Course Code		Course Type	LT
CSA3003	**REINFORCEMENT LEARNING**	**Credits**	**2**

Course Objectives:

- Students with the basic concepts as well as with the state-of-the-art research literature in deep reinforcement learning.
- Students will understand the sequential decision in Machine learning.
- Students will make the formal framework for the problems in Reinforcement learning.

Course Outcomes:

At the end of the course, students should able to

- Understand the structure a reinforcement learning problem,
- Understand and apply basic RL algorithms for simple sequential decision making problems in uncertain conditions.
- Evaluate the performance of the solution
- Interpret state-of-the-art RL research and communicate their results.
- Understand the policy gradient in the Reinforcement learning

Student Outcomes (SO): a, b, c, h

a. An ability to apply the knowledge of mathematics, science and computing appropriate to the discipline
b. An ability to analyze a problem, identify and define the computing requirements appropriate to its solution.
c. An ability to design, implement and evaluate a system / computer-based system, process, component or program to meet desired needs
h. An ability to analyze the local and global impact of computing on individuals, organizations, and society

Module No.	Module Description	Hours	SO
1	**Introduction:** Course logistics and overview. Origin and history of Reinforcement Learning research. Its connections with other related fields and with different branches of machine learning.	**4**	**a, b,**
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.	**6**	**b, c**
3	**Markov Decision Process:** Introduction to Markov decision process (MDP), state and action value functions, Bellman expectation equations, optimality of value functions and policies, Bellman optimality equations.	**6**	**b, c,h**
4	**Prediction and Control by Dynamic Programing:** Overview of dynamic programing for MDP, definition and formulation of planning in MDPs, principle of optimality, iterative policy evaluation, policy	**8**	**a, c**

	iteration, value iteration, Banach fixed point theorem, proof of contraction mapping property of Bellman expectation and optimality operators, proof of convergence of policy evalutation and value iteration algorithms, DP extensions.		
5	**Policy Gradients:** Getting started with policy gradient methods, Log-derivative trick, Naive REINFORCE algorithm, bias and variance in Reinforcement Learning, Reducing variance in policy gradient estimates, baselines, advantage function, actor-critic methods. Privacy ACT - Legal Policies.	**6**	**a, b, h**
6	**Guest Lecture on Contemporary Topics**	**2**	
Total Hours		**30**	

Mode of Teaching and Learning:

Flipped Class Room, Activity Based Teaching/Learning, Digital/Computer based models, wherever possible to augment lecture for practice/tutorial and minimum 2 hours lectures by industry experts on contemporary topics

Mode of Evaluation and assessment:

The assessment and evaluation components may consist of unannounced open book examinations, quizzes, student's portfolio generation and assessment, and any other innovative assessment practices followed by faculty, in addition to the Continuous Assessment Tests and Term End Examinations.

Text Book(s):

1.	Reinforcement Learning: An Introduction", Richard S. Sutton and Andrew G. Barto, 2nd Edition
2	Probability, Statistics, and Random Processes for Electrical Engineering", 3rd Edition, Alberto Leon-Garcia

References:

1.	Machine Learning: A Probabilistic Perspective", Kevin P. Murphy
2	Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar "Foundations of Machine Learning, MIT Press, 2012.
3	Tom Mitchell, Machine Learning, McGraw Hill, 3rd Edition, 1997.
4	Charu C. Aggarwal, Data Classification Algorithms and Applications, CRC Press, 2014.

Recommendation by the Board of Studies on	24.06.2020
Approval by Academic council on:	29.06.2020
Compiled by:	Dr M Ashwin