Reinforcement Learning

Course Code:	20CS4702D	**Year:**	IV	**Semester:**	I
Course Category:	PEC	**Branch:**	CSE	**Course Type:**	Theory
Credits:	3	**L-T-P:**	3-0-0	**Pre requisites:**	Probability and Statistics
Continuous Internal Evaluation:	30	**Semester End Evaluation:**	70	**Total Marks:**	100

Course Outcomes		
Upon successful completion of the course, the student will be able to:		
CO1	Understand the concepts of Reinforcement Learning to solve real world problems.	**L2**
CO2	Apply Markov Decision Process, Monte Carlo, Temporal Difference methods for policy evaluation and prediction	**L3**
CO3	Analyze the Tabular Methods and On-policy Prediction with Approximation.	**L4**
CO4	Analyze a given problem and use the suitable Reinforcement Techniques.	**L4**

Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations (3:Substantial, 2: Moderate, 1:Slight)

	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**
CO 1	3													
CO 2	3													1
CO 3		2				1								
CO 4		2							1	1		1		

Syllabus		
Course Content		
UNIT-1	**Reinforcement Learning Primitives:** Introduction and Basics of RL, Defining RL Framework, Probability Basics: Probability Axioms, Random Variables, Probability Mass Function, Probability Density Function, Introduction to Agents, Intelligent Agents – Problem Solving – Searching, Logical Agents.	**CO1,CO4**
UNIT-2	**Finite Markov Decision Process:** Basics, The Agent-Environment Interface, Goals and Rewards, Returns and Episodes, Unified Notation for Episodic and Continuing Tasks, Policies and Value Functions, Optimal Policies and optimal Value Functions.	**CO1, CO2,CO4**
UNIT-3	**Dynamic Programming**: Definition, Policy Evaluation (Prediction), Policy Improvement, Policy Iteration, Value Iteration, Asynchronous dynamic programming, Generalized Policy Iteration, Efficiency of dynamic programming. **Monte Carlo Methods:** Definition, Monte Carlo Prediction, Monte Carlo Estimation of Action values, Monte Carlo Control, Monte Carlo Control without Exploring Starts.	**CO1, CO2,CO4**
UNIT-4	**MONTE CARLO METHODS FOR MODEL FREE PREDICTION AND CONTROL AND TD METHODS:** **Temporal-Difference Learning:** TD Prediction, Advantages of TD Prediction Methods, Optimality of TD(0), **Sarsa:** On-policy TD control, Q-learning Off-policy TD control.	**CO1, CO2, CO4**
UNIT-5	**ELIGIBILITY TRACES AND REINFORCE** **Planning and Learning with Tabular Methods:** Models and Planning, **Dyna:** Integrated Planning, acting and learning, Prioritized Sweeping, Real-time dynamic programming, Planning at decision time, Heuristic search, Rollout algorithms, Monte carlo tree search.	**CO1, CO3, CO4**
Learning Resources		
Text Books	**1.** Richard S. Sutton and Andrew G. Barto, "Reinforcement learning: An Introduction", Second Edition, MIT Press, 2019. **2.** Russell, Stuart J., and Peter Norvig. "Artificial intelligence: a modern approach.", Pearson	
Reference Books	1. Alberto Leon-Garcia, "Probability, Statistics, and Random Processes for Electrical Engineering", 3rd Edition,	
e-Resources & other	1. https://www.udemy.com/course/artificial-intelligence-az/ 2. https://nptel.ac.in/courses/106106143 3. https://www.coursera.org/specializations/reinforcement-learning	