Course Code	PAD21D06T	Course Name	REINFORCEMENT LEARNING FOR AI	Course Category	D	Discipline Specific Elective	L	T	P	C
							4	0	0	4

Pre-requisite Courses	*Artifical Intelligence*	Co-requisite Courses	*Nil*	Progressive Courses	*Nil*
Course Offering Department	*Computer Applications*	Data Book / Codes/Standards	*Nil*		

Course Learning Rationale (CLR):	*The purpose of learning this course is to:*
CLR-1 :	To understand the concept of algorithms helps approach problems with simple step by step solutions.
CLR-2 :	To apply creative thinking towards the problem which would help in having a solution-oriented mindset
CLR-3 :	To develop the students, have familiarity with them and stays relevant to the future modern world
CLR-4 :	To develop the abilities creates new opportunities in most business sectors and consumer applications.
CLR-5 :	To develop the decision-making knowledge.

Course Learning Outcomes (CLO):	*At the end of this course, learners will be able to:*	Learning			Program Learning Outcomes (PLO)														
		1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Disciplinary Knowledge	Critical Thinking	Problem Solving	Analytical Reasoning	Research Skills	Team Work	Scientific Reasoning	Reflective Thinking	Self-Directed Learning	Multicultural Competence	Ethical Reasoning	Community Engagement	ICT Skills	Leadership Skills	Life Long Learning
CLO-1:	Learn how to define RL tasks and the core principals behind the RL, including policies, value functions	*3*	*85*	*80*	*M*	*L*	*L*	*-*	*L*	*-*	*L*	*-*	*M*	*L*	*L*	*L*	*M*	*M*	*L*
CLO-2 :	Implement in code common algorithms following code standards and libraries used in RL	*3*	*80*	*70*	*-*	*L*	*H*	*-*	*H*	*-*	*L*	*-*	*H*	*M*	*H*	*M*	*L*	*M*	*L*
CLO-3 :	Understand and work with tabular methods to solve classical control problems	*3*	*70*	*65*	*M*	*M*	*H*	*-*	*H*	*-*	*M*	*-*	*M*	*M*	*H*	*M*	*L*	*M-*	*M*
CLO-4 :	Explore imitation learning tasks and solutions	*3*	*70*	*70*	*H*	*H*	*M*	*-*	*M*	*-*	*M*	*-*	*H*	*L*	*M*	*L*	*M*	*H*	*H*
CLO-5 :	Recognize current advanced techniques and applications in RL	*3*	*80*	*70*	*-*	*M*	*M*	*-*	*M*	*-*	*M*	*-*	*H*	*M*	*H*	*M*	*M*	*H*	*M*

Duration (hour)		12	12	12	12	12
S-1	SLO-1	***Reinforcement learning basics***	***RL formalisms and relations***	***Open AI Gym***	***Deep Q-Networks***	***Learning All possible policies with Entropy Methods***
S-2	SLO-1	***Use and applications of RL***	***Rewards in RL***	***The Random Cartpole agent***	***Deep Learning Architecture***	***Maximum Entropy RL***
S-3	SLO-1	*Reinforcement learning as MDF*	***Agents in Reinforcement learning***	***The extra Gym Functionality - Wrapper and monitors***	***Deep Q-Learning***	***Soft Actor-Critic***
S-4	SLO-1	*Learnable Functions in Reinforcement learning*	***The environnent***	***Deep Learning with PyTorch***	***Rainbow DQN***	***Extension to maximum Entropy Methods***

S-5	SLO-1	Reinforcement learning and machine learning	Actions	Tensors	Example: Rainbow DQN on Atari Games	Performance Comparison: SAC Versus PPO
S-6	SLO-1	Taxonomy of RL Approaches	Observations	Gradients	Other DQN Improvements	Industrial Example: Learning to drive with a remote control car
S-7	SLO-1	Reinforcement learning flow	Markov décisions process	NN building blocks	Policy Gradient Methods	Rethinking the MDP
	SLO-2	Deep Reinforcement learning Algorithms	Inventory Control and control simulation	Custom layers	Benefits of Learning a Policy Directly	Hierarchical RL
S-8	SLO-1	On-Policy and Off - Policy Algorithm	Markov reward process	Final glue - loss functions and optimizers	Policy Gradient Theorem	Multi- Agent RL
	SLO-2	The First RL Algorithm	Markov decision process	Monitoring with Tensor Board	n-Step Actor-Critic and Advantage Actor-Critic (A2C)	Expert Guidance
S-9	SLO-1	Compare and contrast RL and ML	Rewards Engineering	GAN on Atari images	Industrial Example: Automatically purchasing products for customers	Other Paradigms
	SLO-2	State change and transition process	Policy Evaluation: The Vale Function	The Cross-Entropy Method	Beyond Policy Gradients	The RL Project Life sysle
S-10	SLO-1	RL as a Discipline	Policy Improvement: Choosing the Best Action	Taxonomy of RL methods	Off-Policy Algorithms	Problem definition in RL
	SLO-2	Deep Learning for Reinforcement learning	Improving the e-greedy Algorithm	Cross entropy on cartpole and Frozem Lake	Deterministic policy Gradients	RL Engineering and refinement
S-11	SLO-1	Reinforcement learning and Supervised Learning	Policies and Value Functions	Theoretical background of the cross-entropy method	Trust Region Methods	Mapping policies and Action spaces
	SLO-2	Lack of an Oracle	Discounted Rewards	Tabulate Learning and the Bellman equation	Using Servos for a Real-Life Reacher	Operational RL Implementation and Deployment
S-12	SLO-1	Sparsity of Feedback	Monte Carlo Policy Generation	Value, state and optimality	Other policy Gradient Algorithms	Conclusion and the future Tips and Tricks
	SLO-2	Data Generation.	Value Iteration with Dynamic Programming	Q-Learning for FrozenLake	Extensions to policy Gradient Algorithms	The future of RL

Learning Resources	1. Deep Reinforcement Learning Hands-On - Second Edition, Maxim Lapan, January 2020 2. Reinforcement Learning, By Phil Winder, March 2020 3. Foundations of Deep Reinforcement Learning: Theory and Practice in Python, By Laura Graesser and Wah Loon Keng, December 2019.

Learning Assessment											
Level	Bloom's Level of Thinking	Continuous Learning Assessment (50% weightage)								Final Examination (50% weightage)	
		CLA – 1 (10%)		CLA – 2 (10%)		CLA – 3 (20%)		CLA – 4 (10%) #			
		Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice	Theory	Practice
Level 1	Remember Understand	*30%*	-	*30%*	-	*30%*	-	*30%*	-	*30%*	-
Level 2	Apply Analyze	*40%*	-	*40%*	-	*40%*	-	*40%*	-	*40%*	-
Level 3	Evaluate Create	*30%*	-	*30%*	-	*30%*	-	*30%*	-	*30%*	-
	Total	100 %		100 %		100 %		100 %		100 %	

CLA – 4 can be from any combination of these: Assignments, Seminars, Short Talks, Mini-Projects, Case-Studies, Self-Study, MOOCs, Certifications, Conf. Paper etc.,

Course Designers		
	Experts from Higher Technical Institutions	**Internal Experts**
	Dr.Muthu, Professor, Loyola College, Chennai	***Dr.B.Rebecca Jeyavadhanam***
	Dr.Vincent, Associate Professor, VIT	