SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

FACULTY OF ENGINERING AND TECHNOLOGY

SCHOOL OF COMPUTING

COURSE PLAN

18AIC304J – REINFORCEMENT LEARNING TECHNIQUES

JANUARY – MAY 2024

Revision History:

Date	Version	Modification done	Modified by	Reviewed by	Authorized by
09-01-2024	1.0	Initial Release	Dr Maivizhi R	Dr. E. Poovammal	

Table of Contents

1.0 General Details

Course Code: 18AIC304J

Course Title: Reinforcement Learning Techniques

Semester: VI

Course Time: JANUARY – MAY 2024

Slot: A

Day	Batch			
	Batch 1		Batch 2	
	Hour	Timing	Hour	Timing
Day order 1	1,2	8:00am - 9:40am	6,7	12:30pm - 2:15pm
Day order 2	10	4:00pm - 4:50pm	5	11:35am - 12:25pm
Day order 3	3	9:45am - 10:35am	8	2:20pm - 3:10pm
Day order 4	-	-	-	-
Day order 5	-	-	-	-

Location: University Building

2.0 Reference Books

1. Richard S. Sutton and Andrew G. Barto, "Reinforcement learning: An introduction", Second Edition, MIT Press, 2019
2. Algorithms for Reinforcement learning, by Csaba Szepesvari, Morgan & Claypool Publishers, 2010.
3. Probability, Statistics, and Random Processes for Electrical Engineering, 3rd Edition, Alberto Leon-Garcia, 2009.
4. "Machine Learning: A Probabilistic Perspective", Kevin P. Murphy, 2012.

3.0 Prerequisites

Nil.

4.0 Instructional Objectives

1. Introduce a range of topics related to Reinforcement Learning and probability concepts
2. Gain knowledge on Markov Decision Process
3. The dynamic programming methods of Reinforcement L
4. The Monte Carlo Prediction and Time Difference Learning
5. Function Approximation methods and Q-learning

5.0 Overall Assessment Plan

#	Component	Type	Marks
1	Cycle Test - I	Written Test	5
		Lab Exercise	5
2	Cycle Test - II	Written Test	7.5
		Lab Exercise	7.5
3	Cycle Test - III	Written Test	7.5
		Lab Exercise	7.5
4	Cycle Test - IV	Mini project	10
Total Marks			50

6.0 Tentative Test Schedule

#	Tentative date	Test		Marks	Portion	Duration
1	01-02-2024	Cycle Test – I	CLA T1	5	Unit 1	50 minutes
			CLA P1	5		
2	18-03-2024	Cycle Test – II	CLA T2	7.5	Unit 2 and 3	100 minutes
			CLA P2	7.5		
3	24-04-2024	Cycle Test – III	CLA T3	7.5	Unit 4 and 5	100 minutes
			CLA P3	7.5		
4	02-05-2024	Cycle Test - IV	CLA T4	5	All 5 units	-
			CLA P4	5	All 5 units	100 minutes

7.0 Detailed Test Plan

Theory

CIA TESTS	MARKS	TEST
Cycle Test (CLAT 1)	PART A : (MCQ) → 5 × 1 = 5 PART –B : 5 Marks →2× 5 = 10 (Descriptive) → 1 X 10 = 10 **Total Marks → 25**	**Offline mode: Written Exam**
Cycle Test (CLA T2)	PART A : (MCQ) → 10 × 1 = 10 PART B : 5 Marks →4× 5 = 20 PART –C : (Descriptive) → 2 × 10 = 20 **Total Marks → 50**	**Offline mode: Written Exam**

Cycle Test (CLA T3)	PART A : (MCQ) → 10 × 1 = 10 PART B : 5 Marks →4× 5 = 20 PART –C : (Descriptive) → 2 × 10 = 20 **Total Marks → 50**	**Offline mode: Written Exam**

