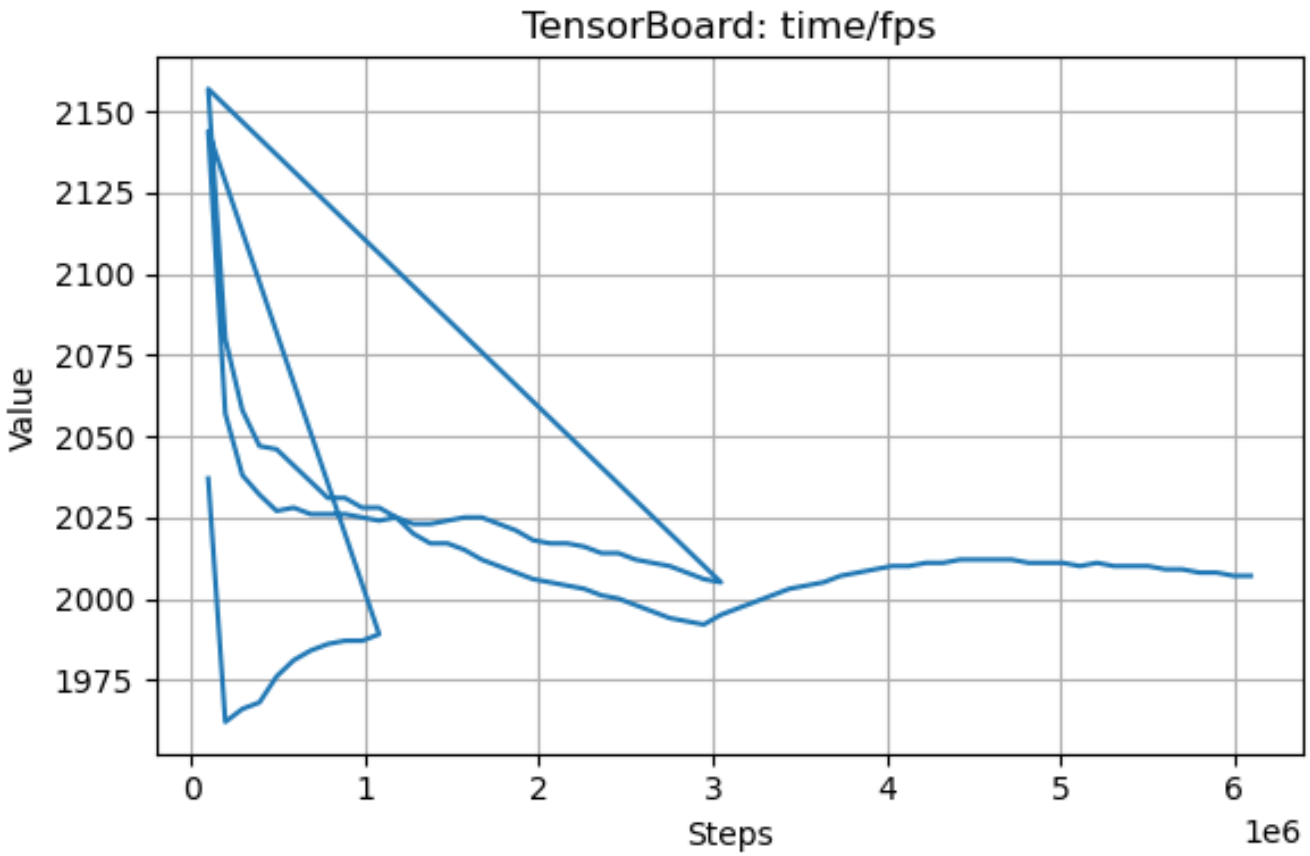
Done Function Source

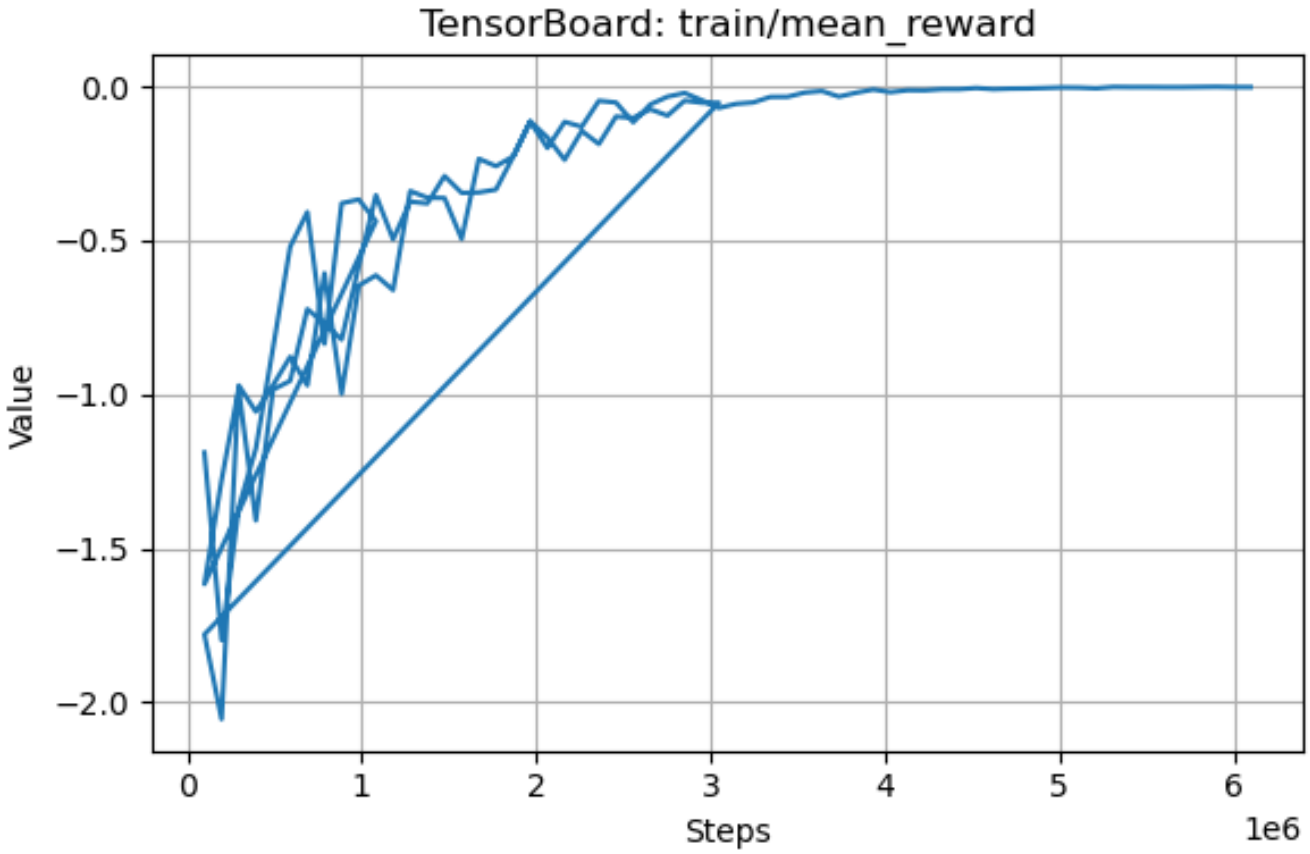
```
def my_custom_done(reward, step,distance,in_step):
    return reward>=-0.5
```

