****

**FIRST SEMESTER 2021-2022**

# Course Handout Part II

Date: 20-08-2021

In addition to part-I (General Handout for all courses appended to the time table) this portion gives further specific details regarding the course.

*Course No. : CS F317*

## *Course Title : Reinforcement Learning*

## *Instructor-in-Charge : Dr. Paresh Saxena*

### Scope and Objectives:

Reinforcement Learning (RL) is based on the idea where we learn by interacting with our environment. It is a collection of machine learning techniques where agent(s) learn how to behave in an environment by performing actions and assessing the results. RL approach is building programs that learn how to predict and act in stochastic environment, based on past experience. This course will provide an overview to students on some of the fundamental ideas on which modern RL is built including markov decision processes, value functions, monte carlo estimation, dynamic programming, TLD methods, approximation methods, Actor-Critic methods, etc. This course will help students to understand and apply RL in several systems including video distribution systems, game development, IOT devices, robotics, clinical decision making, industrial process control, finance portfolio balancing, etc.

This subject aims to achieve the following goals:

* To provide students with the knowledge to structure a reinforcement learning problem.
* To introduce students to learn and apply basic RL algorithms for simple sequential decision-making problems in uncertain conditions
* To introduce students research and development work in reinforcement learning by interpreting state-of-the-art RL research and communicating their results.
* To provide knowledge to students to build a RL system that knows how to make automated decisions.
* To give students opportunities to understand the space of RL algorithms including Temporal Difference Learning, Monte Carlo, Q-Learning, approximation solution methods, A2C, A3C, etc.

##### Textbooks:

* Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. Second Edition, MIT press, 2018.

**Reference Books:**

* Wiering, Marco A., and Martijn Van Otterlo. "Reinforcement learning." *Adaptation, learning, and optimization* 12.3 (2012).

**Course Plan:**

|  |  |  |  |
| --- | --- | --- | --- |
| Lecture No | Learning Outcomes | Topics to be covered | Chapter in the Text Book |
|  | | | |
| 1-2 | To Introduce Reinforcement Learning, its limitations and scope. | Introduction to Reinforcement Learning. A brief summary of the pre-requisites, mathematical and programming tools required. | Class notes |
| PART 1: Tabular Solution Methods | | | |
| 3-4 | To understand the application of RL for single-state Bandit problems. | Bandit problems | T1:Ch2 |
| 5-7 | To introduce methods for solving finite markov decision problems | Finite Markov Decision Processes | T1:Ch3 |
| 8-9 | To understand the utility of DP (require perfect model) in RL | Solution Methods: Dynamic Programming | T1: Ch4 |
| 10-12 | To understand and apply Monte Carlo Methods (require only experience) for RL | Solution Methods: Monte Carlo Methods | T1:Ch5 |
| 13-18 | To understand and apply the application of Temporal Difference Learning including Q-learning and Sarsa for RL | Solution Methods: Temporal-Difference Learning | T1:Ch6, Ch7 and Class notes |
| 19-22 | To develop a unified view of methods that require a model of the environment and methods that can be used without a model. | Planning and Learning with Tabular Methods | T1:Ch8 |
| PART 2: Approximate Solution Methods | | | |
| 23-28 | To apply RL for applications where most states encountered will never have been experienced exactly before. To apply on-policy and off-policy approximation of action values. | Function approximation for generalization | T1:Ch9-Ch11 and Class Notes |
| 29-34 | To understand and learn the reinforcement learning in a multiagent setting | Multi-agent reinforcement learning | Class Notes |
| 35-40 | To understand the application of A2C and A3C (with multiple independent agents) | Actor-Critic (A2C, A3C) | Class Notes |

##### Evaluation Scheme:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Component** | **Duration (Minutes)** | **Weightage** | Date&Time | **Nature of Component** |
| Midsemester Test | 90 Mins | 35% | 19/10/2021 1.30 - 3.00PM | Open Book |
| Course Project (with final presentation/viva) | - | 25%  **(5% will be evaluated before the mid-sem)** | Details will be announced during the lecture sessions. | Open Book |
| Comprehensive Examination | 120 Mins | 40% | 15/12 AN | Closed Book |

**Consultation Hour:** Online consultation hours will be announced in the class.

**Notices:** All notices pertaining to this course will be displayed on the CMS.

**Make-up Policy:** The make-up policy will be in accordance with the AUGSD guidelines.

**Academic Honesty and Integrity Policy**: Academic honesty and integrity are to be maintained by all the students throughout the semester and no type of academic dishonesty is acceptable.

**Instructor-in-charge**