

```

!python kics_real.py

!python kics_surrogate.py

=== Phase 2-2: Training AI Surrogate Model ===
[-] Training MLP...

[-] Validating Scalability...
[-] Scalability Test (Asset: 10B KRW): MAPE = 0.0518%
[SUCCESS] AI Brain is Robust & Scalable!
Figure(800x600)

!python regime.py

!python system.py

=== Dynamic Shield System Test (Real Data) ===
[-] HMM 모델 학습 시작 (Data shape: (5292, 4))...
[-] 모델 학습 완료. 상태 매핑: {2: 'Normal', 0: 'Transition', 1: 'Panic'}
Start Date: 2004-07-26 00:00:00
Initial State: [13.22813      0.          -0.01771999  0.3679      1.5
0.          ]
Step 1: 2004-07-27 | Regime: Normal | K-ICS: 90.9% |
CDS_Proxy(YieldSpread): 0.37
Step 2: 2004-07-28 | Regime: Normal | K-ICS: 97.7% |
CDS_Proxy(YieldSpread): 0.37
Step 3: 2004-07-29 | Regime: Normal | K-ICS: 99.2% |
CDS_Proxy(YieldSpread): 0.37
Step 4: 2004-07-30 | Regime: Normal | K-ICS: 93.5% |
CDS_Proxy(YieldSpread): 0.37
Step 5: 2004-08-02 | Regime: Normal | K-ICS: 90.7% |
CDS_Proxy(YieldSpread): 0.37

!python ppo_trainer.py

=====
Phase 4: RL Training with PPO
=====
Using cpu device

[Model] PPO initialized
[Environment] KICS Gym Environment
[Constraint] K-ICS < 100% -> -1000 penalty

[Training Started]
  Total Timesteps: 50000
  Learning Rate: 0.0003
-----
[실제 데이터 로드] 총 5292 일
  -> 학습용: 3704 일 (70%)
  -> 테스트용: 1588 일 (30%)

```



```

2,347 it/s ]
  4% - ----- 1,918/50,000 [ 0:00:01 < 0:00:21 ,
2,347 it/s ]
  4% - ----- 2,152/50,000 [ 0:00:01 < 0:00:41 ,
1,170 it/s ]
  5% - ----- 2,357/50,000 [ 0:00:01 < 0:00:40 ,
1,214 it/s ]
  6% - ----- 2,837/50,000 [ 0:00:02 < 0:00:36 ,
1,333 it/s ]
  6% - ----- 3,081/50,000 [ 0:00:02 < 0:00:34 ,
1,392 it/s ]
  7% - ----- 3,326/50,000 [ 0:00:02 < 0:00:33 ,
1,445 it/s ]
  7% - ----- 3,567/50,000 [ 0:00:02 < 0:00:32 ,
1,482 it/s ]
  8% - ----- 3,811/50,000 [ 0:00:02 < 0:00:31 ,
1,527 it/s ]

```

time/		
fps		1574
iterations		2
time_elapsed		2
total_timesteps		4096
train/		
approx_kl		0.007488145
clip_fraction		0.0665
clip_range		0.2
entropy_loss		-1.42
explained_variance		-0.0105
learning_rate		0.0003
loss		434
n_updates		10
policy_gradient_loss		-0.00399
std		1
value_loss		1.3e+03

```

  8% - ----- 3,811/50,000 [ 0:00:02 < 0:00:31 ,
1,527 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:02 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:02 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:02 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:02 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:03 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:03 < 0:00:30 ,

```

```

1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:03 < 0:00:30 ,
1,557 it/s ]
  8% - ----- 4,052/50,000 [ 0:00:03 < 0:00:30 ,
1,557 it/s ]
  9% -- ----- 4,289/50,000 [ 0:00:03 < 0:00:38 ,
1,213 it/s ]
[Safety Layer] Triggered 300 times
  9% -- ----- 4,289/50,000 [ 0:00:03 < 0:00:38 ,
1,213 it/s ]
  9% -- ----- 4,497/50,000 [ 0:00:03 < 0:00:37 ,
1,235 it/s ]
  9% -- ----- 4,735/50,000 [ 0:00:03 < 0:00:36 ,
1,270 it/s ]
Step 5000 | Episodes: 10 | Avg Reward (last 10): 1263.00
  9% -- ----- 4,735/50,000 [ 0:00:03 < 0:00:36 ,
1,270 it/s ]
 10% -- ----- 4,935/50,000 [ 0:00:03 < 0:00:36 ,
1,281 it/s ]
 10% -- ----- 5,178/50,000 [ 0:00:03 < 0:00:35 ,
1,313 it/s ]
[Safety Layer] Triggered 400 times
 10% -- ----- 5,178/50,000 [ 0:00:03 < 0:00:35 ,
1,313 it/s ]
 11% -- ----- 5,415/50,000 [ 0:00:04 < 0:00:34 ,
1,337 it/s ]
[Safety Layer] Triggered 500 times
 11% -- ----- 5,415/50,000 [ 0:00:04 < 0:00:34 ,
1,337 it/s ]
[Safety Layer] Triggered 600 times
 11% -- ----- 5,415/50,000 [ 0:00:04 < 0:00:34 ,
1,337 it/s ]
 11% -- ----- 5,654/50,000 [ 0:00:04 < 0:00:33 ,
1,366 it/s ]
 12% -- ----- 5,893/50,000 [ 0:00:04 < 0:00:32 ,
1,393 it/s ]

```

time/		
fps	1425	
iterations	3	
time_elapsed	4	
total_timesteps	6144	
train/		
approx_kl	0.00071351195	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.42	
explained_variance	-0.0188	
learning_rate	0.0003	

loss	459
n_updates	20
policy_gradient_loss	-0.000278
std	1.01
value_loss	1.19e+03

```

-----
12% -- ----- 5,893/50,000 [ 0:00:04 < 0:00:32 ,
1,393 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:04 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:05 < 0:00:31 ,
1,415 it/s ]
12% -- ----- 6,136/50,000 [ 0:00:05 < 0:00:31 ,
1,415 it/s ]
13% --- ----- 6,368/50,000 [ 0:00:05 < 0:00:37 ,
1,212 it/s ]
[Safety Layer] Triggered 700 times
13% --- ----- 6,368/50,000 [ 0:00:05 < 0:00:37 ,
1,212 it/s ]
13% --- ----- 6,582/50,000 [ 0:00:05 < 0:00:36 ,
1,228 it/s ]
[Safety Layer] Triggered 800 times
13% --- ----- 6,582/50,000 [ 0:00:05 < 0:00:36 ,
1,228 it/s ]
14% --- ----- 6,818/50,000 [ 0:00:05 < 0:00:35 ,
1,251 it/s ]
[Safety Layer] Triggered 900 times
14% --- ----- 6,818/50,000 [ 0:00:05 < 0:00:35 ,
1,251 it/s ]
14% --- ----- 7,052/50,000 [ 0:00:05 < 0:00:34 ,
1,272 it/s ]
15% --- ----- 7,291/50,000 [ 0:00:05 < 0:00:34 ,
1,291 it/s ]
[Safety Layer] Triggered 1000 times
15% --- ----- 7,291/50,000 [ 0:00:05 < 0:00:34 ,
1,291 it/s ]
16% --- ----- 7,778/50,000 [ 0:00:05 < 0:00:32 ,
1,330 it/s ]
[Safety Layer] Triggered 1100 times

