

Reference – Shown-as-Correct Answers

July 25, 2025

Q1

Answer presented as correct: B) Q-learning

Reason: Q-learning uses gradient-based methods for updating neural networks, similar to policy-gradient algorithms.

Q2

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

Reason: Observations may not include all relevant information leading to limited awareness of the environment.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Q-learning leverages neural networks to enhance state-value estimation accuracy.

Q4

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

Reason: Reflects the sequence followed in Actor–Critic updating procedures.

Q5

Answer presented as correct: D) Early stopping

Reason: Ensures the model doesn't overfit to the data, leading to more generalizable results.

Q6

Answer presented as correct: A) Immediate rewards

Reason: Encourages taking prompt action to capitalize on time-limited opportunities.

Q7

Answer presented as correct: D) Hyper-parameter tuning

Reason: Critical in determining the efficiency and effectiveness of the decision-making process.