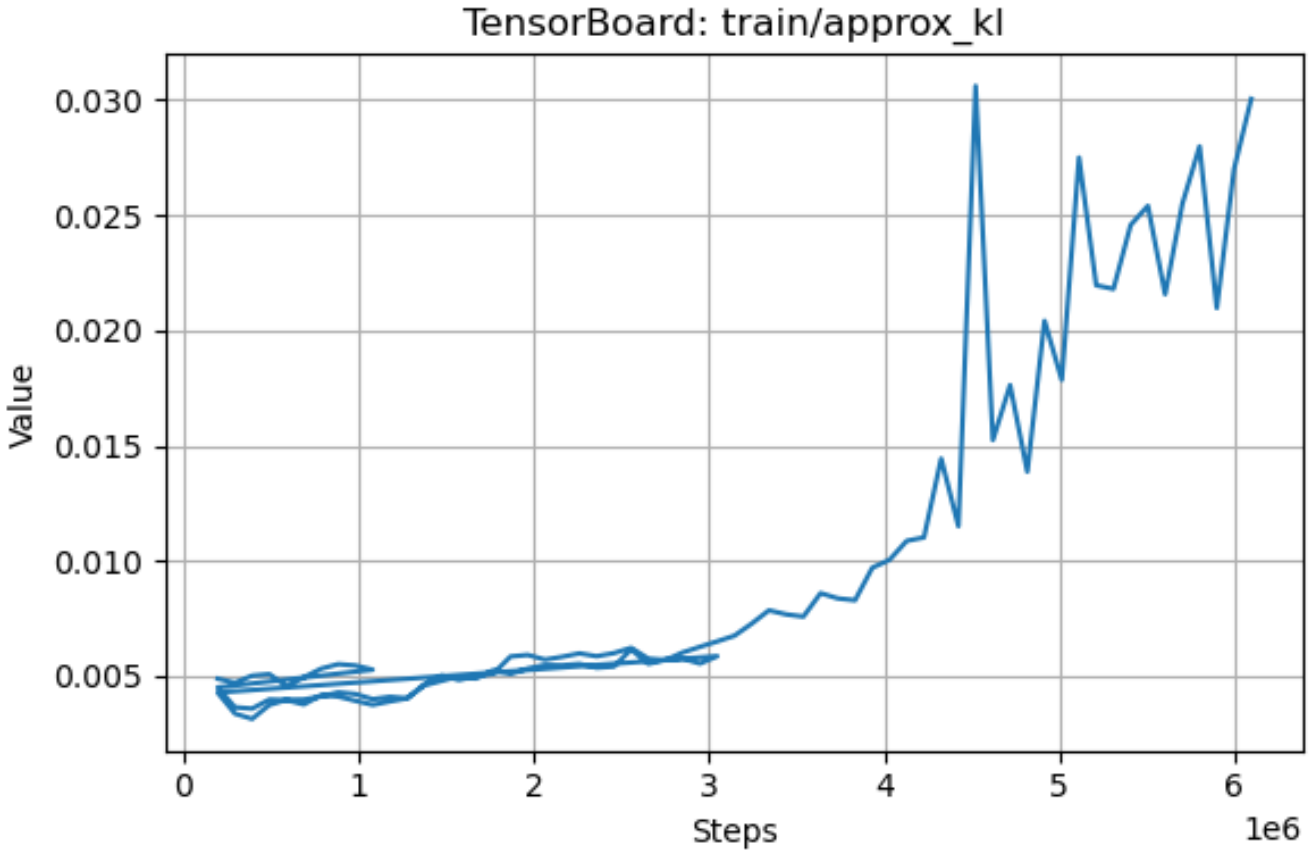
Training Reward Curve

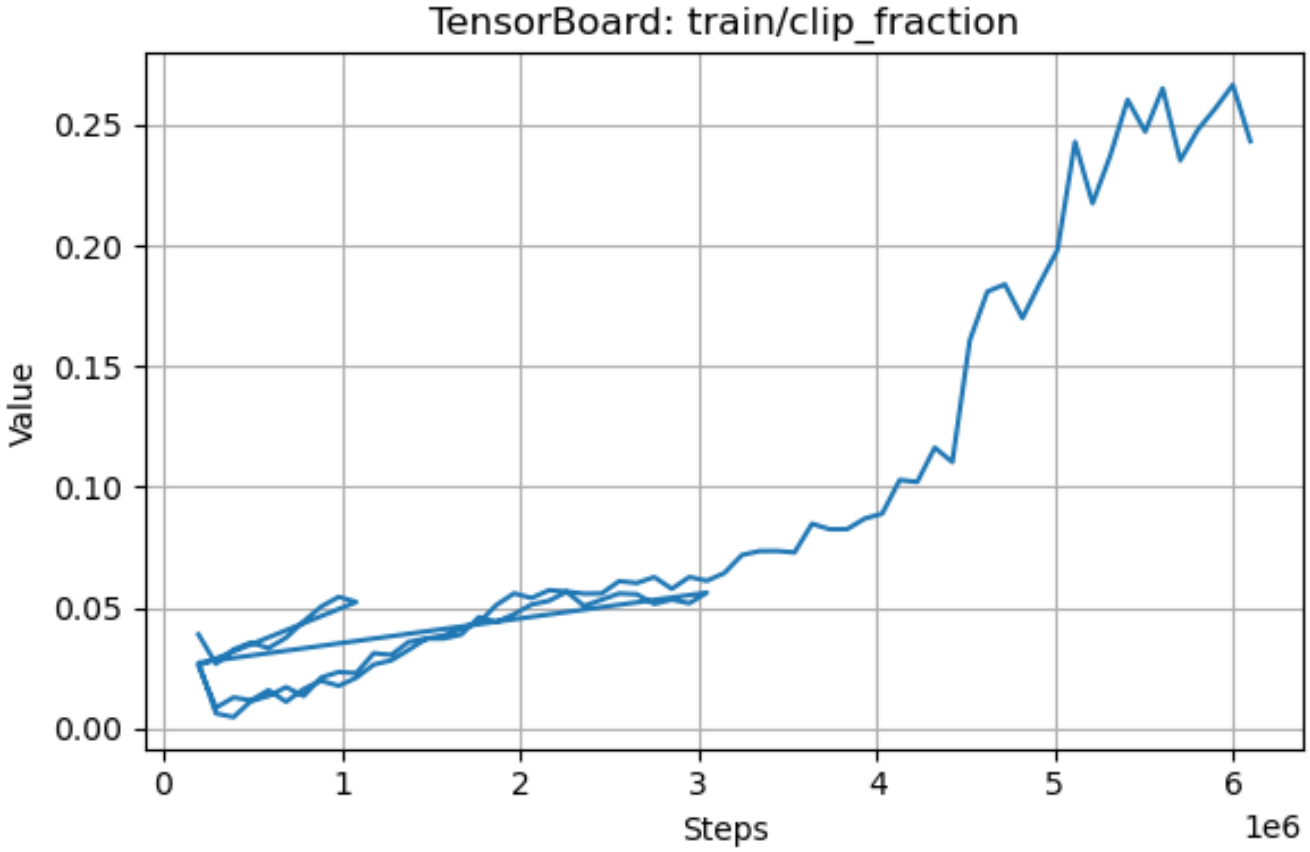


