Deep Reinforcement Learning

Code: CS32211CS	Course Type: Core	Semester: II
Evolution Scheme (TA-MSE-ESE): 20-30-100	Total Marks: 150	L-T-P (Credits): 3-0-0 (3)

Pre-Requisites: Reinforcement Learning, Machine Learning, Deep Learning, Probability and Statistics, Linear Algebra, Python Programming.

Course Objective:

1. To understand deep neural networks–based approaches for reinforcement learning
2. To study advanced value-based and policy-based deep reinforcement learning algorithms.
3. To explore scalable learning methods for high-dimensional state and action spaces.
4. To understand multi-agent deep reinforcement learning and real-world applications.

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

CO-1	Analyze deep reinforcement learning architectures and algorithms for sequential decision-making problems.
CO-2	Implement deep value-based and policy-based reinforcement learning algorithms.
CO-3	Evaluate advanced techniques such as exploration strategies, experience replay, and stability improvements in deep RL.
CO-4	Apply deep reinforcement learning methods to complex environments such as robotics, games, and autonomous systems.

Course Articulation Matrix:

PO	PEO-1	PEO-2	PEO-3	PEO-4	PEO-5	PEO-6
CO-1	3	3	2	2	1	-
CO-2	3	3	2	3	-	1
CO-3	3	2	3	2	1	-
CO-4	3	3	3	3	1	2

1 - Slightly; 2 - Moderately; 3 – Substantially

Syllabus:

Unit-1: Foundations of Deep Reinforcement Learning

Review of Neural Networks for Reinforcement Learning, Representation Learning for Sequential Decision Making, Function Approximation using Deep Neural Networks, Stability Challenges in Deep Reinforcement Learning, Exploration vs Exploitation in Deep RL, Experience Replay Mechanisms, Target Networks, Neural Network Architectures for RL.

Unit-2: Deep Value-Based Reinforcement Learning

Deep Q-Network (DQN) Architecture, Experience Replay in DQN, Target Network Stabilization, Double Deep Q-Network (Double DQN), Dueling Network Architectures, Prioritized Experience Replay, Distributional Reinforcement Learning, Noisy Networks for Exploration.

Policy Gradient Networks, Advantage Actor-Critic (A2C), Asynchronous Advantage Actor-Critic (A3C), Deep Deterministic Policy Gradient (DDPG),

Unit-3: Model-Based and Advanced Deep Reinforcement Learning

Trust Region Policy Optimization (TRPO), Proximal Policy Optimization (PPO), Soft Actor-Critic (SAC), Entropy-Regularized Reinforcement Learning.

Model-Based Reinforcement Learning with Neural Networks, World Models, Imagination-Augmented Agents, Planning with Learned Models, Offline Reinforcement Learning, Batch Reinforcement Learning, Hierarchical Reinforcement Learning, Meta Reinforcement Learning.

Unit-4: Multi-Agent and Scalable Deep Reinforcement Learning

Multi-Agent Reinforcement Learning Architectures, Cooperative and Competitive Multi-Agent Systems, Centralized Training with Decentralized Execution, Multi-Agent Deep Q Networks, Communication Learning among Agents, Scalable RL Systems, Distributed Deep Reinforcement Learning.

Learning Resources:

2. Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2nd Edition.
3. Maxim Lapan, *Deep Reinforcement Learning Hands-On*, Packt Publishing.
4. Li, Yuxi. *Deep Reinforcement Learning: An Overview*, arXiv Tutorial.
5. Francois-Lavet, Henderson, Islam, Pineau. *An Introduction to Deep Reinforcement Learning*.
6. Silver, Sutton et al., *Mastering the Game of Go with Deep Neural Networks and Tree Search* (Research Paper).