Lab

CLA COMPONENTS	MARKS	RULES
CLAP-1	Lab Exercises Completion (Ex : 1 to 3) → 5 Marks **Total Marks→5 Marks**	• Students must maintain Observation Hand written in Neat and Proper Order • Observation to be signed before CLAT1 • Students will be evaluated based on the number of experiments completed and viva voce
CLAP-2	Lab Exercises Completion (Ex : 4 to 10) → 7.5 Marks **Total Marks → 7.5 Marks**	• Students must maintain Observation Hand written in Neat and Proper Order • Observation to be signed before CLAT2 • Students will be evaluated based on the number of experiments completed and viva voce
CLAP-3	Lab Exercises Completion (Ex : 11 to 15) →3 Marks **Total Marks→7.5 Marks**	• Students must maintain Observation Hand written in Neat and Proper Order • Observation to be signed before CLAT3 • Students will be evaluated based on the number of experiments completed and viva voce
CLAP-4	Model exam 5 Marks **Total Marks → 5 Marks**	• Students will be evaluated based on the Model Exam and viva voce

8.0 Detailed Session Plan

Theory

#	Topics to be covered	Hours	Ref	Teaching method	Testing method
Unit 1					
1	Introduction to Reinforcement Learning, Examples, Elements of Reinforcement Learning- Limitations and Scope	1	1	BB / PPT	Group discussion, Quiz
2	Tic-Tac-Toe example, History of Reinforcement Learning, Probability concepts, Axioms of probability	1	1	BB / PPT	Group discussion, Quiz
3	Concepts of random variables, PMF, PDFs, CDFs,	1	1	BB / PPT	Group discussion, Quiz
4	Expectation, Concepts of joint and multiple random variables	1	1	BB / PPT	Group discussion, Quiz
5	joint, conditional and marginal distributions, Correlation and independence	1	1	BB / PPT	Group discussion, Quiz
6	An-Armed Bandit Problem, Action-Value Methods	1	1	BB / PPT	Group discussion, Quiz
Unit 2					
7	Markov Decision Process, The Agent–Environment Interface, Goals and Rewards– Returns, Unified Notation for Episodic and Continuing Tasks	1	1	BB / PPT	Group discussion, Quiz
8	The Markov Property, Markov Decision Processes	1	1	BB / PPT	Group Discussion, Illustration by examples
9	Value Functions, Optimal Value Functions, Optimality and Approximation	1	1	BB / PPT	Group Discussion, Illustration by examples
10	Bellman expectation equations	1	1	BB / PPT	Group Discussion, Illustration by examples
11	Bellman optimality equations	1	1	BB / PPT	Group Discussion, Illustration by examples
12	Markov Reward Process	1	1	BB / PPT	Group Discussion, Illustration by examples

Unit 3					
13	Overview of dynamic programming for MDP, Definition and formulation of planning in MDPs	1	1	BB / PPT	Group discussion, Quiz
14	principle of optimality, Policy Evaluation	1	1	BB / PPT	Group Discussion, Illustration by examples
15	Policy Improvement, Policy Iteration, Value Iteration	1	1	BB / PPT	Group Discussion, Illustration by examples
16	Generalized Policy Iteration, Efficiency of Dynamic Programming	1	1	BB / PPT	Group Discussion, Illustration by examples
17	Banach fixed point theorem	1	1	BB / PPT	Group Discussion, Illustration by examples
18	proof of convergence of policy evaluation and value iteration algorithms	1	1	BB / PPT	Group Discussion, Illustration by examples
Unit 4					
19	Monte Carlo Prediction, Monte Carlo Estimation of Action Values, Monte Carlo Control	1	1	BB / PPT	Group discussion, Quiz
20	Off-policy Prediction via Importance Sampling, Incremental Implementation, Off-Policy Monte Carlo Control	1	1	BB / PPT	Group Discussion, Illustration by examples
21	Temporal-Difference Learning: TD Prediction	1	1	BB / PPT	Group Discussion, Illustration by examples
22	Advantages of TD Prediction Methods, Optimality of TD(0)	1	1	BB / PPT	Group Discussion, Illustration by examples
23	TD(1), TD(λ), Sarsa: On-Policy TD Control	1	1	BB / PPT	Group Discussion, Illustration by examples
24	Q-Learning: Off-Policy TD Control, unified view of DP, MC and TD evaluation methods	1	1	BB / PPT	Group Discussion, Illustration by examples
Unit 5					
25	Getting started with the function approximation methods, Revisiting risk minimization	1	1	BB / PPT	Group discussion, Quiz