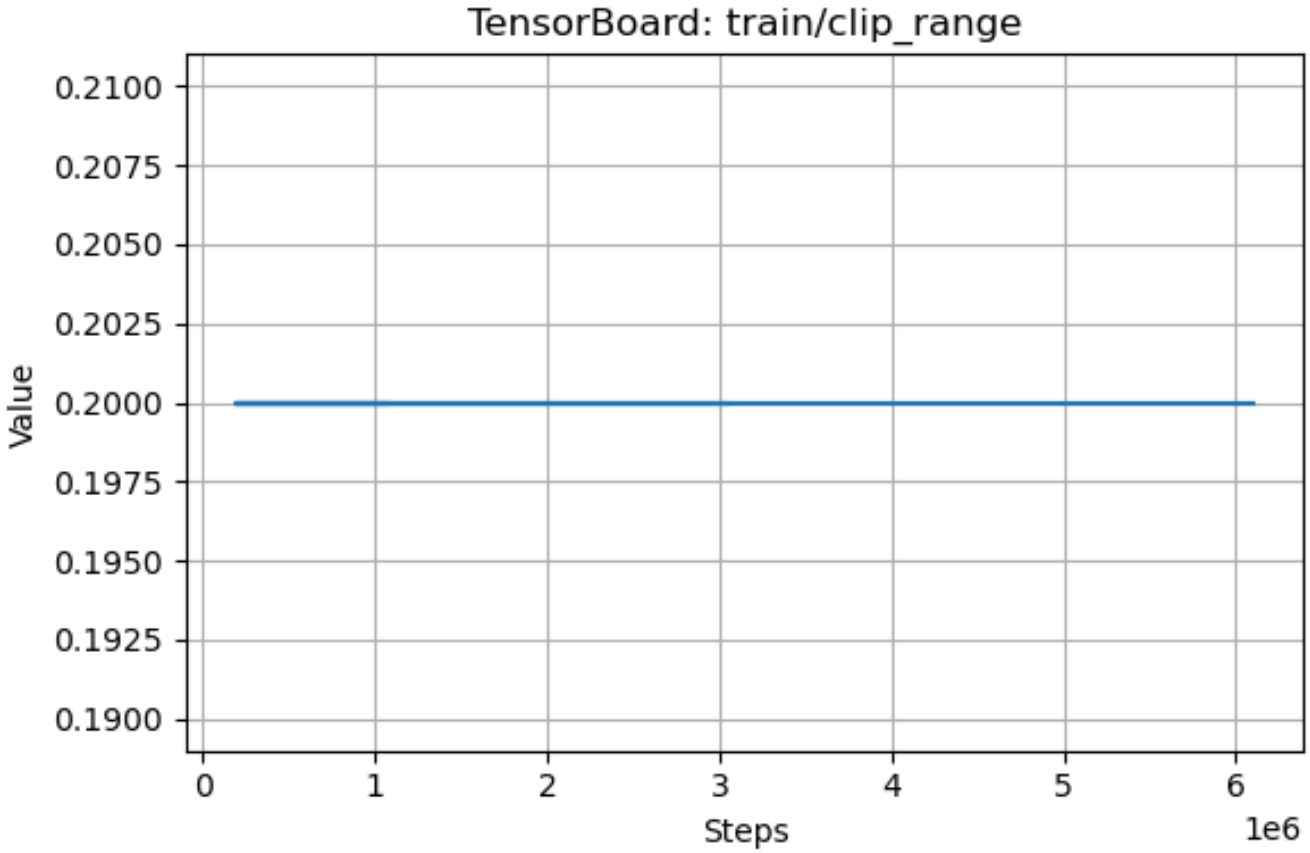


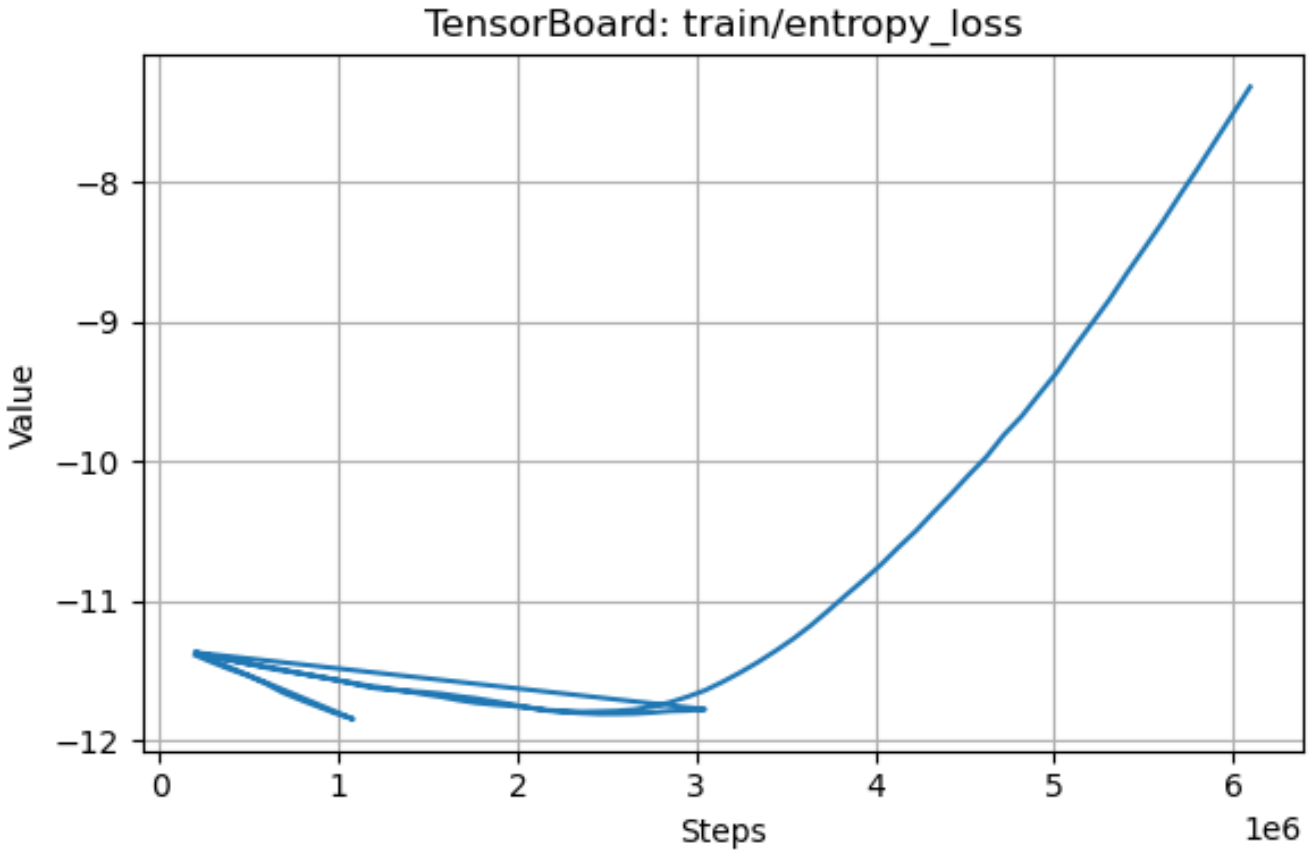

TensorBoard: time/fps
2150
2125
2100
2075
2050
2025
2000
1975
Value
0
1
2
3
4
5
6
Steps
1e6

