GANPAT UNIVERSITY

FACULTY OF ENGINEERING & TECHNOLOGY

Programme	Master of Technology	Branch/Spec.	Computer Engineering (Artificial Intelligence)
Semester	II	Version	1.0.0.0
Effective from Academic Year	2025-26	Effective for the Batch admitted in	July 2025
Subject code	3CEAIPE107	Subject Name	Reinforcement Learning

Teaching scheme						Examination scheme (Marks)			
(Per week)	Lecture (DT)		Practical (Lab.)		Total		CE	SEE	Total
	L	TU	P	TW					
Credit	3	-	1	-	4	Theory	50	50	100
Hours	3	-	2	-	5	Practical	30	20	50

Pre-requisites

Artificial Intelligence, Probability

Course Outcomes

On successful completion of the course, the students will be able to:

CO1	Get knowledge of basic and advanced reinforcement learning techniques.
CO2	Identify suitable learning tasks to which learning techniques can be applied.
CO3	Get knowledge of some of the current limitations of reinforcement learning techniques.
CO4	Design, implement and apply novel AI techniques based on emerging real-world requirements. Model a control task in the framework of MDPs

Theory syllabus

Unit	Content	Hrs.
1	**Introduction to RL and Immediate RL:** RL framework and applications, Introduction to immediate RL, Bandit optimality, Value function-based methods	02
2	**Probability Primer:** Brush up of Probability concepts - Axioms of probability, concepts of random variables, PMF, PDFs, CDFs, Expectation, Concepts of joint and multiple random variables, joint, conditional and marginal distributions, Correlation and independence	02
3	**Bandit Algorithms:** UCB 1, Concentration bounds, UCB 1 Theorem, PAC bounds, Median elimination, Thompson sampling	06
4	**Policy Gradient Methods & Introduction to Full RL:** Policy search, REINFORCE, Contextual bandits, Full RL introduction, Returns, value functions & MDPs	05
5	**MDP Formulation, Bellman Equations & Optimality Proofs:** MDP modelling, Bellman equation, Bellman optimality equation, Cauchy sequence & Green's equation, Banach fixed point theorem, Convergence proof	06
6	**Monte Carlo Methods:** LPI convergence, Value iteration, Policy iteration, Dynamic programming, Monte Carlo, Control in Monte Carlo, Off Policy MC, UCT, TD(0), TD(0) control, Q-learning	08
7	**Eligibility Traces:** Eligibility traces, Backward view of eligibility traces, Eligibility trace control, Thompson sampling recap	05
8	**Function Approximation:** Function approximation, Linear parameterization, State aggregation methods, Function approximation & eligibility traces, LSTD & LSTDQ, LSPI & Fitted Q	07
9	**DQN, Fitted Q & Policy Gradient Approaches:** DQN & Fitted Q-iteration, Policy gradient approach, Actor critic & REINFORCE, Policy gradient with function approximation	04

Practical content

Practicals, assignments and tutorials are based on the above syllabus.

Text Books

1. Introduction to Reinforcement learning by Richard S. Sutton and Andrew G. Barto, MIT press.

Reference Books

1. Deep Learning by Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press.

2	Stochastic Approximation: A Dynamical Systems Viewpoint by V. S. Borkar, Hindustan Book Agency
ICT/MOOCs Reference	
1	https://nptel.ac.in/courses/106/106/106106143/

Mapping of CO with PO and PSO:														
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	2	1	1	1	0	1	0	1	0	1	3	1	1
CO2	3	3	1	2	1	0	2	0	2	0	1	1	2	1
CO3	3	3	1	1	1	0	1	0	2	0	1	1	1	1
CO4	3	3	3	3	3	0	3	0	2	0	2	3	3	3