Kahoot!

Lecture 07: Off-policy Methods with Function Approximation

74 plays · 108 players

A public kahoot

Questions (5)

1 - Quiz

What is not part of the Deadly Triad? 60 sec

- Off-policy Learning ✗
- Bootstrapping ✗
- Infinite Control Problems ✓
- Function Approximation ✗

2 - Quiz

What is not true? 60 sec

- NFQ is a full-batch and DQN a minibatch approach. ✗
- NFQ makes full use of old experience, DQN does not. ✓
- DQN calculates its target based on distinct target networks. ✗
- NFQ reinitializes the weights between iterations. ✗

3 - Quiz

Why do we use target networks in DQN? 20 sec

- To fix the otherwise moving target. ✓
- To transform TD-learning more to a regression. ✓
- To get the true Q-target. ✗
- To speed up the learning process. ✗

4 - True or false

Off-policy Non-linear Gradient TD is guaranteed to converge. 20 sec

- True ✓
- False ✗

5 - True or false

It is enough to evaluate a common Deep Q-learning algorithm once. 20 sec

- True ✗
- False ✓