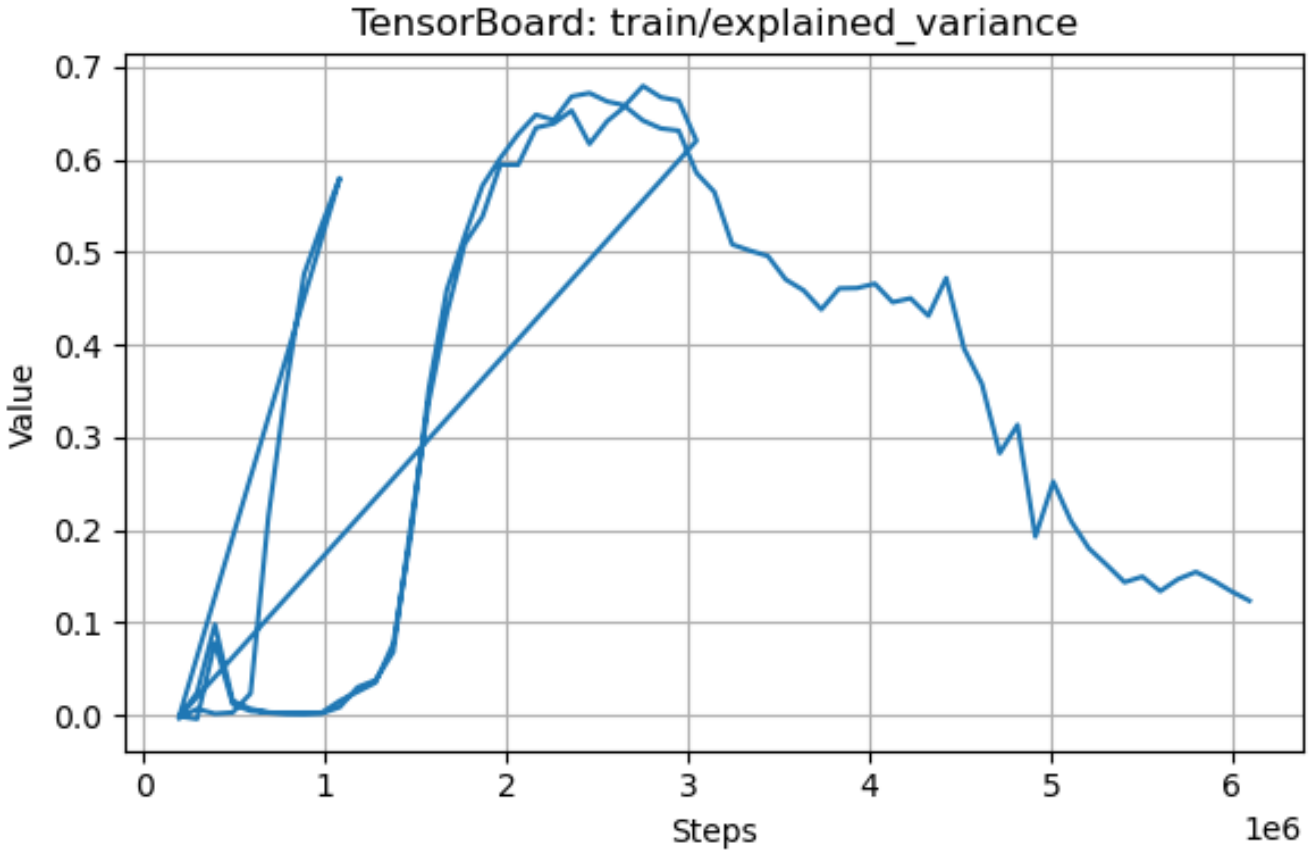

TensorBoard: train/mean_reward
0.0
−0.5
−1.0
−1.5
−2.0
Value
0
1
2
3
4
5
6
Steps
1e6

