IT565 – Reinforcement Learning

1. Credit Structure (L-T-P-Cr): 3-0-2-4

2. Course Code: TBD

3. Program/Semester: 5-xx [Btech (Sem VI MnC, ICT & ICT-CS; SEM II Mtech (ICT-ML)]

4. Category: Technical elective (TE) /Specialization elective (I/II)

5. Prerequisites: Probability, statistics & information theory/Probability & random processes

6. Foundation for: B. Tech. projects, M. Tech. major projects and M. Tech. thesis.

7. Abstract Content:

Reinforcement learning (RL) is one of the paradigms of machine learning which delves on the concept of learning by maintaining a balance between exploration and exploitation. It is different from supervised and unsupervised learning, taught in the courses on machine learning.

This course on RL aims to equip students with a computational approach to learning through interaction. Specifically, the students will be introduced to goal-directed learning and trial-and-error search with delayed reward. The notions and definitions of policy and policy function, value function and reward will be introduced. Once the theoretical formulation is done, Monte-carlo methods will be introduced for possible computation of the strategies based on RL. Students will learn the applications of concepts learned in the course on probability theory, stochastic processes and optimization in solving problems in the realm of RL.

The laboratory problems for the course will try to enhance the understanding of the concepts through several hands-on exercises. A few of the lab sessions will be reserved for case studies including games such as Tic-Tac-Toe, game of go, etc., finance, image processing and digital communication.

8. Suggested Text/s and other resources:

[1]. Reinforcement learning: An Introduction by R. S. Sutton and A. G. Barto, 2nd edition, MIT press.
[2]. Reinforcement Learning: State-of-the-Art, M. Wiering and M. van Otterlo, Springer.
[3]. Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions by W. Powell, Wiley.
[4]. Reinforcement Learning and Optimal Control by D. Bertsekas, Athena Scientific
[5]. Deep Reinforcement Learning Hands-On, M. Lapan, Packt publication
[6]. Distributional Reinforcement Learning by M. Bellemare et. al. MIT press.
[7]. Deep Reinforcement Learning in Action by Zai and Brown, Manning Public.
[8]. David Silver's course on Reinforcement Learning, available at https://www.davidsilver.uk/teaching/

[9]. Balaraman Ravindran's course on Reinforcement Learning, available at https://nptel.ac.in/courses/106106143

10. Detailed Contents:

Topic Name	***Content (2 -3 lines per 4 – 6 lectures)***	***No of lectures***
A. Introduction to reinforcement learning (RL)	*RL vs machine learning, applications of RL, elements of RL, a few examples of RL.*	***2***
B. Markov decision processes	*Formulation of agent and environment using the Markov model, concept and usage of of goal, rewards, returns and episodes in problem formulation and solution, definition of policy and value functions and their conceptual understanding, derivation of optimal policies and optimal value functions.*	***7***
C. Dynamic programming	*Policy evaluation, policy improvement, iteration of policy and value, asynchronous dynamic programming.*	***7***
D. Monte-carlo (MC) methods	*MC prediction, estimation and control, importance sampling.*	***8***
E. Temporal difference (TD) learning	*TD prediction methods, TD(lambda), advantages and optimality of TD methods, Sarsa, expected sarsa, Q learning.*	***8***
F. Approximate solution methods:	*model free prediction and control, policy approximation, policy gradient, actor-critic methods.*	***8***

11. Evaluation:

In semester 1 exam: 22%
In semester 2 exam: 23%
End semester exam: 25%
Laboratory exercises: 30%

There is no explicit project component. However, laboratory exercises will be sufficient to develop a project based learning.

12. Course Outcome:

1. Familiarity with both fundamental and sophisticated reinforcement learning methods.
2. Determining the appropriate learning assignments for which these strategies can be used.

3. Formulation of decision problems, set up and run computational experiments, evaluation of results from experiments.

13. POs-COs Matrix:

PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
■	■	■	■	■							■