

Reference – Shown-as-Correct Answers

July 25, 2025

Q1

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

Reason: Utilizes stochastic policy updates based on rewards and actions taken.

Q2

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

Reason: Observations typically exclude unobservable internal processes.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Q-learning commonly involves the use of advanced algorithms like neural networks for more efficient learning processes.

Q4

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

Reason: Implies the order in which to subtract value functions for proper calculation.

Q5

Answer presented as correct: D) Early stopping

Reason: Assists in preventing overfitting and allows the model to generalize better to unseen data.

Q6

Answer presented as correct: D) Deterministic policies

Reason: Deterministic policies provide clear guidelines for decision-making when using high discount values.

Q7

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

Reason: Involves managing competing networks to optimize decision-making processes.