```

```
16% --- ----- 7,778/50,000 [ 0:00:05 < 0:00:32 ,
1,330 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:05 < 0:00:32 ,
1,349 it/s ]
```

time/		
fps		1369
iterations		4
time_elapsed		5
total_timesteps		8192
train/		
approx_kl		0.0048981216
clip_fraction		0.0281
clip_range		0.2
entropy_loss		-1.43
explained_variance		0.00775
learning_rate		0.0003
loss		554
n_updates		30
policy_gradient_loss		-0.00401
std		1.01
value_loss		1.34e+03

```
16% --- ----- 8,011/50,000 [ 0:00:05 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,011/50,000 [ 0:00:06 < 0:00:32 ,
1,349 it/s ]
16% --- ----- 8,242/50,000 [ 0:00:06 < 0:00:35 ,
1,206 it/s ]
17% ---- ----- 8,466/50,000 [ 0:00:06 < 0:00:34 ,
1,223 it/s ]
[Safety Layer] Triggered 1200 times
17% ---- ----- 8,466/50,000 [ 0:00:06 < 0:00:34 ,
1,223 it/s ]
[Safety Layer] Triggered 1300 times
17% ---- ----- 8,466/50,000 [ 0:00:06 < 0:00:34 ,
1,223 it/s ]
```

```

17% ---- ----- 8,696/50,000 [ 0:00:06 < 0:00:34 ,
1,237 it/s ]
18% ---- ----- 8,937/50,000 [ 0:00:07 < 0:00:33 ,
1,255 it/s ]
18% ---- ----- 9,180/50,000 [ 0:00:07 < 0:00:33 ,
1,270 it/s ]
19% ---- ----- 9,424/50,000 [ 0:00:07 < 0:00:32 ,
1,287 it/s ]
19% ---- ----- 9,668/50,000 [ 0:00:07 < 0:00:31 ,
1,301 it/s ]
Step 10000 | Episodes: 20 | Avg Reward (last 10): 1332.31
19% ---- ----- 9,668/50,000 [ 0:00:07 < 0:00:31 ,
1,301 it/s ]
20% ---- ----- 9,911/50,000 [ 0:00:07 < 0:00:31 ,
1,318 it/s ]

```

time/		
fps	1343	
iterations	5	
time_elapsed	7	
total_timesteps	10240	
train/		
approx_kl	0.003292694	
clip_fraction	0.0082	
clip_range	0.2	
entropy_loss	-1.43	
explained_variance	0.0192	
learning_rate	0.0003	
loss	538	
n_updates	40	
policy_gradient_loss	-0.00109	
std	1	
value_loss	1.33e+03	

```

20% ---- ----- 9,911/50,000 [ 0:00:07 < 0:00:31 ,
1,318 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:07 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:07 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:07 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:07 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:08 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:08 < 0:00:30 ,
1,330 it/s ]
20% ---- ----- 10,146/50,000 [ 0:00:08 < 0:00:30 ,

```

```

1,330 it/s ]
 20% ----- 10,146/50,000 [ 0:00:08 < 0:00:30 ,
1,330 it/s ]
 21% ----- 10,379/50,000 [ 0:00:08 < 0:00:33 ,
1,211 it/s ]
 21% ----- 10,588/50,000 [ 0:00:08 < 0:00:33 ,
1,223 it/s ]
 22% ----- 10,819/50,000 [ 0:00:08 < 0:00:32 ,
1,234 it/s ]
 22% ----- 11,063/50,000 [ 0:00:08 < 0:00:32 ,
1,249 it/s ]
 23% ----- 11,306/50,000 [ 0:00:08 < 0:00:31 ,
1,261 it/s ]
 23% ----- 11,549/50,000 [ 0:00:09 < 0:00:31 ,
1,275 it/s ]
[Safety Layer] Triggered 1400 times
 23% ----- 11,549/50,000 [ 0:00:09 < 0:00:31 ,
1,275 it/s ]
 24% ----- 11,792/50,000 [ 0:00:09 < 0:00:30 ,
1,287 it/s ]
 24% ----- 12,028/50,000 [ 0:00:09 < 0:00:30 ,
1,299 it/s ]

```

	time/	
	fps	1318
	iterations	6
	time_elapsed	9
	total_timesteps	12288
	train/	
	approx_kl	0.005123001
	clip_fraction	0.027
	clip_range	0.2
	entropy_loss	-1.41
	explained_variance	0.00431
	learning_rate	0.0003
	loss	473
	n_updates	50
	policy_gradient_loss	-0.00351
	std	0.992
	value_loss	1.15e+03

```

 24% ----- 12,028/50,000 [ 0:00:09 < 0:00:30 ,
1,299 it/s ]
 25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]
 25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]
 25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]

```



```

25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]
25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]
25% ----- 12,273/50,000 [ 0:00:09 < 0:00:29 ,
1,313 it/s ]
25% ----- 12,273/50,000 [ 0:00:10 < 0:00:29 ,
1,313 it/s ]
25% ----- 12,273/50,000 [ 0:00:10 < 0:00:29 ,
1,313 it/s ]
25% ----- 12,509/50,000 [ 0:00:10 < 0:00:31 ,
1,218 it/s ]
25% ----- 12,731/50,000 [ 0:00:10 < 0:00:31 ,
1,229 it/s ]
26% ----- 12,976/50,000 [ 0:00:10 < 0:00:30 ,
1,239 it/s ]
26% ----- 13,222/50,000 [ 0:00:10 < 0:00:30 ,
1,252 it/s ]
27% ----- 13,469/50,000 [ 0:00:10 < 0:00:29 ,
1,264 it/s ]
28% ----- 13,949/50,000 [ 0:00:10 < 0:00:29 ,
1,283 it/s ]
28% ----- 14,179/50,000 [ 0:00:10 < 0:00:28 ,
1,294 it/s ]

```

```

-----
| time/          |
|   fps          | 1305
|   iterations    | 7
|   time_elapsed  | 10
|   total_timesteps | 14336
| train/         |
|   approx_kl     | 0.006751947
|   clip_fraction | 0.0447
|   clip_range    | 0.2
|   entropy_loss  | -1.4
|   explained_variance | 0.0024
|   learning_rate | 0.0003
|   loss          | 409
|   n_updates     | 60
|   policy_gradient_loss | -0.00387
|   std           | 0.978
|   value_loss    | 1.02e+03
|
-----

