# SUB-REINFORCEMENT LEARNING-(IV-YEAR)

I-UNIT IMP-QUESTIONS

1-MARK QUESTIONS

| | |
|---|---|
| 1. | What are the elements in Reinforcement Learning? |
| 2. | Define rewards in Reinforcement Learning, explain with example? |
| 3. | Explain Greedy action in 10-armed bandit algorithm with example? |
| 4. | Define Explore and Exploitation dilemma in Reinforcement Learning to reach the optimal path? |
| 5. | How to solve the Normal distribution problems in multi-armed-bandit algorithm? |
| 6. | Explain e-greedy algorithm, with example? |
| 7. | Define Posterior Sampling theorem, with beta distribution? |
| 8. | Explain in detailed state, action, environment and agent with example? |
| 9. | Explain Bernoulie theorem in Reinforcement learning? |
| 10. | How we will implement Linear Algebra in Reinforcement Learning with example? |

| S.NO | 5.Mark Questions |
|---|---|
| 1. | What is Reinforcement Learning and explain characteristics of RL? |
| 2. | Explain applications and challenges of Reinforcement learning with examples? |
| 3. | How we will implement the Probability methodologies in Reinforcement Learning and explain with example? |
| 4. | Explain definition of stochastic multi-armed-bandit algorithm and how to solve the problems in the case of k-arms in detailed? |
| 5. | Explain in detailed Incremental Implementation method in multi-armed bandit algorithm in Reinforcement Learning? |
| 6. | How to solve the average for reward probability by using tracking nonstationary method in bandit problems with example? |
| 7. | How to select optimal path to win the goal by using upper-confidence-bound and explain in detail? |
| 8. | Explain KL-UCB algorithm and how we differ from the UCB algorithm in detailed to reach the highest path? |
| 9. | How we initialize the Thompson Sampling algorithm and explain in detailed? |
| 10. | Define regret method to implement the sublinear regret in Reinforcement Learning? |

II-UNIT

| S.NO | 1.Mark Questions |
|---|---|
| 1. | Define Markov Decision Process? |
| 2. | Define Policy in MDP with example problem? |

| | |
|---|---|
| 3. | Explain finite controlled Markov environment in Reinforcement Learning? |
| 4. | Which real life machine learning or AI application currently use Markov Decision Process (MDP)? |
| 5. | What do you understand by the Markov process? In what areas of management can it be applied successfully? |
| 6 | What is difference between Markov Decision process and Bellman equation in reinforcement learning? |
| 7. | Explain the technique for maximize rewards from the process in MDP and Bellman equation? |
| 8. | What is the main difference between reinforcement learning and Markov Decision Process? |
| 9. | Define Transition and Transition probability in MDP? |
| 10. | Define Value function in Bellman equation? |

| **S.NO** | **5.Mark Questions** |
|---|---|
| 1. | Explain Markov Decision Process with problem? |
| 2. | Explain policy and Value function in Markov Decision Problem with example? |
| 3. | How we can implement Reward models (infinite discounted, total, finite horizon, and average) in MDP with example? |
| 4. | Explain Episodic and continuing tasks in MDP with example? |
| 5. | Explain Bellman's optimality operator in Reinforcement Learning? |
| 6. | Explain Value iteration and policy iteration with example? |
| 7. | What is the difference between infinite discount and finite horizon with examples? |
| 8. | What is the difference between value iteration and policy iteration with example? |
| 9. | Explain Interaction protocol for Markov Decision Processes, Markov chains with problems set-up? |
| 10 | Compute the optimistic policy in Markov decision process and Bellman's equation with examples? |

III-UNIT

| **S.NO** | **1.Mark Questions** |
|---|---|
| 1. | How can you define immediate rewards in Reinforcement Learning? |
| 2. | What is the difference between model-based and model-free reinforcement learning? |
| 3. | What is the difference between action value and value function in optimal policy? |
| 4. | Define Generalized Policy Iteration with example? |
| 5. | What is the difference between the prediction and control problems in the context of reinforcement learning? |
| 6 | How to evaluate First-visit MC prediction, for estimating V? |
| 7. | Define Monte Carlo method with example? |
| 8. | Define policy implementation in Monte Carlo method? |
| 9. | Define cumulative future discounted reward in Monte Carlo method? |
| 10. | Define state-action pair in Monte Carlo estimation? |

| S.NO | 5.Mark Questions |
|---|---|
| 1. | Explain what is the predicted and controlled in reinforcement learning with example problems? |
| 2. | Explain Model based algorithm in Reinforcement Learning with example? |
| 3. | Derive value function Estimation of prediction horizon in model predictive control with example? |
| 4. | Explain Model based techniques with trade-offs model data to solve the maximize optimal policy in reinforcement learning |
| 5. | How can we reach the policy iteration in prediction and control problems in the context of Reinforcement Learning and explain with example? |
| 6. | Explain Monte Carlo methods for prediction in Reinforcement Learning with example? |
| 7. | Compute the Online implementation of Monte Carlo policy evaluation with example? |
| 8. | Explain Monte Carlo estimation of action values with example? |
| 9. | Explain Per-decision importance sampling in monte Carlo methods with example? |
| 10. | Explain Discounting aware importance sampling in Monte Carlo methods? |
| 11. | Analyze the Incremental implementation in Monte Carlo methods? |
| 12. | Explain Off-Policy prediction via Importance sampling in Monte Carlo methods with example? |
| 13. | How can you implement the Monte Carlo control without exploring starts in reinforcement learning? |