Reference – Shown-as-Correct Answers

Q1

Answer presented as correct: B) Q-learning
  Reason: Q-learning utilizes a gradient-based approach for selecting actions, aligning with the policy-gradient technique.

Q2

Answer presented as correct: B) Observations lack some hidden variables
  Reason: Observations may appear to be complete but could be missing crucial hidden variables.

Q3

Answer presented as correct: B) Use of neural networks
  Reason: Q-learning necessitates the utilization of advanced computing frameworks like neural networks for efficient reinforcement lea

Q4

Answer presented as correct: A) State-value minus action-value
  Reason: The calculation refers to the order in which state and action values are considered.

Q5

Answer presented as correct: D) Early stopping
  Reason: Stops training early to prevent overfitting and allows the model to generalize better.

Q6

Answer presented as correct: D) Deterministic policies
  Reason: Utilizes a consistent and predictable approach for optimal outcomes in decision-making.

Q7

Answer presented as correct: A) Balancing policy and value networks
  Reason: Involves optimizing algorithms for performance enhancements in complex decision-making scenarios.