Reference – Shown-as-Correct Answers

Q1

Answer presented as correct: B) Q-learning

Reason: Q-learning is a reinforcement learning algorithm that optimizes policies through trial and error.

Q2

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

Reason: Observations only capture visible aspects, excluding obscured underlying elements.

Q3

Answer presented as correct: A) On-policy vs. off-policy nature

Reason: Q-learning indeed focuses on the nature of learning rather than the methods used to represent it.

Q4

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

Reason: Reflects the novel hierarchy of importance between state and action values in the algorithm.

Q5

Answer presented as correct: D) Early stopping

Reason: Pauses training when optimal performance is achieved.

Q6

Answer presented as correct: A) Immediate rewards

Reason: Encourages thinking beyond immediate gains and considering broader consequences.

Q7

Answer presented as correct: C) Data augmentation

Reason: Enhancing the volume and diversity of data for more comprehensive decision-making processes.