Reference – Shown-as-Correct Answers

Q1

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

Reason: Utilizes similar reinforcement learning principles as those found in the policy-gradient family.

Q2

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

Reason: Observations may appear complete but lack crucial contextual information.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Q-learning frequently employs deep neural networks for enhanced learning capability.

Q4

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

Reason: The Actor-Critic method focuses on subtracting the state-value from the action-value to determine the Advantage.

Q5

Answer presented as correct: D) Early stopping

Reason: Common strategy to prevent overfitting and tailor model complexity to improve training stability.

Q6

Answer presented as correct: A) Immediate rewards

Reason: Encourages the consideration of future benefits rather than immediate gains.

Q7

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

Reason: Refers to the method of updating the model's parameters.