Course code: Course Title	Course Structure			Pre-Requisite
SE407: Deep Learning	**L**	**T**	**P**	**NIL**
	3	**0**	**2**	

Course Objective: To make one understand concepts and application of reinforcement learning.

S. NO	Course Outcomes (CO)
CO1	Demonstrate deep understanding of reinforcement learning fundamentals and apply action-value methods to optimize k-armed bandit problems.
CO2	Analyze Markov Decision Processes and evaluate policies, value functions, and optimality in sequential decision-making.
CO3	Implement dynamic programming techniques for policy evaluation, policy iteration, and value iteration.
CO4	Apply and analyze temporal-difference in reinforcement learning.
CO5	Evaluate N-step bootstrapping techniques for improving reinforcement learning algorithms.

S. NO	Contents	Contact Hours
UNIT 1	**Introduction:** Elements of Reinforcement Learning, Episodic vs Continuous Tasks, The Rewards Hypothesis, Cumulative Reward, Multi-armed Bandits: A k -armed Bandit Problem, Action-value Methods, The 10-armed Testbed, Optimistic Initial Values, Gradient Bandit Algorithms.	8
UNIT 2	**Markov Decision Process:** The Agent–Environment Interface, Returns and Episodes, Episodic vs Continuous Tasks, Policies and Value Functions, Optimal Policies and Optimal Value Functions.	8
UNIT 3	**Dynamic Programming:** Policy Evaluation, Policy Improvement, Policy Iteration, Value Iteration, Asynchronous, Dynamic Programming, Generalized Policy Iteration .	8
UNIT 4	Temporal-Difference Methods, TD Prediction, Advantages of TD Prediction Methods, TD control – Sarsa, TD control- Q-Learning, TD control- Expected Sarsa, Maximization Bias and Double Learning.	9
UNIT 5	N-step Bootstrapping, N-step TD prediction, N-step Sarsa, N-step Off -policy Learning.	9
	TOTAL	**42**

REFERENCES		
S.No.	**Name of Books/Authors/Publishers**	**Year of Publication / Reprint**
1	Richard S. Sultton, Andrew G. Barto, "Reinforcement Learning", MIT Press, 2nd Edition.	**2018**
2	Marco Wiering (Editor), Martijn van Otterlo, "Reinforcement Learning: State-of-the-Art: 12 (Adaptation, Learning, and Optimization)", Springer-Verlag Berlin and Heidelberg GmbH & Co. K.	**2012**