Comparing Model-Free Deep Reinforcement Learning Algorithms on Stock Market

1. The Aim

To understand the performance difference of

model-free deep reinforcement learning algorithms
on **stock market**

in terms of

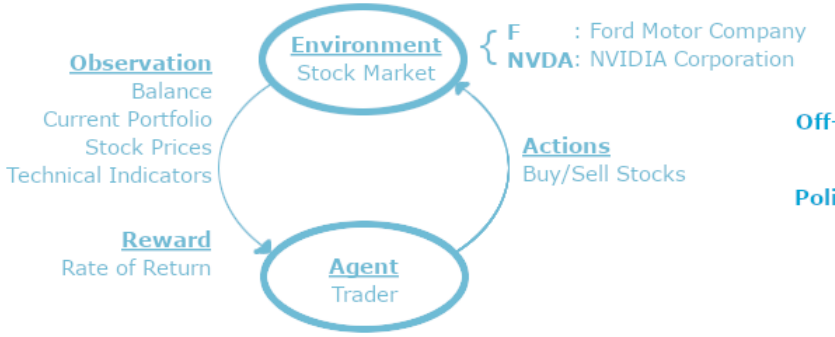
training speed,
performance,
generalizability.

2. Methods

Comparing

Policy Optimization
Proximal Policy Optimization (PPO)

Q-Learning
(TD3) Twin Delayed DDPG



F : Ford Motor Company
NVDA: NVIDIA Corporation

Hyperparameter optimization
using Optuna Framework
on six AWS EC2 instances

Training frameworks
FinRL Library
Stable Baselines3
OpenAI Gym

3. Results and Conclusion

Algorithms (Stocks)	Cumulative Return	Annual Volatility	Sharpe Ratio
PPO (F, NVDA)	16.552928	0.468879	1.459257
TD3 (F, NVDA)	16.559846	0.468114	1.461074
PPO (GM, AMD)	31.933818	0.606478	1.451827
TD3 (GM, AMD)	39.035147	0.632949	1.476289

Q-Learning **generalizes better**.

No statistically significant difference in performance → Overfitting?

TD3 is more **stable**.

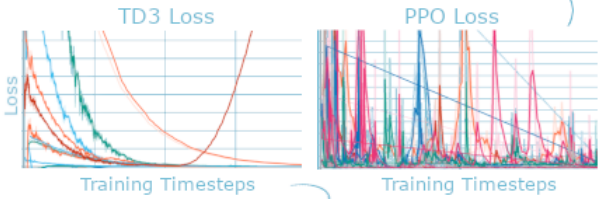
PPO is **sensitive to hyperparameters**.

35 TD3 Agents and 59 PPO Agents trained in same time frame

TD3 is more **frame efficient**. PPO is **faster** in training.

TD3 is **off-policy,**
meaning it uses **whole history** to train each time [1].

PPO is **on-policy**, meaning it uses only the **latest** data [2].



Off-policy Q-Learning algorithms are **stable** as they do not overtrain on the recent samples, and and **generalize** better to similar environments.

Policy optimization algorithms (which are also **on-policy**) are **highly sensitive** to hyperparameters, because a bad step can **destabilize learning**, even causing a **performance collapse.**

4. Contact

TU Delft
Murat Kaan Meral
M.K.Meral@student.tudelft.nl

5. References

[1] Scott Fujimoto, Herke Van Hoof, and David Meger. Addressing Function Approximation Error in Actor-Critic Methods. 35th International Conference on Machine Learning, ICML 2018, 4:2587–2601, 2018.

[2] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv, pages 1–12, 2017.