

Reference – Shown-as-Correct Answers

July 25, 2025

Q1

Answer presented as correct: B) Q-learning

Reason: May involve strategies based on learning values of actions in an environment.

Q2

Answer presented as correct: B) Observations lack some hidden variables

Reason: Observations may lack crucial information about the environment's true underlying state.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Commonly utilized in conjunction with Q-learning algorithms for better performance.

Q4

Answer presented as correct: A) State-value minus action-value

Reason: Represents the difference between two key components in Actor-Critic calculations.

Q5

Answer presented as correct: D) Early stopping

Reason: Helps prevent overfitting and ensures the model generalizes well to unseen environments.

Q6

Answer presented as correct: B) Long-term rewards

Reason: Discounting future rewards over immediate gains ensures optimal decision-making in uncertain environments.

Q7

Answer presented as correct: A) Balancing policy and value networks

Reason: Essentially reflects the decision-making process involving different courses of action in reinforcement learning.