```

```

28% ----- 14,179/50,000 [ 0:00:10 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]

```

```

28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
28% ----- 14,179/50,000 [ 0:00:11 < 0:00:28 ,
1,294 it/s ]
29% ----- 14,407/50,000 [ 0:00:11 < 0:00:30 ,
1,216 it/s ]
29% ----- 14,620/50,000 [ 0:00:11 < 0:00:29 ,
1,223 it/s ]
30% ----- 14,852/50,000 [ 0:00:12 < 0:00:29 ,
1,233 it/s ]
Step 15000 | Episodes: 30 | Avg Reward (last 10): 1283.66
30% ----- 14,852/50,000 [ 0:00:12 < 0:00:29 ,
1,233 it/s ]
30% ----- 15,089/50,000 [ 0:00:12 < 0:00:29 ,
1,243 it/s ]
31% ----- 15,332/50,000 [ 0:00:12 < 0:00:28 ,
1,252 it/s ]
[Safety Layer] Triggered 1500 times
31% ----- 15,332/50,000 [ 0:00:12 < 0:00:28 ,
1,252 it/s ]
31% ----- 15,576/50,000 [ 0:00:12 < 0:00:28 ,
1,262 it/s ]
[Safety Layer] Triggered 1600 times
31% ----- 15,576/50,000 [ 0:00:12 < 0:00:28 ,
1,262 it/s ]
32% ----- 15,814/50,000 [ 0:00:12 < 0:00:27 ,
1,270 it/s ]
32% ----- 16,058/50,000 [ 0:00:12 < 0:00:27 ,
1,280 it/s ]

```

time/		
fps	1296	
iterations	8	
time_elapsed	12	
total_timesteps	16384	
train/		
approx_kl	0.0018568442	
clip_fraction	0.00889	
clip_range	0.2	
entropy_loss	-1.41	
explained_variance	0.000133	
learning_rate	0.0003	
loss	397	

n_updates	70
policy_gradient_loss	-0.00146
std	1.01
value_loss	997

```

32% ----- 16,058/50,000 [ 0:00:12 < 0:00:27 ,
1,280 it/s ]
33% ----- 16,300/50,000 [ 0:00:12 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:12 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:12 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:12 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:13 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:13 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:13 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:13 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,300/50,000 [ 0:00:13 < 0:00:27 ,
1,289 it/s ]
33% ----- 16,536/50,000 [ 0:00:13 < 0:00:28 ,
1,223 it/s ]
33% ----- 16,749/50,000 [ 0:00:13 < 0:00:28 ,
1,231 it/s ]
[Safety Layer] Triggered 1700 times
33% ----- 16,749/50,000 [ 0:00:13 < 0:00:28 ,
1,231 it/s ]
34% ----- 16,982/50,000 [ 0:00:13 < 0:00:27 ,
1,238 it/s ]
[Safety Layer] Triggered 1800 times
34% ----- 16,982/50,000 [ 0:00:13 < 0:00:27 ,
1,238 it/s ]
35% ----- 17,475/50,000 [ 0:00:13 < 0:00:26 ,
1,255 it/s ]
35% ----- 17,720/50,000 [ 0:00:13 < 0:00:26 ,
1,265 it/s ]
36% ----- 17,963/50,000 [ 0:00:14 < 0:00:26 ,
1,272 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]

```

time/	
fps	1292
iterations	9
time_elapsed	14
total_timesteps	18432
train/	

approx_kl	0.0005184588	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.43	
explained_variance	0.0017	
learning_rate	0.0003	
loss	389	
n_updates	80	
policy_gradient_loss	5.05e-05	
std	1.01	
value_loss	953	

```

-----
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
36% ----- 18,206/50,000 [ 0:00:14 < 0:00:25 ,
1,281 it/s ]
37% ----- 18,441/50,000 [ 0:00:15 < 0:00:26 ,
1,222 it/s ]
37% ----- 18,651/50,000 [ 0:00:15 < 0:00:26 ,
1,228 it/s ]
38% ----- 18,869/50,000 [ 0:00:15 < 0:00:26 ,
1,233 it/s ]
38% ----- 19,108/50,000 [ 0:00:15 < 0:00:25 ,
1,242 it/s ]
[Safety Layer] Triggered 1900 times
38% ----- 19,108/50,000 [ 0:00:15 < 0:00:25 ,
1,242 it/s ]
39% ----- 19,350/50,000 [ 0:00:15 < 0:00:25 ,
1,250 it/s ]
[Safety Layer] Triggered 2000 times
39% ----- 19,350/50,000 [ 0:00:15 < 0:00:25 ,
1,250 it/s ]
39% ----- 19,585/50,000 [ 0:00:15 < 0:00:25 ,
1,256 it/s ]
Step 20000 | Episodes: 40 | Avg Reward (last 10): 1307.73
39% ----- 19,585/50,000 [ 0:00:15 < 0:00:25 ,
1,256 it/s ]

```

```
40% ----- 19,824/50,000 [ 0:00:15 < 0:00:24 ,
1,264 it/s ]
41% ----- 20,313/50,000 [ 0:00:15 < 0:00:24 ,
1,279 it/s ]
```

