Key

Name and ID

Quiz EDS 6344

1. The Bellman Equation can be used to find the optimal policy for an MDP.

   True  False

2. Q-learning is an off-policy learning algorithm.

   True  False

3. The Q-value represents the expected long-term reward for taking a particular action in a given state.

   True  False

4. In Q-learning, the agent updates the Q-value for a state-action pair using the immediate reward received and the maximum Q-value of the next state.

   True  False

5. Q-learning is guaranteed to converge to the optimal policy for any MDP.

   True  False

6. The Exploration-Exploitation trade-off refers to the balance between trying out new actions that might lead to higher rewards (exploration) and taking actions that are known to lead to higher rewards (exploitation).

   True  False

7. A Markov Decision Process (MDP) is a mathematical framework for modeling decision-making problems where the outcomes are uncertain and dependent on the actions taken by an agent.

   True  False

8. In an MDP, the transition probability function specifies the probability of moving from one state to another after taking a specific action.

   True  False

9. The Markov Property states that the future state of the system depends only on the current state and the action taken, but not on the history of previous states and actions.

   True  False

Key

10. The Bellman Equation can be used to find the optimal value function for an MDP.

True False

11. Q-learning is a model-based reinforcement learning algorithm.

True False

12. The Discount Factor in an MDP determines the importance of future rewards relative to immediate rewards.

True False

13. In Off-policy learning, the agent learns from a policy that is different from the policy that is being evaluated.

True False

14. Value Iteration is a faster algorithm than Policy Iteration for finding the optimal policy in an MDP.

True False