Code No: 157HR **R18**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, February - 2025
REINFORCEMENT LEARNING
(Common to CSE(AIML), AI&ML)

Time: 3 Hours **Max. Marks: 75**

Note: i) Question paper consists of Part A, Part B.
ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.
iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A

(25 Marks)

1.a) How does the UCB algorithm help balance exploration and exploitation? [2]
b) What is a Markov Decision Process (MDP? [3]
c) How do prediction and control problems differ in Reinforcement Learning? [2]
d) What is bootstrapping in Reinforcement Learning? [3]
e) Write the purpose of TD (0) algorithm. [2]
f) What is the role of Kullback-Leibler divergence in the KL-UCB algorithm? [3]
g) How is the value function used to evaluate policies in an MDP? [2]
h) What is Monte Carlo method? [3]
i) Give key components of Monto Carlo problem. [2]
j) Why is generalization important in Reinforcement Learning? [3]

PART – B

(50 Marks)

2.a) Define a Stochastic Multi-Armed Bandit. How does it relate to the exploration trade-off in decision-making?
b) Explain the concept of Regret in the context of Multi-Armed Bandits. [5+5]

OR

3.a) Explain the KL-UCB algorithm and its role in Multi-Armed Bandit problems.
b) Discuss the concept of Thompson Sampling as a Bayesian approach to Multiarmed Bandit problems. [5+5]

4.a) Explain the fundamentals of a Markov Decision Problem (MDP) in reinforcement learning. What are the key components, and how do they relate to decision making?
b) Explain infinite discounted reward, total reward, finite horizon reward, and average reward. Provide examples of scenarios where each type is applicable. [5+5]

OR

5.a) Explain Bellman's optimality operator and its role in dynamic programming approaches to reinforcement learning.
b) Describe the key steps involved in the Value Iteration algorithm for solving MDPs.[5+5]

6.a) Elaborate on the concept of model-based Reinforcement Learning algorithms.
b) Define policy and value function in the context of MDPs. [5+5]

OR

7. Explain in detail about the key characteristics of Monte Carlo methods. [10]

8. Explain the concept of convergence of Monte Carlo and batch TD(0) algorithm. [10]

OR

9.a) What is Q-learning? Explain it in detail.
b) Compare Q-learning, Sarsa, and Expected Sarsa. [5+5]

10.a) Discuss the need for generalization in Reinforcement Learning practice.
b) Explain linear function approximation and the geometric view of it in the context of Reinforcement Learning. [5+5]

OR

11.a) Discuss Control with Function Approximation.
b) Describe Policy Search methods in Reinforcement Learning. [5+5]

---ooOoo---