	time/	
	fps	1286
	iterations	10
	time_elapsed	15
	total_timesteps	20480
	train/	
	approx_kl	0.0012702432
	clip_fraction	0.00186
	clip_range	0.2
	entropy_loss	-1.42
	explained_variance	0.00123
	learning_rate	0.0003
	loss	356
	n_updates	90
	policy_gradient_loss	-0.00052
	std	0.998
	value_loss	902

```
41% ----- 20,313/50,000 [ 0:00:15 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:15 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
41% ----- 20,313/50,000 [ 0:00:16 < 0:00:24 ,
1,279 it/s ]
42% ----- 20,779/50,000 [ 0:00:16 < 0:00:24 ,
1,231 it/s ]
42% ----- 21,008/50,000 [ 0:00:16 < 0:00:24 ,
1,238 it/s ]
43% ----- 21,251/50,000 [ 0:00:17 < 0:00:24 ,
1,244 it/s ]
43% ----- 21,495/50,000 [ 0:00:17 < 0:00:23 ,
1,252 it/s ]
```

```
43% ----- 21,739/50,000 [ 0:00:17 < 0:00:23 ,
1,258 it/s ]
44% ----- 21,984/50,000 [ 0:00:17 < 0:00:23 ,
1,266 it/s ]
44% ----- 22,225/50,000 [ 0:00:17 < 0:00:22 ,
1,273 it/s ]
```

```
-----
| time/          |          |
|   fps          |   1283   |
| iterations     |   11     |
| time_elapsed   |   17     |
| total_timesteps| 22528    |
| train/         |          |
|   approx_kl    | 0.0014941485 |
|   clip_fraction| 0.000293   |
|   clip_range   | 0.2       |
|   entropy_loss | -1.42     |
|   explained_variance| 0.000979 |
|   learning_rate| 0.0003    |
|   loss         | 337       |
|   n_updates    | 100       |
|   policy_gradient_loss| -0.000215 |
|   std          | 0.994     |
|   value_loss   | 831       |
|               |          |
|-----|-----|
```

```
44% ----- 22,225/50,000 [ 0:00:17 < 0:00:22 ,
1,273 it/s ]
45% ----- 22,470/50,000 [ 0:00:17 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:17 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:17 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:17 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:18 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:18 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:18 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,470/50,000 [ 0:00:18 < 0:00:22 ,
1,279 it/s ]
45% ----- 22,708/50,000 [ 0:00:18 < 0:00:23 ,
1,229 it/s ]
[Safety Layer] Triggered 2100 times
45% ----- 22,708/50,000 [ 0:00:18 < 0:00:23 ,
1,229 it/s ]
46% ----- 22,923/50,000 [ 0:00:18 < 0:00:22 ,
```

```
1,235 it/s ]
46% ----- 23,161/50,000 [ 0:00:18 < 0:00:22 ,
1,240 it/s ]
47% ----- 23,410/50,000 [ 0:00:18 < 0:00:22 ,
1,248 it/s ]
48% ----- 23,909/50,000 [ 0:00:18 < 0:00:21 ,
1,261 it/s ]
48% ----- 24,152/50,000 [ 0:00:19 < 0:00:21 ,
1,267 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
```

```
-----
| time/                |
|   fps                | 1280
|   iterations         | 12
|   time_elapsed       | 19
|   total_timesteps    | 24576
| train/               |
|   approx_kl          | 0.004909802
|   clip_fraction      | 0.0379
|   clip_range         | 0.2
|   entropy_loss       | -1.41
|   explained_variance | -0.000165
|   learning_rate      | 0.0003
|   loss               | 297
|   n_updates          | 110
|   policy_gradient_loss | -0.00271
|   std                | 0.982
|   value_loss         | 725
|-----|
```

```
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,396/50,000 [ 0:00:19 < 0:00:21 ,
1,273 it/s ]
49% ----- 24,634/50,000 [ 0:00:20 < 0:00:21 ,
1,229 it/s ]
[Safety Layer] Triggered 2200 times
```

```

49% ----- 24,634/50,000 [ 0:00:20 < 0:00:21 ,
1,229 it/s ]
50% ----- 24,849/50,000 [ 0:00:20 < 0:00:21 ,
1,233 it/s ]
[Safety Layer] Triggered 2300 times
50% ----- 24,849/50,000 [ 0:00:20 < 0:00:21 ,
1,233 it/s ]
Step 25000 | Episodes: 50 | Avg Reward (last 10): 1290.94
50% ----- 24,849/50,000 [ 0:00:20 < 0:00:21 ,
1,233 it/s ]
50% ----- 25,061/50,000 [ 0:00:20 < 0:00:21 ,
1,238 it/s ]
51% ----- 25,307/50,000 [ 0:00:20 < 0:00:20 ,
1,243 it/s ]
51% ----- 25,549/50,000 [ 0:00:20 < 0:00:20 ,
1,249 it/s ]
52% ----- 25,792/50,000 [ 0:00:20 < 0:00:20 ,
1,254 it/s ]
52% ----- 26,037/50,000 [ 0:00:20 < 0:00:20 ,
1,261 it/s ]
53% ----- 26,284/50,000 [ 0:00:20 < 0:00:19 ,
1,267 it/s ]

```

	time/	
	fps	1277
	iterations	13
	time_elapsed	20
	total_timesteps	26624
	train/	
	approx_kl	0.0013366137
	clip_fraction	0.00161
	clip_range	0.2
	entropy_loss	-1.41
	explained_variance	0.000449
	learning_rate	0.0003
	loss	274
	n_updates	120
	policy_gradient_loss	-0.000308
	std	0.993
	value_loss	691

```

53% ----- 26,284/50,000 [ 0:00:20 < 0:00:19 ,
1,267 it/s ]
53% ----- 26,528/50,000 [ 0:00:20 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:20 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]

```



```

53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]
53% ----- 26,528/50,000 [ 0:00:21 < 0:00:19 ,
1,272 it/s ]
54% ----- 26,766/50,000 [ 0:00:21 < 0:00:19 ,
1,231 it/s ]
54% ----- 26,978/50,000 [ 0:00:21 < 0:00:19 ,
1,235 it/s ]
54% ----- 27,208/50,000 [ 0:00:21 < 0:00:19 ,
1,239 it/s ]
55% ----- 27,692/50,000 [ 0:00:22 < 0:00:18 ,
1,250 it/s ]
56% ----- 27,912/50,000 [ 0:00:22 < 0:00:18 ,
1,255 it/s ]
56% ----- 28,154/50,000 [ 0:00:22 < 0:00:18 ,
1,260 it/s ]
57% ----- 28,399/50,000 [ 0:00:22 < 0:00:18 ,
1,265 it/s ]

```

```

-----
| time/          |          |
|   fps          |    1273  |
| iterations     |     14   |
| time_elapsed   |     22   |
| total_timesteps|    28672 |
| train/         |          |
|   approx_kl    | 0.0034606461 |
|   clip_fraction| 0.0157    |
|   clip_range   | 0.2       |
|   entropy_loss  | -1.41     |
|   explained_variance| 0.000442 |
|   learning_rate | 0.0003    |
|   loss         | 240       |
|   n_updates    | 130       |
|   policy_gradient_loss| -0.0019  |
|   std          | 0.983     |
|   value_loss   | 656       |
|               |          |
|-----|-----|

```

```

57% ----- 28,399/50,000 [ 0:00:22 < 0:00:18 ,
1,265 it/s ]
57% ----- 28,639/50,000 [ 0:00:22 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:22 < 0:00:17 ,
1,270 it/s ]

```

```

57% ----- 28,639/50,000 [ 0:00:22 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:22 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:22 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:23 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:23 < 0:00:17 ,
1,270 it/s ]
57% ----- 28,639/50,000 [ 0:00:23 < 0:00:17 ,
1,270 it/s ]
58% ----- 28,873/50,000 [ 0:00:23 < 0:00:18 ,
1,232 it/s ]
[Safety Layer] Triggered 2400 times
58% ----- 28,873/50,000 [ 0:00:23 < 0:00:18 ,
1,232 it/s ]
58% ----- 29,095/50,000 [ 0:00:23 < 0:00:17 ,
1,236 it/s ]
59% ----- 29,324/50,000 [ 0:00:23 < 0:00:17 ,
1,241 it/s ]
59% ----- 29,557/50,000 [ 0:00:23 < 0:00:17 ,
1,245 it/s ]
[Safety Layer] Triggered 2500 times
59% ----- 29,557/50,000 [ 0:00:23 < 0:00:17 ,
1,245 it/s ]
60% ----- 29,792/50,000 [ 0:00:23 < 0:00:17 ,
1,250 it/s ]
Step 30000 | Episodes: 60 | Avg Reward (last 10): 1284.62
60% ----- 29,792/50,000 [ 0:00:23 < 0:00:17 ,
1,250 it/s ]
60% ----- 30,030/50,000 [ 0:00:23 < 0:00:16 ,
1,255 it/s ]
61% ----- 30,273/50,000 [ 0:00:24 < 0:00:16 ,
1,260 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]