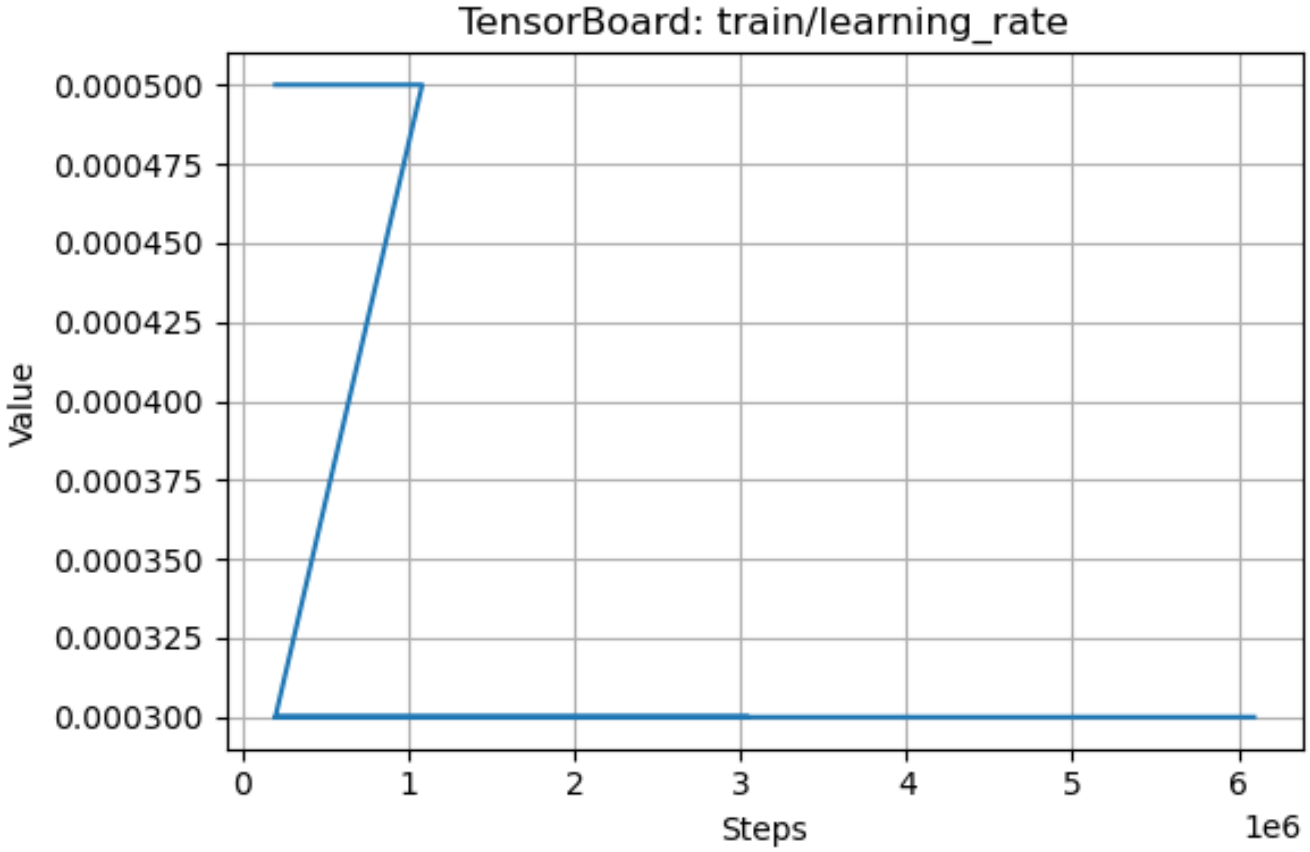

TensorBoard: train/approx_kl
0.030
0.025
0.020
0.015
0.010
0.005
Value
0
1
2
3
4
5
6
Steps
1e6

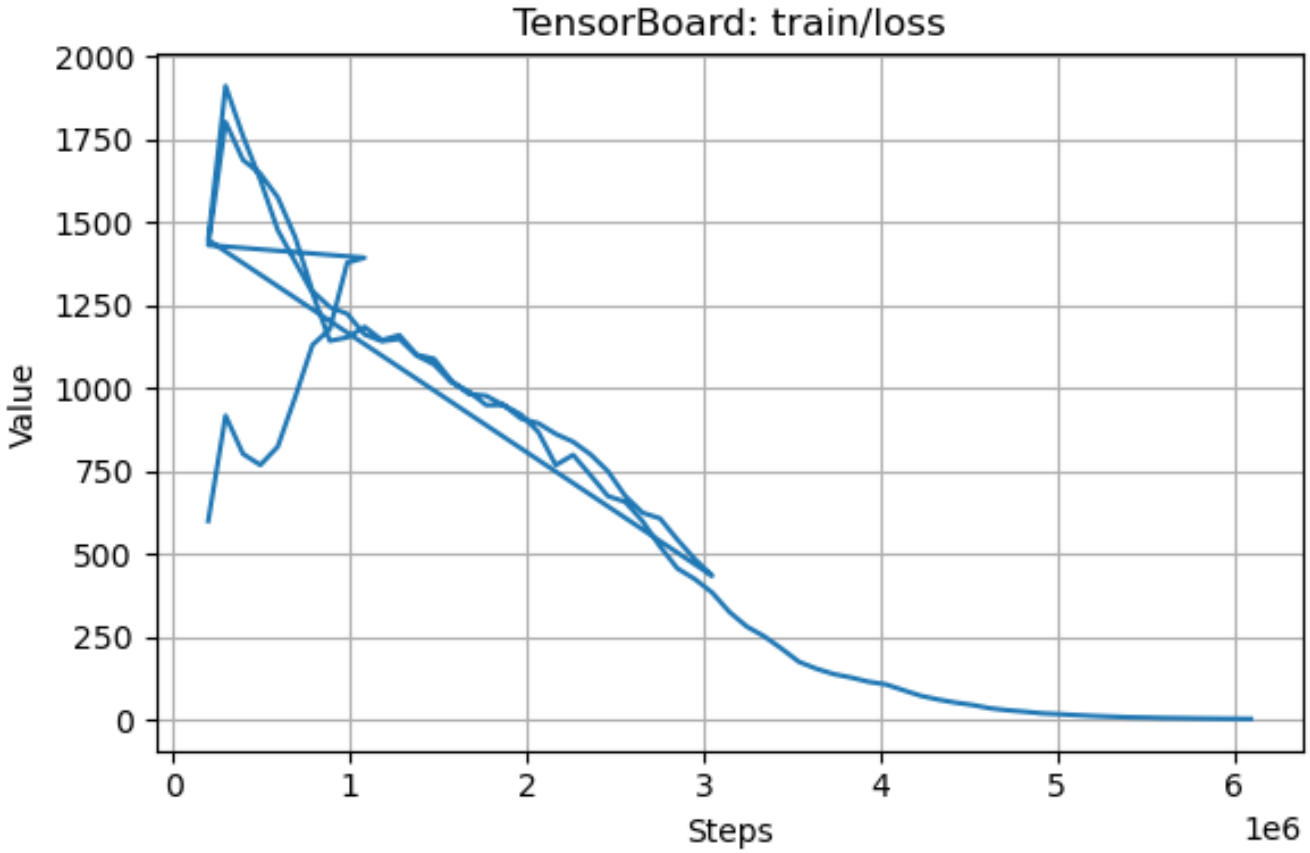

TensorBoard: train/clip_fraction
0.25
0.20
0.15
0.10
0.05
0.00
Value
0
1
2
3
4
5
6
Steps
1e6

