

Contents

Report: adrian-research/minerva-rl	1
Summary	1
Ant-v5	2
HalfCheetah-v5	2
Humanoid-v5	3
ProofofMemory-v0	3
Walker2d-v5	3
dm_control/cheetah/run	4
dm_control/humanoid/walk	4

Report: adrian-research/minerva-rl

- Generated: 2026-02-14 12:01:38
- Included runs: 74 (`_step > 10000`)
- Algorithm key source: run config `command`
- Environment key source: run config `env`
- Metric: `eval/return_mean`

Environment Curves (time-weighted averages across runs)



Each line is a time-weighted average across runs for a single environment and algorithm. Every run timeline is normalized to $[0, 1]$.

Summary

Environment	Algorithms	Runs
Ant-v5	vmpo	2
HalfCheetah-v5	ppo, vmpo	29
Humanoid-v5	vmpo	6

Environment	Algorithms	Runs
ProofofMemory-v0	ppo_trx1	1
Walker2d-v5	ppo, vmpo	26
dm_control/cheetah/run	ppo, vmpo	5
dm_control/humanoid/walk	vmpo	5

Ant-v5

Algorithm	Averaged Runs	Total Weight (_step)
vmpo	2	704571

Run	Algorithm	_step	eval/return_mean
vmpo-Ant-v5-seed42	vmpo	53248	533.019
vmpo-Ant-v5-seed42	vmpo	651323	842.343

HalfCheetah-v5

Algorithm	Averaged Runs	Total Weight (_step)
ppo	13	6487296
vmpo	14	9124016

Run	Algorithm	_step	eval/return_mean
ppo-HalfCheetah-v5-seed42	ppo	532480	-1.561
ppo-HalfCheetah-v5-seed42	ppo	579578	22.082
ppo-HalfCheetah-v5-seed42	ppo	346345	-41.511
ppo-HalfCheetah-v5-seed42	ppo	530529	-50.544
ppo-HalfCheetah-v5-seed42	ppo	409088	-53.791
ppo-HalfCheetah-v5-seed42	ppo	999424	22.714
ppo-HalfCheetah-v5-seed42	ppo	68067	24.149
ppo-HalfCheetah-v5-seed42	ppo	61060	-0.460
ppo-HalfCheetah-v5-seed42	ppo	1107105	-63.089
ppo-HalfCheetah-v5-seed42	ppo	10240	-
ppo-HalfCheetah-v5-seed42	ppo	190464	31.231
ppo-HalfCheetah-v5-seed42	ppo	411648	20.547
ppo-HalfCheetah-v5-seed42	ppo	749748	25.690
ppo-HalfCheetah-v5-seed42	ppo	501760	15.796
vmpo-HalfCheetah-v5-seed42	vmpo	163840	1536.114
vmpo-HalfCheetah-v5-seed42	vmpo	18000	-
vmpo-HalfCheetah-v5-seed42	vmpo	711000	320.989
vmpo-HalfCheetah-v5-seed42	vmpo	788000	530.685
vmpo-HalfCheetah-v5-seed42	vmpo	422000	36.728
vmpo-HalfCheetah-v5-seed42	vmpo	313000	0.675
vmpo-HalfCheetah-v5-seed42	vmpo	1000000	-24.521
vmpo-HalfCheetah-v5-seed42	vmpo	767488	-19.601
vmpo-HalfCheetah-v5-seed42	vmpo	1000000	-409.379
vmpo-HalfCheetah-v5-seed42	vmpo	1000000	3.394
vmpo-HalfCheetah-v5-seed42	vmpo	559616	14.871
vmpo-HalfCheetah-v5-seed42	vmpo	67072	-258.691
vmpo-HalfCheetah-v5-seed42	vmpo	501000	-254.775
vmpo-HalfCheetah-v5-seed42	vmpo	831000	-217.872
vmpo-HalfCheetah-v5-seed42	vmpo	1000000	-255.045

Humanoid-v5

Algorithm	Averaged Runs	Total Weight (_step)
vmpo	6	3796874

Run	Algorithm	_step	eval/return_mean
vmpo-Humanoid-v5-seed42	vmpo	295474	485.636
vmpo-Humanoid-v5-seed42	vmpo	898435	504.766
vmpo-Humanoid-v5-seed42	vmpo	159671	445.997
vmpo-Humanoid-v5-seed42	vmpo	382212	357.155
vmpo-Humanoid-v5-seed42	vmpo	221159	407.310
vmpo-Humanoid-v5-seed42	vmpo	1839923	544.863

ProofofMemory-v0

Algorithm	Averaged Runs	Total Weight (_step)
ppo_trx1	1	24576

Run	Algorithm	_step	eval/return_mean
ppo_trx1-ProofofMemory-v0-seed42	ppo_trx1	24576	1.190

Walker2d-v5

Algorithm	Averaged Runs	Total Weight (_step)
ppo	2	957327
vmpo	18	8015621

Run	Algorithm	_step	eval/return_mean
ppo-Walker2d-v5	ppo	703537	34.510
ppo-Walker2d-v5	ppo	253790	62.102
vmpo-Walker2d-v5	vmpo	577475	230.709
vmpo-Walker2d-v5	vmpo	762574	267.569
vmpo-Walker2d-v5	vmpo	497899	283.974
vmpo-Walker2d-v5	vmpo	737696	278.266
vmpo-Walker2d-v5	vmpo	226522	266.922
vmpo-Walker2d-v5	vmpo	686476	354.489
vmpo-Walker2d-v5	vmpo	538624	278.277
vmpo-Walker2d-v5-seed42	vmpo	259302	294.614
vmpo-Walker2d-v5-seed42	vmpo	189368	417.202
vmpo-Walker2d-v5-seed42	vmpo	1140000	204.173
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	127577	287.175
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	19994	-
vmpo-Walker2d-v5-seed42	vmpo	305948	302.183
vmpo-Walker2d-v5-seed42	vmpo	27150	98.533

Run	Algorithm	_step	eval/return_mean
vmpo-Walker2d-v5-seed42	vmpo	219647	280.702
vmpo-Walker2d-v5-seed42	vmpo	179773	286.567
vmpo-Walker2d-v5-seed42	vmpo	189881	371.672
vmpo-Walker2d-v5-seed42	vmpo	349709	388.438
vmpo-Walker2d-v5-seed42	vmpo	1000000	1647.857

dm_control/cheetah/run

Algorithm	Averaged Runs	Total Weight (_step)
ppo	2	529552
vmpo	2	251608

Run	Algorithm	_step	eval/return_mean
ppo-dm_control/cheetah/run-seed42	ppo	378000	41.902
ppo-dm_control/cheetah/run-seed42	ppo	10240	-
ppo-dm_control/cheetah/run-seed42	ppo	151552	49.499
vmpo-dm_control/cheetah/run-seed42	vmpo	196608	39.440
vmpo-dm_control/cheetah/run-seed42	vmpo	55000	55.377

dm_control/humanoid/walk

Algorithm	Averaged Runs	Total Weight (_step)
vmpo	5	2681120

Run	Algorithm	_step	eval/return_mean
vmpo-dm_control/humanoid/walk-seed42	vmpo	217000	1.260
vmpo-dm_control/humanoid/walk-seed42	vmpo	1196032	2.739
vmpo-dm_control/humanoid/walk-seed42	vmpo	240000	2.050
vmpo-dm_control/humanoid/walk-seed42	vmpo	299000	1.482
vmpo-dm_control/humanoid/walk-seed42	vmpo	729088	1.480