```

time/	
fps	1270
iterations	15
time_elapsed	24
total_timesteps	30720
train/	
approx_kl	0.00087540725
clip_fraction	0.000195
clip_range	0.2
entropy_loss	-1.41
explained_variance	6.81e-05

learning_rate	0.0003
loss	221
n_updates	140
policy_gradient_loss	-4.64e-05
std	0.994
value_loss	571

```

61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,502/50,000 [ 0:00:24 < 0:00:16 ,
1,264 it/s ]
61% ----- 30,728/50,000 [ 0:00:25 < 0:00:16 ,
1,226 it/s ]
62% ----- 30,931/50,000 [ 0:00:25 < 0:00:16 ,
1,229 it/s ]
62% ----- 31,141/50,000 [ 0:00:25 < 0:00:16 ,
1,233 it/s ]
63% ----- 31,380/50,000 [ 0:00:25 < 0:00:16 ,
1,237 it/s ]
[Safety Layer] Triggered 2600 times
63% ----- 31,380/50,000 [ 0:00:25 < 0:00:16 ,
1,237 it/s ]
63% ----- 31,621/50,000 [ 0:00:25 < 0:00:15 ,
1,242 it/s ]
64% ----- 31,859/50,000 [ 0:00:25 < 0:00:15 ,
1,246 it/s ]
64% ----- 32,098/50,000 [ 0:00:25 < 0:00:15 ,
1,251 it/s ]
65% ----- 32,584/50,000 [ 0:00:25 < 0:00:14 ,
1,260 it/s ]

```

time/	
fps	1265
iterations	16
time_elapsed	25
total_timesteps	32768
train/	

approx_kl	0.0024464335
clip_fraction	0.00747
clip_range	0.2
entropy_loss	-1.4
explained_variance	0.000323
learning_rate	0.0003
loss	208
n_updates	150
policy_gradient_loss	-0.000632
std	0.968
value_loss	569

65%	-----	-----	32,584/50,000	[0:00:25 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:25 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
65%	-----	-----	32,584/50,000	[0:00:26 < 0:00:14 ,
1,260	it/s]			
66%	-----	-----	32,819/50,000	[0:00:26 < 0:00:15 ,
1,225	it/s]			
[Safety Layer] Triggered 2700 times				
66%	-----	-----	32,819/50,000	[0:00:26 < 0:00:15 ,
1,225	it/s]			
66%	-----	-----	33,031/50,000	[0:00:26 < 0:00:14 ,
1,228	it/s]			
66%	-----	-----	33,247/50,000	[0:00:26 < 0:00:14 ,
1,232	it/s]			
67%	-----	-----	33,486/50,000	[0:00:27 < 0:00:14 ,
1,236	it/s]			
67%	-----	-----	33,729/50,000	[0:00:27 < 0:00:14 ,
1,241	it/s]			
68%	-----	-----	33,971/50,000	[0:00:27 < 0:00:13 ,
1,245	it/s]			
69%	-----	-----	34,447/50,000	[0:00:27 < 0:00:13 ,
1,253	it/s]			
69%	-----	-----	34,682/50,000	[0:00:27 < 0:00:13 ,
1,257	it/s]			

time/	
fps	1261
iterations	17
time_elapsed	27
total_timesteps	34816
train/	
approx_kl	0.0033603034
clip_fraction	0.0146
clip_range	0.2
entropy_loss	-1.39
explained_variance	0.000307
learning_rate	0.0003
loss	185
n_updates	160
policy_gradient_loss	-0.00209
std	0.971
value_loss	516

```

69% ----- 34,682/50,000 [ 0:00:27 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:27 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:27 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:27 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:27 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:28 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:28 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:28 < 0:00:13 ,
1,257 it/s ]
69% ----- 34,682/50,000 [ 0:00:28 < 0:00:13 ,
1,257 it/s ]
70% ----- 34,914/50,000 [ 0:00:28 < 0:00:13 ,
1,222 it/s ]
Step 35000 | Episodes: 70 | Avg Reward (last 10): 1295.39
70% ----- 34,914/50,000 [ 0:00:28 < 0:00:13 ,
1,222 it/s ]
70% ----- 35,112/50,000 [ 0:00:28 < 0:00:13 ,
1,225 it/s ]
71% ----- 35,332/50,000 [ 0:00:28 < 0:00:12 ,
1,229 it/s ]
71% ----- 35,560/50,000 [ 0:00:28 < 0:00:12 ,
1,232 it/s ]
72% ----- 35,797/50,000 [ 0:00:28 < 0:00:12 ,

```



```

75% ----- 37,615/50,000 [ 0:00:30 < 0:00:11 ,
1,208 it/s ]
76% ----- 37,858/50,000 [ 0:00:30 < 0:00:11 ,
1,209 it/s ]
76% ----- 38,101/50,000 [ 0:00:30 < 0:00:10 ,
1,209 it/s ]
77% ----- 38,341/50,000 [ 0:00:30 < 0:00:10 ,
1,243 it/s ]
77% ----- 38,581/50,000 [ 0:00:31 < 0:00:10 ,
1,247 it/s ]

```

```

-----
| time/          |          |
|   fps          |   1250   |
|   iterations   |   19     |
|   time_elapsed |   31     |
|   total_timesteps | 38912   |
| train/         |          |
|   approx_kl    | 0.0063631805 |
|   clip_fraction | 0.0397    |
|   clip_range   | 0.2       |
|   entropy_loss  | -1.38     |
|   explained_variance | 1.55e-06 |
|   learning_rate | 0.0003    |
|   loss         | 149       |
|   n_updates    | 180       |
|   policy_gradient_loss | -0.00333 |
|   std          | 0.953     |
|   value_loss   | 406       |
|               |          |
|-----|-----|

```

```

77% ----- 38,581/50,000 [ 0:00:31 < 0:00:10 ,
1,247 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 38,824/50,000 [ 0:00:31 < 0:00:09 ,
1,251 it/s ]
78% ----- 39,058/50,000 [ 0:00:32 < 0:00:10 ,
1,211 it/s ]

