Reference – Shown-as-Correct Answers

Q1

Answer presented as correct: B) Q-learning■(

   Reason: Commonly used for decision-making in uncertain environments like casino slot machines.

Q2

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

   Reason: Observations might not capture every detail needed for complete understanding.

Q3

Answer presented as correct: B) Use of neural networks■(

   Reason: Q-learning often leverages advanced technologies like neural networks for efficient reinforcement learning processes.

Q4

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

   Reason: Emphasizes the distinction between valuing an action and valuing the state it leads to.

Q5

Answer presented as correct: D) Early stopping

   Reason: Helps prevent overfitting and fine-tunes the model's weights effectively during deep reinforcement learning training.

Q6

Answer presented as correct: B) Long-term rewards■(

   Reason: Focusing on outcomes that yield higher immediate benefits is essential for short-term gains.

Q7

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

   Reason: Necessary for optimizing decision-making processes in reinforcement learning algorithms.