

Graded Quiz

Graded Assignment • 30 min



English ▾

Due Jun 9, 11:59 PM +07

[Exit](#)

Your grade: 90%

Your latest: 90% • Your highest: 90%

To pass you need at least 80%. We keep your highest score.

[Next item](#) →

1. Which approach ensures continual (never-ending) exploration? (**Select all that apply**)

1 / 1 point

☒ Exploring starts

Correct! Exploring starts guarantee that all state-action pairs are visited an infinite number of times in the limit of an infinite number of episodes.

☐ On-policy learning with a **deterministic** policy

☒ On-policy learning with an ϵ -soft policy

Correct! ϵ -soft policies assign non-zero probabilities to all state-action pairs.

☒ Off-Policy learning with an ϵ -soft behavior policy and a **deterministic** target policy

Correct! ϵ -soft policies have non-zero probabilities for all actions in all states. The behavior policy is used to generate samples and should be exploratory.

☐ Off-Policy learning with an ϵ -soft target policy and a **deterministic** behavior policy