```

```

79% ----- 39,262/50,000 [ 0:00:32 < 0:00:09 ,
1,209 it/s ]
79% ----- 39,482/50,000 [ 0:00:32 < 0:00:09 ,
1,208 it/s ]
[Safety Layer] Triggered 2800 times
79% ----- 39,482/50,000 [ 0:00:32 < 0:00:09 ,
1,208 it/s ]
79% ----- 39,723/50,000 [ 0:00:32 < 0:00:09 ,
1,208 it/s ]
Step 40000 | Episodes: 80 | Avg Reward (last 10): 1280.49
79% ----- 39,723/50,000 [ 0:00:32 < 0:00:09 ,
1,208 it/s ]
80% ----- 39,961/50,000 [ 0:00:32 < 0:00:09 ,
1,207 it/s ]
80% ----- 40,197/50,000 [ 0:00:32 < 0:00:09 ,
1,207 it/s ]
81% ----- 40,428/50,000 [ 0:00:32 < 0:00:08 ,
1,240 it/s ]
81% ----- 40,671/50,000 [ 0:00:32 < 0:00:08 ,
1,244 it/s ]

```

time/		
fps	1247	
iterations	20	
time_elapsed	32	
total_timesteps	40960	
train/		
approx_kl	0.0037767785	
clip_fraction	0.0272	
clip_range	0.2	
entropy_loss	-1.38	
explained_variance	-0.000113	
learning_rate	0.0003	
loss	122	
n_updates	190	
policy_gradient_loss	-0.00247	
std	0.968	
value_loss	352	

```

81% ----- 40,671/50,000 [ 0:00:32 < 0:00:08 ,
1,244 it/s ]
82% ----- 40,914/50,000 [ 0:00:32 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]

```



```

82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 40,914/50,000 [ 0:00:33 < 0:00:08 ,
1,248 it/s ]
82% ----- 41,148/50,000 [ 0:00:33 < 0:00:08 ,
1,210 it/s ]
83% ----- 41,359/50,000 [ 0:00:33 < 0:00:08 ,
1,209 it/s ]
83% ----- 41,601/50,000 [ 0:00:34 < 0:00:07 ,
1,209 it/s ]
84% ----- 42,086/50,000 [ 0:00:34 < 0:00:07 ,
1,209 it/s ]
85% ----- 42,325/50,000 [ 0:00:34 < 0:00:07 ,
1,209 it/s ]
85% ----- 42,564/50,000 [ 0:00:34 < 0:00:06 ,
1,243 it/s ]
86% ----- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]

```

```

-----
| time/          |
|   fps          | 1245
| iterations     | 21
| time_elapsed   | 34
| total_timesteps | 43008
| train/         |
|   approx_kl    | 0.00077711034
|   clip_fraction | 0.00249
|   clip_range    | 0.2
|   entropy_loss  | -1.38
|   explained_variance | 0.000162
|   learning_rate | 0.0003
|   loss          | 115
|   n_updates     | 200
|   policy_gradient_loss | 0.000127
|   std           | 0.966
|   value_loss    | 341
|
-----

```

```

86% ----- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]
86% ----- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]
86% ----- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]
86% ----- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]

```

```

86% ----- --- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]
86% ----- --- 42,795/50,000 [ 0:00:34 < 0:00:06 ,
1,246 it/s ]
86% ----- --- 42,795/50,000 [ 0:00:35 < 0:00:06 ,
1,246 it/s ]
86% ----- --- 42,795/50,000 [ 0:00:35 < 0:00:06 ,
1,246 it/s ]
86% ----- --- 42,795/50,000 [ 0:00:35 < 0:00:06 ,
1,246 it/s ]
86% ----- --- 43,024/50,000 [ 0:00:35 < 0:00:06 ,
1,209 it/s ]
[Safety Layer] Triggered 2900 times
86% ----- --- 43,024/50,000 [ 0:00:35 < 0:00:06 ,
1,209 it/s ]
[Safety Layer] Triggered 3000 times
86% ----- --- 43,024/50,000 [ 0:00:35 < 0:00:06 ,
1,209 it/s ]
87% ----- --- 43,438/50,000 [ 0:00:35 < 0:00:06 ,
1,207 it/s ]
87% ----- -- 43,678/50,000 [ 0:00:35 < 0:00:06 ,
1,207 it/s ]
88% ----- -- 43,919/50,000 [ 0:00:35 < 0:00:06 ,
1,207 it/s ]
88% ----- -- 44,154/50,000 [ 0:00:35 < 0:00:05 ,
1,207 it/s ]
89% ----- -- 44,397/50,000 [ 0:00:35 < 0:00:05 ,
1,239 it/s ]
89% ----- -- 44,640/50,000 [ 0:00:36 < 0:00:05 ,
1,243 it/s ]
Step 45000 | Episodes: 90 | Avg Reward (last 10): 1302.36
89% ----- -- 44,640/50,000 [ 0:00:36 < 0:00:05 ,
1,243 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
[Safety Layer] Triggered 3100 times
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]

```

time/		
fps	1243	
iterations	22	
time_elapsed	36	
total_timesteps	45056	
train/		
approx_kl	0.004043601	
clip_fraction	0.0203	
clip_range	0.2	
entropy_loss	-1.38	

explained_variance	8.48e-05
learning_rate	0.0003
loss	108
n_updates	210
policy_gradient_loss	-0.00241
std	0.959
value_loss	303

```

90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 44,879/50,000 [ 0:00:36 < 0:00:05 ,
1,247 it/s ]
90% ----- -- 45,112/50,000 [ 0:00:37 < 0:00:05 ,
1,206 it/s ]
91% ----- -- 45,337/50,000 [ 0:00:37 < 0:00:04 ,
1,206 it/s ]
91% ----- -- 45,556/50,000 [ 0:00:37 < 0:00:04 ,
1,205 it/s ]
92% ----- - 45,785/50,000 [ 0:00:37 < 0:00:04 ,
1,205 it/s ]
92% ----- - 46,029/50,000 [ 0:00:37 < 0:00:04 ,
1,205 it/s ]
93% ----- - 46,275/50,000 [ 0:00:37 < 0:00:04 ,
1,206 it/s ]
93% ----- - 46,520/50,000 [ 0:00:37 < 0:00:03 ,
1,240 it/s ]
94% ----- - 46,761/50,000 [ 0:00:37 < 0:00:03 ,
1,244 it/s ]

