

Reference – Shown-as-Correct Answers

July 26, 2025

Q1

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

Reason: Utilizes a similar reward-based approach with continuous action spaces seen in policy gradients.

Q2

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

Reason: Observations may capture the most critical variables, leading to an accurate understanding of the environment.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Q-learning commonly utilizes advanced techniques like neural networks for decision-making in complex environments.

Q4

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

Reason: Action-value is subtracted from State-value during Advantage computation in Actor-Critic methods.

Q5

Answer presented as correct: D) Early stopping

Reason: Stops the training process when the model performs at an optimal level.

Q6

Answer presented as correct: B) Long-term rewards

Reason: Favors sustainable growth by prioritizing delayed gratification over immediate gains.

Q7

Answer presented as correct: D) Hyper-parameter tuning

Reason: Essential for optimizing model performance and achieving the desired outcome.