

**Year: B. Tech IV (Semester VII)**

**Subject Name:** Reinforcement Learning

**Subject Code:** BTAI14702

**Type of course:** Professional Elective Course

**Prerequisite (if any):** Basic knowledge of machine learning

**Rationale:** Reinforcement learning is an area of machine learning in which an agent learns how to behave in an environment by the trial-and-error learning process. It has roots in operations research, behavioral psychology and Artificial Intelligence. The goal of the course is to introduce the basic mathematical foundations of reinforcement learning, as well as highlight some of the recent directions of research.

**Teaching and Examination Scheme:**

Teaching Scheme				Theory Marks			Practical Marks		Total
L	T	P	C	TEE	CA1	CA2	TEP	CA3	
3	0	0	3	60	25	15	-	-	100

CA1: Continuous Assessment (assignments / projects / open book tests / closed book tests) CA2: Sincerity in attending classes / class tests / timely submissions of assignments / self-learning attitude / solving advanced problems TEE: Term End Examination TEP: Term End Practical Exam (Performance and viva on practical skills learned in course) CA3: Regular submission of Lab work / Quality of work submitted / Active participation in lab sessions / viva on practical skills learned in course.

**Contents:**

Sr. No.	Contents	Total Hours
1.	<b>Introduction to Reinforcement Learning</b> Reinforcement Learning terminologies, Examples, Elements of Reinforcement Learning, Limitations and Scope, an Extended Example: Tic-Tac-Toe, History of Reinforcement Learning	05
2.	<b>Multi-armed Bandits</b> A k-armed Bandit Problem, Action-value Methods, The 10-armed Test bed, Incremental Implementation, Tracking a Non-stationary Problem, Optimistic Initial Values, Upper-Confidence-Bound Action Selection, Gradient Bandit Algorithms, Associative Search	09
3.	<b>Markov Decision Processes</b> The Agent–Environment Interface, Goals and Rewards, Returns and Episodes, Unified Notation for Episodic and Continuing Tasks, Policies and Value Functions, Optimal	10

	Policies and Optimal Value Functions, Optimality and Approximation	
4.	<b>Dynamic Programming</b> Policy Evaluation (Prediction), Policy Improvement, Policy Iteration, Value Iteration, Asynchronous Dynamic Programming, Generalized Policy Iteration, Efficiency of Dynamic Programming	07
5.	<b>Monte Carlo Methods</b> Monte Carlo Prediction , Monte Carlo Estimation of Action Values, Monte Carlo Control, Monte Carlo Control without Exploring Starts, policy Prediction via Importance Sampling, Off-Policy Monte Carlo Control	06
6.	<b>Temporal-Difference Learning</b> Temporal Difference (TD) Prediction, Advantages of TD Prediction Methods, Optimality of TD(0), Sarsa: On-Policy TD Control, Q-learning	08

**Suggested Specification table with Marks (Theory): (For B. Tech only)**

Distribution of Theory Marks					
R Level	U Level	A Level	N Level	E Level	C Level
15	25	15	5	0	0

Legends: R: Remembrance; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (Revised Bloom’s Taxonomy)

**Reference Books:**

Sr no	Title of book /article	Author(s)	Publisher and details like ISBN	Year of publication	Publication Edition
1.	Reinforcement Learning: An Introduction	Richard S. Sutton and Andrew G. Barto	MIT Press, Cambridge	2018	2 <sup>nd</sup> Edition
2.	Reinforcement Learning: State of the Art	Dr. Marco Wiering, Dr. Martijn Van Otterlo	Springer	2012	-

3.	Statistical reinforcement learning: modern machine learning approaches	Masashi Sugiyama	Chapman and Hall ISBN 9780367575861	2015	-
4.	Artificial Intelligence: A Modern Approach	Stuart J. Russell, Peter Norvig	PHI	2020	-

**Course Outcomes (CO):**

Sr. No.	CO statements	Marks % weightage
CO-1	Comprehend the terminologies of reinforcement learning and the basics of bandit problems	30%
CO-2	Analyze and apply policy iteration, value iteration, Markov Decision Processes and dynamic programming reinforcement learning algorithms.	40%
CO-3	Employ Monte Carlo reinforcement learning algorithms	12%
CO-4	Demonstrate temporal-difference reinforcement learning algorithms on real world problems.	18%

**List of Open learning website:**

- NPTEL online course: NOC: Reinforcement Learning <https://nptel.ac.in/courses/106106143>