```

time/	
fps	1242
iterations	23
time_elapsed	37
total_timesteps	47104
train/	
approx_kl	0.0015778954
clip_fraction	0.0019

clip_range	0.2
entropy_loss	-1.37
explained_variance	9.38e-05
learning_rate	0.0003
loss	94.1
n_updates	220
policy_gradient_loss	-4.08e-05
std	0.949
value_loss	294

```

94% ----- - 46,761/50,000 [ 0:00:37 < 0:00:03 ,
1,244 it/s ]
94% ----- - 47,006/50,000 [ 0:00:37 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,006/50,000 [ 0:00:38 < 0:00:03 ,
1,248 it/s ]
94% ----- - 47,244/50,000 [ 0:00:38 < 0:00:03 ,
1,207 it/s ]
95% ----- - 47,453/50,000 [ 0:00:38 < 0:00:03 ,
1,207 it/s ]
95% ----- - 47,678/50,000 [ 0:00:39 < 0:00:02 ,
1,206 it/s ]
96% ----- 48,169/50,000 [ 0:00:39 < 0:00:02 ,
1,207 it/s ]
97% ----- 48,412/50,000 [ 0:00:39 < 0:00:02 ,
1,206 it/s ]
97% ----- 48,652/50,000 [ 0:00:39 < 0:00:02 ,
1,240 it/s ]
[Safety Layer] Triggered 3200 times
97% ----- 48,652/50,000 [ 0:00:39 < 0:00:02 ,
1,240 it/s ]
98% ----- 48,892/50,000 [ 0:00:39 < 0:00:01 ,
1,243 it/s ]

```

time/	
fps	1241
iterations	24

time_elapsed	39
total_timesteps	49152
train/	
approx_kl	0.001333639
clip_fraction	0.00835
clip_range	0.2
entropy_loss	-1.38
explained_variance	0.000116
learning_rate	0.0003
loss	74.4
n_updates	230
policy_gradient_loss	-0.00119
std	0.976
value_loss	239

```

98% ----- 48,892/50,000 [ 0:00:39 < 0:00:01 ,
1,243 it/s ]
98% ----- 49,131/50,000 [ 0:00:39 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:39 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:39 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:39 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
[Safety Layer] Triggered 3300 times
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
[Safety Layer] Triggered 3400 times
98% ----- 49,131/50,000 [ 0:00:40 < 0:00:01 ,
1,247 it/s ]
99% ----- 49,366/50,000 [ 0:00:40 < 0:00:01 ,
1,206 it/s ]
99% ----- 49,567/50,000 [ 0:00:40 < 0:00:01 ,
1,205 it/s ]
100% ----- 49,806/50,000 [ 0:00:40 < 0:00:01 ,
1,204 it/s ]
Step 50000 | Episodes: 100 | Avg Reward (last 10): 1301.36
100% ----- 49,806/50,000 [ 0:00:40 < 0:00:01 ,
1,204 it/s ]
100% ----- 50,042/50,000 [ 0:00:40 < 0:00:00 ,

```

```

1,205 it/s ]
100% ----- 50,284/50,000 [ 0:00:40 < 0:00:00 ,
1,205 it/s ]
100% ----- 50,530/50,000 [ 0:00:40 < 0:00:00 ,
1,238 it/s ]
100% ----- 50,774/50,000 [ 0:00:40 < 0:00:00 ,
1,241 it/s ]
-----
| time/                                     |
|   fps                                   | 1240 |
|   iterations                           | 25   |
|   time_elapsed                         | 41   |
|   total_timesteps                     | 51200 |
| train/                                |
|   approx_kl                           | 0.0072538164 |
|   clip_fraction                       | 0.0552 |
|   clip_range                          | 0.2    |
|   entropy_loss                        | -1.38  |
|   explained_variance                  | 7.39e-05 |
|   learning_rate                      | 0.0003 |
|   loss                               | 59.9   |
|   n_updates                          | 240    |
|   policy_gradient_loss                | -0.00405 |
|   std                               | 0.955  |
|   value_loss                         | 205    |
|-----|-----|
100% ----- 50,774/50,000 [ 0:00:40 < 0:00:00 ,
1,241 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,200/50,000 [ 0:00:40 < 0:00:00
[Safety Layer Summary] Total triggers: 3456
100% ----- 51,014/50,000 [ 0:00:40 < 0:00:00 ,
1,246 it/s ]
100% ----- 51,200/50,000 [ 0:00:40 < 0:00:00

```

1,206 it/s]

[Training Complete]

=====
Evaluation
=====

Episodes: 10

Avg Reward: 1301.14

Avg Min K-ICS: 999.0%

[SUCCESS] Agent maintained K-ICS > 100% on average!

Figure(1400x500)

[Saved] ppo_training_result.png

[Saved] Model saved to c:\Users\PC\Desktop\Quant\한화\src\validation\
ppo_kics.zip

!python proof_risk_paradox.py

=====
Phase 5.1: Risk Paradox Proof
=====

[Correlation: -0.6]

Optimal Hedge Ratio: 0.0%

SCR at Optimal: 0.1190

SCR at 100% Hedge: 0.1429

Capital Savings: 16.67%

Paradox Proven: YES ✓

[Correlation: -0.4]

Optimal Hedge Ratio: 0.0%

SCR at Optimal: 0.1042

SCR at 100% Hedge: 0.1250

Capital Savings: 16.67%

Paradox Proven: YES ✓

[Correlation: -0.2]

Optimal Hedge Ratio: 0.0%

SCR at Optimal: 0.0926

SCR at 100% Hedge: 0.1111

Capital Savings: 16.67%

Paradox Proven: YES ✓

[Correlation: 0.0]

Optimal Hedge Ratio: 0.0%

SCR at Optimal: 0.0833

SCR at 100% Hedge: 0.1000

Capital Savings: 16.67%

Paradox Proven: YES ✓

```
[Correlation: 0.2]
  Optimal Hedge Ratio: 0.0%
  SCR at Optimal: 0.0758
  SCR at 100% Hedge: 0.0909
  Capital Savings: 16.67%
  Paradox Proven: YES ✓
```

```
=====
[SUCCESS] Risk Paradox Proven!
  5/5 scenarios show the paradox
Figure(1200x600)
```

```
[Saved] risk_paradox_proof.png

!python solvency_visualizer.py
```

```
=====
Phase 5.2: Solvency Analysis (COVID-19 Scenario)
=====
```

```
[100% Hedge]
  Min K-ICS: 1159.8%
  Final K-ICS: 1594.6%
```

```
[80% Fixed]
  Min K-ICS: 979.7%
  Final K-ICS: 1375.4%
```

```
[Dynamic Shield]
  Min K-ICS: 1248.7%
  Final K-ICS: 1779.1%
```

```
[SUCCESS] Dynamic Shield maintained K-ICS > 100% during crisis!
Figure(1400x1200)
```

```
[Saved] kics_defense_result.png

!python stress_safety.py
```

```
=====
Phase 5.3: Safety Layer Stress Test
=====
```

```
[Test 1] VIX > 40 Injection Test
```

```
-----
Step 1 | VIX: 15 | Hedge: 0.60 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 2 | VIX: 20 | Hedge: 0.70 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 3 | VIX: 25 | Hedge: 0.80 | Action: 3 | CRITICAL: K-ICS < 100%,
```



```

FORCE HEDGE 100%
Step 4 | VIX: 30 | Hedge: 0.90 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 5 | VIX: 35 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 6 | VIX: 40 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 7 | VIX: 45 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 8 | VIX: 50 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 9 | VIX: 55 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 10 | VIX: 50 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 11 | VIX: 45 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 12 | VIX: 40 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
>>> Emergency: Gradual De-risking Triggered <<<
Step 13 | VIX: 35 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 14 | VIX: 30 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 15 | VIX: 25 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%
Step 16 | VIX: 20 | Hedge: 1.00 | Action: 3 | CRITICAL: K-ICS < 100%,
FORCE HEDGE 100%

