REINFORCEMENT LEARNING

Code: CS31122CS	Course Type: Elective	Semester: I
Evolution Scheme (TA-MSE-ESE): 20-30-100	Total Marks: 150	L-T-P (Credits): 3-0-0 (3)

Pre-Requisites: Machine Learning, Probability and Statistics, Linear Algebra, and Programming Knowledge.

Course Objective: To develop understanding of reinforcement learning foundations, sequential decision-making, advanced neural network-based function approximation, and concurrent interactive learning among multiple intelligent reinforcement learning agents in dynamic environments.

Course Outcomes: At the end of the course, the student will be able to

CO-1	Demonstrate knowledge of basic and advanced reinforcement learning techniques.
CO-2	Identify suitable learning tasks to which reinforcement learning techniques can be applied.
CO-3	Analyze the limitations and challenges of reinforcement learning techniques.
CO-4	Formulate decision problems, design computational experiments, and evaluate results obtained from reinforcement learning algorithms.

Course Articulation Matrix:

PO	PEO-1	PEO-2	PEO-3	PEO-4	PEO-5
CO-1	3	2	-	-	2
CO-2	2	3	1	-	2
CO-3	2	3	-	-	3
CO-4	3	2	2	1	3

1 - Slightly; 2 - Moderately; 3 – Substantially

Syllabus:

Unit-1: Foundations of Reinforcement Learning and Probability:

Introduction to Reinforcement Learning, Course Overview, Origin and History of Reinforcement Learning Research, Connections with Related Fields and Branches of Machine Learning, Trial-and-Error Learning Framework, Exploration–Exploitation Trade-off, Multi-Armed Bandit Problems, Action-Value Methods, Upper Confidence Bound (UCB) Algorithms, Axioms of Probability, Random Variables, Probability Mass Functions (PMF), Probability Density Functions (PDF), Cumulative Distribution Functions (CDF), Expectation, Joint and Multiple Random Variables, Joint Distributions, Conditional Distributions, Marginal Distributions, Correlation, Independence.

Unit 2: Markov Decision Processes and Dynamic Programming:

Reinforcement Learning Terminology, Markov Property, Markov Chains, Markov Reward Processes (MRP), Bellman Equations for MRPs, Existence of Solutions to Bellman Equations, Markov Decision Processes (MDP), State Value Functions, Action Value Functions, Bellman Expectation Equations, Optimal Policies, Optimal Value Functions, Bellman Optimality Equations, Planning in MDPs, Principle of Optimality, Iterative Policy Evaluation, Policy Iteration, Value Iteration.

Unit-3: Model-Free Reinforcement Learning:

Monte Carlo Methods, First-Visit and Every-Visit Monte Carlo Methods, Monte Carlo Policy Evaluation, Monte Carlo Control, On-Policy Learning, Off-Policy Learning, Importance Sampling, Temporal Difference Learning, TD(0), TD(1), TD(λ), k-Step Estimators, Unified View of Dynamic Programming, Monte Carlo and Temporal Difference Methods, TD Control Methods, SARSA, Q-Learning and Variants.

Unit-4: Function Approximation Methods:

Function Approximation in Reinforcement Learning, Risk Minimization, Gradient Descent Methods, Gradient Monte Carlo Algorithms, Semi-Gradient TD(0), Eligibility Traces, Least Squares Methods, Linear Function Approximation, Introduction to Deep Q-Networks (DQN), Prioritized Experience Replay, Policy Gradient Methods, Log-Derivative Trick, Actor–Critic Methods.

Learning Resources:

Text Books:

1.Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, 2nd Edition.

2.Alberto Leon-Garcia, Probability, Statistics, and Random Processes for Electrical Engineering, 3rd Edition.

3.Kevin P. Murphy, Machine Learning: A Probabilistic Perspective.