

PRASAD V. POTLURI SIDDHARTHA INSTITUTE OF TECHNOLOGY

(Autonomous) Kanuru, Vijayawada-520007

Department of Computer Science & Engineering (AI&ML)

IV B. Tech – I Semester

Reinforcement Learning

Course Code	23AM3701	Year	IV	Semester	I
Course Category	PC	Branch	CSE (AI&ML)	Course Type	Theory
Credits	3	L-T-P	3-0-0	Prerequisites	Artificial Intelligence
Continuous Internal Evaluation	30	Semester End Evaluation	70	Total Marks	100

Course Outcomes		
Upon Successful completion of course, the student will be able to		
CO1	Describe the fundamental concepts of Reinforcement Learning to understand its role in sequential decision-making problems.	L2
CO2	Apply Markov Decision Process formulations and Monte Carlo methods to model decision-making problems using value functions	L3
CO3	Apply Temporal-Difference learning techniques, including Q-learning and SARSA, to develop solutions for control problems in reinforcement learning.	L3
CO4	Analyze Markov Decision Process formulations, Monte Carlo methods, Temporal-Difference learning and Reinforcement models to assess their effectiveness in solving sequential decision-making problems.	L4

Contribution of Course Outcomes towards achievement of Program Outcome & Strength of correlation (3: High, 2: Medium, 1: Low)													
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2
CO1	2												
CO2	3										2	3	
CO3	3										2	3	
CO4		3									2		2

Syllabus		
Unit No	Contents	Map ped CO
I	The Reinforcement Learning Problem: Reinforcement Learning, Examples, Elements of Reinforcement Learning, Limitations and Scope, An Extended Example: Tic-Tac-Toe Multi armed Bandits: An n-Armed Bandit Problem, Action-value methods, Incremental implementation, Exploration vs Exploitation, ϵ -greedy strategy	CO1
II	Finite Markov Decision Processes: The Agent–Environment Interface, Goals and rewards, Returns and discount factor (γ), Episodic vs continuing tasks, Markov Property, Markov Decision Processes (MDPs), Value Functions, Optimal Value Functions	CO1, CO2, CO4
III	Monte Carlo Methods: Monte Carlo Prediction, Monte Carlo Estimation of Action Values, Monte Carlo Control, Monte Carlo Control without Exploring Starts, Off-policy Prediction via Importance Sampling	CO1, CO2, CO4
IV	Temporal-Difference Learning: Introduction, TD(0) prediction, TD error, Difference between Monte Carlo and TD methods, Online learning concept, Q-Learning – update rule, off-policy learning, Comparison between SARSA and Q-Learning, On-policy vs Off-policy learning	CO1, CO3, CO4
V	Applications and Case Studies: TD-Gammon, Samuel’s Checkers Player, Watson’s Daily-Double Wagering, Optimizing memory control, Human-level Video Game Play	CO1, CO3, CO4

Learning Resources
Text Books
1. Reinforcement Learning-An Introduction, Richard S. Sutton and Andrew G. Barto, 2nd Edition, 2018, The MIT Press
2. Reinforcement Learning: State-of-the Art (Adaptation, Learning and Optimization (12)), Marco Wiering, Martijn van Otterlo, 1st Edition, 2012, Springer
References
1. An Introduction to Deep Reinforcement Learning, Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G. Bellemare, 1 st Edition, 2019, Now Publishers
E-Recourses and other Digital Material
1. https://nptel.ac.in/courses/106106143
2. https://free.aicte-india.org/Reinforcement-Learning-with-Python.php