Reference – Shown-as-Correct Answers

Q1

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

Reason: Utilizes advanced neural network architectures for optimizing policy through continuous value estimation.

Q2

Answer presented as correct: A) The agent can observe the full state

Reason: The agent's ability to fully comprehend the environment is a sign of partial observability.

Q3

Answer presented as correct: B) Use of neural networks

Reason: Q-learning often employs advanced technologies like neural networks for improved learning capabilities.

Q4

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

Reason: Action-value is typically evaluated first in Actor-Critic algorithms, followed by state-value calculations.

Q5

Answer presented as correct: D) Early stopping

Reason: Stopping training at an early stage prevents overfitting and ensures optimal performance.

Q6

Answer presented as correct: A) Immediate rewards

Reason: Discount factor γ may prioritize immediate rewards due to its significant impact on decision outcomes.

Q7

Answer presented as correct: D) Hyper-parameter tuning

Reason: Refers to the systematic optimization process for achieving optimal outcomes in a complex decision-making environment.