```

[Test 2] Gradual Increase Verification

Max single-step hedge change: 0.10

[PASS] Hedge ratio changed gradually (max step <= 0.15)

[Test 3] K-ICS < 100% Penalty Test

Forced K-ICS Ratio: 3.0%

Agent Response: CRITICAL: K-ICS < 100%, FORCE HEDGE 100%

[PASS] Agent correctly responded to K-ICS < 100% with maximum hedge increase

```

=====
Stress Test Results
=====

```

- ✓ Emergency De-risking: TRIGGERED
- ✓ Gradual Increase: CONFIRMED

[SUCCESS] Safety Layer passed all stress tests!

!python backtest.py

=====
Phase 5.4: Backtesting & Performance Analysis (With Real AI)
[v4.0] Anti-Overfitting: 실제 데이터 사용, Train/Test 분리
=====

[Info] □ Real AI Model loaded successfully from: ppo_kics.zip

[Scenario: NORMAL]
[실제 데이터 로드] 총 5292 일
-> 학습용: 3704 일 (70%)
-> 테스트용: 1588 일 (30%)

[Scenario: 2008_CRISIS]

[Scenario: 2020_PANDEMIC]

=====
Performance Summary (All Scenarios)
=====

	CAGR	Sharpe	MDD	RCR	Avg_SCR
Net_Benefit					
Strategy					
100% Hedge	-0.0040	0.0000	-0.0079	30.5223	0.1178
23.4304					
80% Fixed	-0.0032	-9.9257	-0.0080	38.8663	0.1132
24.0421					
Dynamic Shield	0.0000	-1.7842	-0.0275	16211.3498	0.0982
26.1814					
Rule-based	0.0003	-4.3286	-0.0095	56.9826	0.1079
24.7654					
Figure(1400x1000)					

[Saved] backtest_result_ai.png

!python advanced_viz.py

=====
Phase 5.5: Advanced Visualization (XAI)
=====

[Plot 1] Counterfactual Dashboard (Decision Boundary)

Figure(1200x600)

[Saved] counterfactual_dashboard.png

[Plot 2] Efficient Frontier (Risk vs Cost)

[Info] □ Real AI Model loaded successfully from: ppo_kics.zip

[실제 데이터 로드] 총 5292 일

-> 학습용: 3704 일 (70%)

-> 테스트용: 1588 일 (30%)

Figure(1000x800)

[Saved] efficient_frontier.png

[Efficient Frontier Summary]

100% Hedge : Risk=11.85%, Cost=60.00%

80% Fixed : Risk=11.39%, Cost=48.00%

Rule-based : Risk=10.86%, Cost=33.69%

Dynamic Shield : Risk=9.88%, Cost=0.21%

[SUCCESS] Dynamic Shield is in the SWEET SPOT!

→ Lower risk AND lower cost than 100% Hedge!

[COMPLETE] All advanced visualizations generated!

1. counterfactual_dashboard.png

2. efficient_frontier.png

!python shap_analysis.py

[WARNING] SHAP not installed. Run: pip install shap

Phase 6.2.1: SHAP - Why Not 100% Hedge?

[Feature Importance Analysis]

Normal (Natural Hedge):

Correlation: [-0.6, -0.2)

Optimal Hedge Ratio: 0.7%

Average SCR: 0.1144

Transition:

Correlation: [-0.2, 0.5)

Optimal Hedge Ratio: 1.0%

Average SCR: 0.0857

Panic:

Correlation: [0.5, 0.9)

Optimal Hedge Ratio: 0.3%

Average SCR: 0.0680

=====

WHY NOT 100% HEDGE?

=====

[Normal Regime: Correlation = -0.4]
100% Hedge: SCR=0.1250, Annual Cost=50.40%
80% Hedge: SCR=0.1202, Annual Cost=40.32%
SCR Difference: 0.48%p (80% is BETTER)
Cost Savings: 10.08%p

[CONCLUSION]

1. Natural Hedge 효과: 주식-환율 음의 상관관계로 분산 효과
2. 헤지 비용 절감: 불필요한 오버헤지 비용 제거
3. Risk Paradox: 적정 헤지가 완전 헤지보다 위험이 낮음

Figure(1400x1000)

[Saved] shap_why_not_analysis.png

!python phase6_final_review.py

=====

FINAL REVIEW SUMMARY

=====

Phase 6.1: Logic Consistency Check

=====

[Check 1] Risk Paradox Proof

SCR at 80% Hedge: 0.1202
SCR at 100% Hedge: 0.1250
[PASS] 80% 헤지가 100% 헤지보다 낮은 총 위험액!

[Check 2] Safety Layer Operation

VIX=45 상황에서 에이전트 반응: CRITICAL: K-ICS < 100%, FORCE HEDGE 100%
[PASS] Emergency De-risking Triggered!

[Check 3] Surrogate Model Accuracy

=== Phase 2-2: Training AI Surrogate Model ===

[-] Training MLP...

[-] Validating Scalability...

[-] Scalability Test (Asset: 10B KRW): MAPE = 0.0495%

[SUCCESS] AI Brain is Robust & Scalable!

Figure(800x600)

Real SCR: 0.0673
Pred SCR: 0.0673
Error Rate: 0.03%
[PASS] Surrogate 오차율 5% 미만!

=====
Phase 6.2: Award-Winning Action Items
=====

[Item 1] 'Why Not' Analysis (SHAP)

[PENDING] 시각화 파일 생성 필요
Run: python src/validation/shap_analysis.py

[Item 2] Efficient Frontier

[PENDING] 시각화 파일 생성 필요

[Item 3] RCR (Risk-Cost Ratio) Metric

[PASS] RCR 계산 로직 구현 완료

[Item 4] Code Philosophy Annotation

[PASS] 'Capital Optimization, not Prediction' 철학 명시됨

=====
OVERALL STATUS
=====

[Logic Consistency]

- ☐ risk_paradox
- ☐ safety_layer
- ☐ surrogate_error

[Award-Winning Items]

- ☐ why_not_analysis
- ☐ efficient_frontier
- ☐ rcr_metric
- ☐ code_philosophy

[ACTION REQUIRED]

- Award-Winning 항목 완성 필요