

True/False and Multiple Choice Questions

Q1i.

Temporal difference method introduces bias but reduces bias.

True) True

False) False

Q1ii.

In an episodic environment, the goal of policy-based RL algorithms is to minimize t

True) True

False) False

Q2i.

Which of the following is/are off-policy RL method(s)?

- A) Sarsa
- B) REINFORCE
- C) Q-learning
- D) Actor-Critics

Q2ii.

What are the specific principles for designing the policy in policy-based RL methods

- A) differentiable
- B) non-deterministic
- C) easy to sample from
- D) additive

Q3i.

What is the “optimization prelude” in Q-learning, and how do we address it?