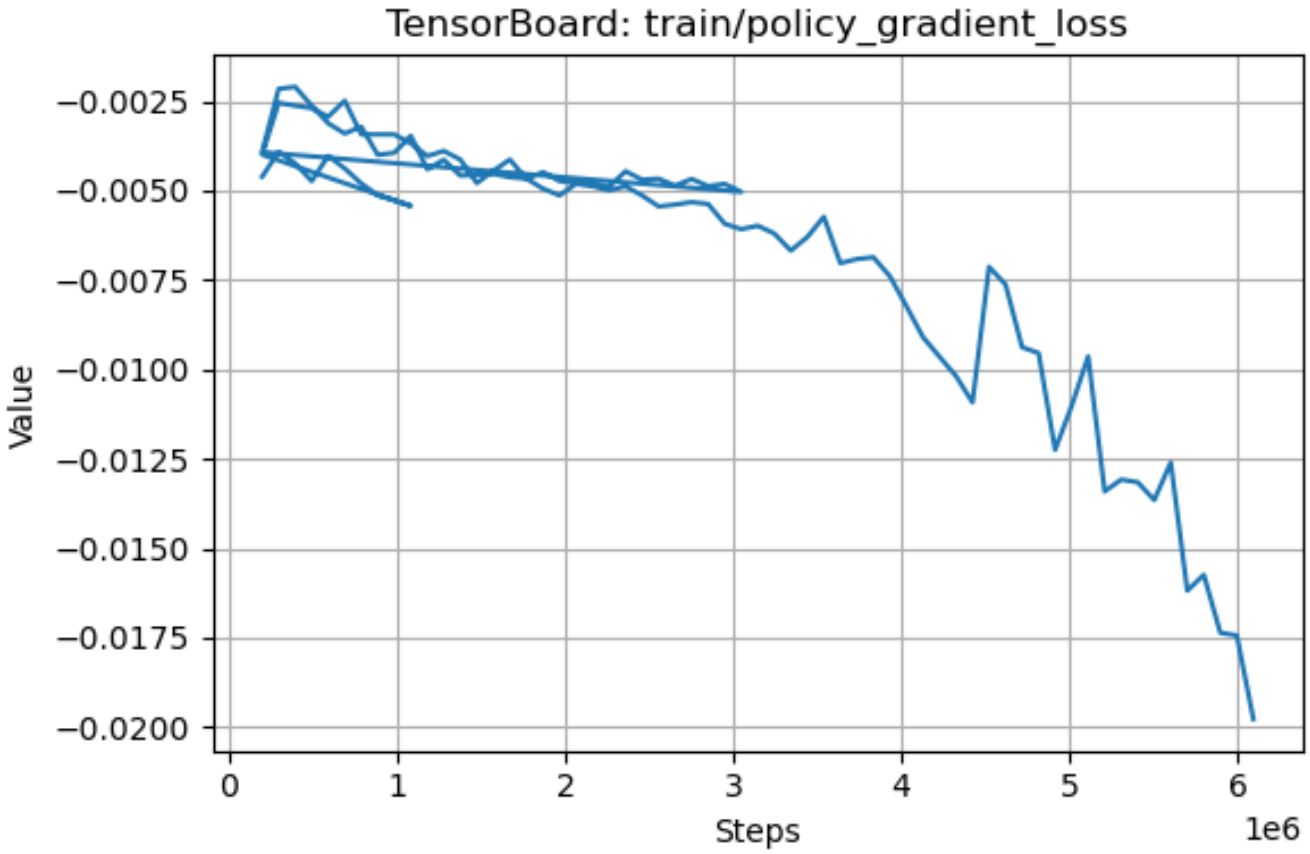

TensorBoard: train/clip_range
0.2100
0.2075
0.2050
0.2025
0.2000
0.1975
0.1950
0.1925
0.1900
Value
0
1
2
3
4
5
6
Steps
1e6

