REINFORCEMENT LEARNING

Code: CS33222CS	Course Type: Elective	Semester: II
Evolution Scheme (TA-MSE-ESE): 20-30-100	Total Marks: 150	L-T-P (Credits): 3-0-0 (3)

Prerequisites:

Machine Learning, Probability and Statistics, Linear Algebra, and Programming Knowledge**.**

Course Objectives:

1. To understand the collection of machine learning techniques that solve sequential decision-making problems using a process of trial-and-error.
2. To understand the foundational models and algorithms used in Reinforcement Learning.
3. To familiarize students with advanced topics such as scalable function approximation using neural network representations.
4. To understand concurrent interactive learning of multiple reinforcement learning agents.

Course Outcomes:

Upon completion of the course, the student will be able to:

CO1: Demonstrate knowledge of basic and advanced reinforcement learning techniques.

CO2: Identify suitable learning tasks to which reinforcement learning techniques can be applied.

CO3: Analyze the limitations and challenges of reinforcement learning techniques.

CO4: Formulate decision problems, design computational experiments, and evaluate results obtained from reinforcement learning algorithms.

Syllabus:

UNIT I: 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, 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, Banach Fixed Point Theorem, Contraction Mapping Property, Convergence of Policy Evaluation and Value Iteration, Dynamic Programming Extensions.

UNIT 3: Model-Free Reinforcement Learning:

Monte Carlo Methods, First-Visit and Every-Visit Monte Carlo Methods, 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 DP, MC and TD 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, Afterstates, Control with Function Approximation, Least Squares Methods, Experience Replay in Deep Q-Networks. Policy Gradient Methods: Policy Gradient Methods, Log-Derivative Trick, REINFORCE Algorithm, Bias and Variance, Variance Reduction Techniques, Baselines, Advantage Functions, Actor-Critic Methods.

Suggested Reading

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*.