Course Code	Course Title	L	T	P	C
1152CS164	**REINFORCEMENT LEARNING**	**3**	**0**	**0**	**3**

A. Preamble:

- To provide a clear and simple account of the key ideas and algorithms of reinforcement learning.
- To explore how the learning is valuable to achieve goals in the real world.
- To explore about how Reinforcement learning algorithms perform better and better in more ambiguous, real-life environments while choosing from an arbitrary number of possible actions, rather than from the limited options of a video game.

B. Pre-requisite

Sl. No	Course Code	Course Name
1	1151CS107	Database Management System

C. Link to Other courses

Sl. No	Course Code	Course Name
1	1152CS137	Artificial Intelligence
2	1152CS140	Machine Learning Techniques

D. Course Outcomes

CO Nos.	Course Outcomes	Level of learning domain (Based on revised Bloom's taxonomy)
CO1	Understand the need for machine learning for various problem solving	K2
CO2	Familiarize the basics of Reinforcement Learning	K2
CO3	Explain various tabular solution methods	K2
CO4	Familiarize in approximate solution methods	K2
CO5	Explain about classic conditioning and explore few applications	K2

E. Correlation of COs with POs :

Cos	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	**H**														
CO2	**H**	**H**													
CO3	**H**	**H**	**M**										**L**		
CO4	**H**	**L**	**M**										**L**		
CO5	**M**	**M**									**M**				

H- High; M-Medium; L-Low

F. Course Content

Unit-I Introduction to Machine Learning **8**

Learning Problems – Perspectives and Issues – Concept Learning – Version Spaces and Candidate Eliminations – Inductive bias – Decision Tree learning – Representation – Algorithm – Heuristic Space Search.

Unit –II Introduction to Reinforcement Learning and optimization 9

Reinforcement Learning: Introduction - Elements of Reinforcement Learning - Limitations and Scope- An Extended Example: Tic-Tac-Toe- Multi-armed Bandits: K armed, test beds, incremental implementation, Optimal initialization- Gradient Bandit, associative Search.

Unit- III Basic Tabular Solution Methods 10

Finite Markov Decision Processes- Goals, Rewards, Returns, Episodes- Optimal policies and optimal valued functions. **Dynamic Programming:** Policy Evaluation (Prediction) - Policy Improvement - Policy Iteration - Value Iteration- Asynchronous Dynamic Programming - Generalized Policy Iteration. **Monte Carlo Methods:** Monte Carlo Prediction - Monte Carlo Estimation of Action Values - Monte Carlo Control - Monte Carlo Control without Exploring Starts - Off-policy Prediction via Importance Sampling. **Temporal-Difference Learning:** TD Prediction - Advantages of TD -Incremental Implementation - Off-policy Monte Carlo Control.

Unit IV -Approximate Solution Methods 9

On-policy Prediction with Approximation : Value-function Approximation -The Prediction Objective (VE) - Stochastic-gradient and Semi-gradient Methods - Linear Methods –Feature Construction for Linear Methods- Nonlinear Function Approximation: Artificial Neural Networks - Least-Squares TD - Memory-based Function Approximation - Kernel-based Function Approximation

Unit-V Classical Conditioning & Case studies

Classical Conditioning : Blocking and Higher-order Conditioning -The Rescorla - Wagner Model - TD Model -Simulations - Instrumental Conditioning - Delayed Reinforcement- Cognitive Maps. **Case Studies:** Samuel's Checkers Player, Optimizing Memory Control, Human-level Video Game Play- Autonomous UAV Navigation and path planning -Drones for Field Coverage

Total : 45 periods

Text books:

1. Richard S.Sutton and Andrew G. Barto, “ Introduction to Reinforcement Learning”, 2nd Edition, MIT Press, 2017.
2. Tom M. Mitchell, ―Machine Learning, McGraw-Hill Education (India) Private Limited, 2013.

Reference books:

1. Sigaud O. & Buffet O. “Markov Decision Processes in Artificial Intelligence”, editors, ISTE Ld., Wiley and Sons Inc, 2010.
2. Draguna Vrabie, Kyriakos G. Vamvoudakis , Frank L. Lewis. “Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles “ 2012.

Web References:

1. B. Zhang, Z. Mao, W. Liu, and J. Liu, "Geometric reinforcement learning for path planning of uavs," Journal of Intelligent & Robotic Systems, vol. 77, no. 2, pp. 391–409, 2015.
2. Huy Xuan Pham, Hung Manh La, David Feil-Seifer, Luan Van Nguyen, "Cooperative and Distributed Reinforcement Learning of Drones for Field Coverage", arXiv:1803.07250v1 [cs.RO] ,20 Mar 2018.