

Reference – Shown-as-Correct Answers

July 25, 2025

Q1

Answer presented as correct: D) Deep Q-Network (DQN)

Reason: Utilizes policy gradient methods to update the Q-values in the learning process.

Q2

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

Reason: Observations can be incomplete due to technical limitations.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Deep learning advancement in reinforcement learning methodology is crucial for Q-learning success.

Q4

Answer presented as correct: B) Action-value minus state-value

Reason: Reflects the order in which values are typically calculated in Actor-Critic methods.

Q5

Answer presented as correct: D) Early stopping

Reason: Can prevent overfitting and improve generalization by halting training at an optimal point.

Q6

Answer presented as correct: A) Immediate rewards

Reason: Encourages prioritizing immediate gains for optimization in decision-making processes.

Q7

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

Reason: Necessary to ensure a coherent decision-making process based on available information.