26	gradient descent from Machine Learning, Gradient MC and Semi-gradient, TD(0) algorithms	1	1	BB / PPT	Group Discussion, Illustration by examples
27	Linear Methods, Eligibility trace for function approximation	1	1	BB / PPT	Group Discussion, Illustration by examples
28	Control with function Approximation, Least squares, Experience replay in deep Q-Networks	1	1	BB / PPT	Group Discussion, Illustration by examples
29	Naive REINFORCE algorithm	1	1	BB / PPT	Group Discussion, Illustration by examples
30	Bias and variance in Reinforcement Learning, Actor–Critic Methods	1	1	BB / PPT	Group Discussion, Illustration by examples

Lab:

Week no.	**Exercise**
Week 1	Installation of Code Standards and Libraries used in RL (Python/Keras/Tensorflow)
Week 2	Implement Tic-tac-toe problem
Week 3	Implement Armed Bandit Problem
Week 4	Dynamic programming algorithms for solving MDPs.
Week 5	Dynamic Programming: Policy Evaluation and Policy Iteration
Week 6	Dynamic Programming: Policy Improvement and Value Iteration
Week 7	Monte Carlo Prediction
Week 8	Monte Carlo Off-Policy Control with Importance Sampling
Week 9	SARSA (On Policy TD Learning)
Week 10	Q-Learning (Off Policy TD Learning)
Week 11	Q-Learning with Linear Function Approximation
Week 12	Deep Q-Learning for Atari Games
Week 13	Policy Gradient: REINFORCE with Baseline
Week 14	Policy Gradient: Actor Critic with Baseline
Week 15	Policy Gradient: Actor Critic with Baseline for Continuous Action Spaces

09. Overall Execution Plan:

#	Activity	Target Dates	Responsibilities	Assigned to
1	Video Content Preparation	15-02-2024	**Guidelines for video preparation:** 1. Each video should cover separate topic in each unit 2. Duration of video to be from 7 to 10 mins only. 3. Common template to be used by all. 4. Formal Dress code while recording. 5. Video should cover - Introduction about the topic, overview and concept explanation.	All faculties, Team Heads
2	Lab Program Exercises Questions Preparation	01-02-2024	1. Each faculty to prepare Maximum 5 programs for the topic assigned.	All faculties, Team Heads
3	Question Bank Preparation	13-02-2024	1. Each faculty to prepare for the respective units assigned. 2. Questions have to be framed on own and not to be taken as such from any other source. Other sources can be referred, but the question has to be modified, say with different example program, and so on. 3. Solution is required for all questions. Multiple Choice Questions - 5 Concept Understanding Questions - 2 Scenario based / HOTs Questions - 1 4. Team Heads are responsible for distributing topics to team members and no topics are missed.	All faculties, Team Heads
4	Cycle Test	22-01-2024 08-03-2024 12-04-2024	1. Select the question from Question Bank 2. Share the QP to audit professor for review 3. Plan for cycle tests question paper printing, print and distribute. 4. Coordinate with CC.	Team
5	Course File Preparation	13-03-2024 25-04-2024 15-05-2024 26-05-2024	1. Responsible for the preparation of course file as per the checklist. 2. At the end of each CT exam, files should be updated and got verified from the Team Head. 3. Participate in result analysis activity. 4. Course Files are to be prepared for each department and the faculties listed are responsible for the preparation including CO-PO Mapping, attainment of Cos, etc. 5. Coordinate with CC.	Team
6	Feedback Collection and Minutes of Meeting	17-02-2024 17-03-2024 14-04-2024 18-05-2024 30-05-2024	1. Scribe and prepare minutes of meeting for all meetings conducted. 2. Share the MoM to CC and Audit professors on the same day or the next of meeting.	Team