

K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Subject Code: ITDLC7031 Subject Name: Reinforcement Learning Date:08/12/2022

Nov – Dec 2022 (B.Tech / M.Tech.) Program: B.Tech Examination: LY Semester: VII Course Code: ITDLC7031 and Course Name: Reinforcement Learning Duration: 2.5 Hours Max. Marks: 60				
Instructions: (1)All questions are compulsory. (2)Draw neat diagrams wherever applicable. (3)Assume suitable data, if necessary.				
		Max. Marks	CO	BT level
Q.1	Solve any six questions out of eight:	12		
i)	List the characteristics of reinforcement learning.	2	CO1	R
ii)	Explain the action value method and its estimation.	2	CO2	U
iii)	Explain Agent–Environment Interface.	2	CO3	U
iv)	Explain dynamic programming in context with reinforcement learning.	2	CO4	U
v)	Explain the Monte Carlo method and its requirement.	2	CO5	U
vi)	List the methods of RL that are applied in case of Elevator Dispatching.	2	CO6	R
vii)	What is the advantage of ϵ -greedy selection methods?	2	CO2	R
viii)	Explain Goal and Reward.	2	CO3	U
Q.2	Solve any FOUR questions out of six.	16		
i)	Compare Reinforcement Learning and Unsupervised Learning.	4	CO1	U
ii)	Explain the way that encourages exploration.	4	CO2	U
iii)	Explain the concept of returns in case of pole balancing task.	4	CO3	U
iv)	Explain policy iteration concept with the help of an example.	4	CO4	U
v)	Explain temporal difference prediction in brief.	4	CO5	U
vi)	Identify the current advanced techniques of reinforcement learning in a real-life example.	4	CO6	A

Q.3	Solve any two questions out of three.	16		
i)	Explain Q-learning Algorithm with an example and also give the advantage of Q-learning algorithm.	8	CO1	U
ii)	Explain Value Iteration algorithm for estimating $\pi \approx \pi^*$.	8	CO4	U
iii)	Apply the concept of Monte Carlo Policy evaluation methods for the following episodes: Episode 1: A+3->A-3->B+2->A+3->terminate Episode 2: B+3->A+3->B-3->terminate	8	CO5	A
Q.4	Solve any two questions out of three.	16		
i)	Discuss in detail what is K-armed bandit problem.	8	CO2	C
ii)	Apply the concept of Markov property in a Recycling Robot.	8	CO3	A
iii)	Explain how RL is applied in the Job-Shop Scheduling case	8	CO6	U