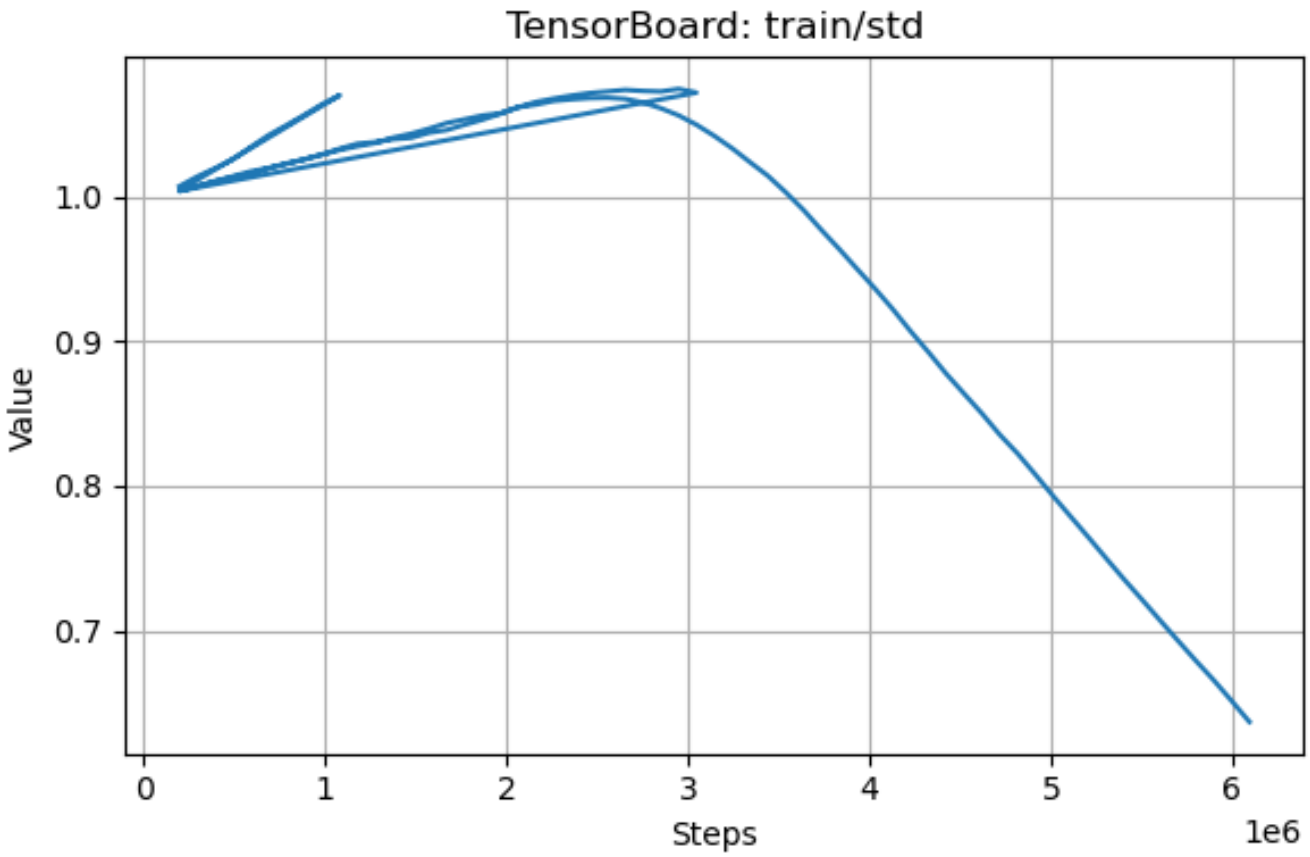

TensorBoard: train/entropy_loss
−8
−9
−10
−11
−12
Value
0
1
2
3
4
5
6
Steps
1e6

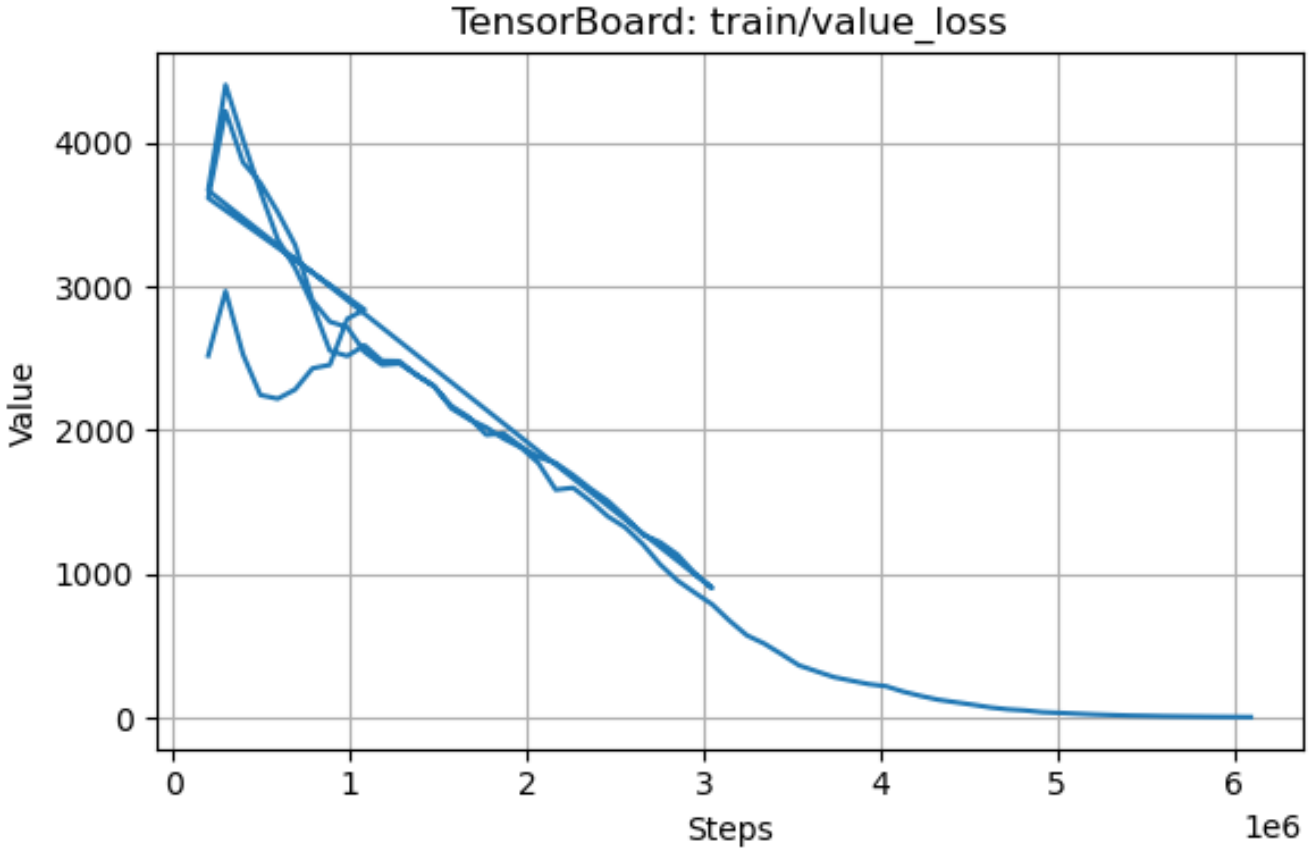

TensorBoard: train/explained_variance
Value
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
6
Steps
1e6


TensorBoard: train/learning_rate
Value
0.000500
0.000475
0.000450
0.000425
0.000400
0.000375
0.000350
0.000325
0.000300
0
1
2
3
4
5
6
Steps
1e6


TensorBoard: train/loss
2000
1750
1500
1250
1000
750
500
250
0
Value
0
1
2
3
4
5
6
Steps
1e6


TensorBoard: train/policy_gradient_loss
−0.0025
−0.0050
−0.0075
−0.0100
−0.0125
−0.0150
−0.0175
−0.0200
Value
0
1
2
3
4
5
6
Steps
1e6


TensorBoard: train/std
1.0
0.9
0.8
0.7
Value
0
1
2
3
4
5
6
Steps
1e6


TensorBoard: train/value_loss
4000
3000
2000
1000
0
Value
0
1
2
3
4
5
6
Steps
1e6