Department of Computer Science
Ramakrishna Mission Vivekananda Educational and Research Institute

Reinforcement Learning

(August - December 2023)

Instructor: Swami Vidyapradananda
Office: MB 113 and AV 106, email: vidyapradananda@gm.rkmvu.ac.in

Course Overview: Reinforcement Learning (RL) is a branch of machine learning that deals with sequential decision-making. RL, however is different from other machine learning paradigms, since there is no supervisor for training. Instead, there is a trial-and-error learning process, that involves reward, feedback from environment, action, and time of action. The agent interacts with a dynamic, stochastic, and incompletely known environment, with the goal of finding an action-selection strategy, or policy, to optimize some measure of its long-term performance. RL has a wide range of applications in many problem situations, such as robotics, healthcare, smart grids, finance, self-driving cars etc., where explicit instructive signals are not available.The goal of the course is to introduce the mathematical foundations of reinforcement learning drawing from Operations Research, and do hands-on project on some small scale recent problems.

Prerequisite(s): Optimization Techniques, Matrix Algebra, Python

Credit Hours: 4

Course Objectives: Reinforcement Learning (RL) is a general purpose formalism for automated decision-making where an agent explicitly takes actions and interacts with the environment. Understanding the importance and challenges of learning agents that make decisions is of vital importance in designing interactive AI agents. This course focuses on various important aspects of RL, viz., formulating problems as Markov Decision Processes, understanding the basic exploration methods and the exploration/exploitation tradeoff, understanding value functions and how to implement dynamic programming as an efficient solution approach, identifying the impact of choices on performance and validating the expected behaviour of algorithms

Knowledge Gained: Students will gain the following:

- Understanding of how RL relates to and fits under the broader umbrella of machine learning, deep learning, supervised and unsupervised learning.
- Understanding of the space of RL algorithms (Temporal- Difference learning, Monte Carlo, Q-learning, Policy Gradient etc.).
- Understanding of how to implement reinforcement learning in RL platforms such as, **TensorFlow/Open AI Gym, PyTorch** etc.

Text Book:

The course material will be drawn from multiple book chapters, journal articles, reviewed tutorials etc. However, the following books are some good recommended texts for this subject.

- *Reinforcement Learning: An Introduction*, Richard S. Sutton and Andrew G. Barto, 2nd Edition, MIT Press, Cambridge, MA, 2018
- *Reinforcement Learning: State-of-the-Art*, Marco Wiering and Martijn van Otterlo, Eds. Springer-Verlag Berlin Heidelberg 2012
- *Artificial Intelligence: A Modern Approach*, Stuart J. Russell and Peter Norvig. Fourth edition, 2020
- *Deep Reinforcement Learning: Fundamentals, Research and Applications* Hao Dong, Zihan Ding, Shanghang Zhang Ed.s © Springer Nature Singapore Pte Ltd. 2020
- *Deep Reinforcement Learning: Frontiers of Artificial Intelligence*, Mohit Sewak © Springer Nature Singapore Pte Ltd. 2019

Table 1: **Topics (for 16 weeks)**

A. Preliminaries –*1 week*

- The Reinforcement Learning (RL) framework, RL vis-a-vis supervised/unsupervised ML
- Interesting applications of RL

B. Mathematical and Algorithmic Foundations of RL (4 weeks) – *The following topics will be covered in detail*

- *Sequential decision making*: Dynamic Programming Problem, Bellman Optimality Equations
- *Markov Decision Process* (MDP): State, stage, action, total expected reward criteria, average reward criteria. Optimal policies and optimal value functions. Value iteration and Policy Iteration. Optimal solution vs. approximate solution of Bellman optimality equation.
- *Partially Observable Markov Decision Process* (POMDP): The observation model, Monahan's enumeration algorithm, pruning via Linear Programming, the Witness algorithm, Parsimonious representation of value functions, near-optimal solutions for POMDPs

C. Methods of RL (8-10 weeks) – *The following topics will be covered in detail*

- *Model based methods*: Value iteration, Policy Iteration, Conditions on convergence Exploration/Exploitation Tradeoff (the k-armed bandit problem)
- *Model free methods*: Watkins' Q-learning,Sutton's Dyna architecture, Monte Carlo and Gaussian Process Temporal Difference Learning
- *Approximate solution methods*: Policy gradient, Monte Carlo policy gradient, Actor-critic algorithm, Performance evaluation of policy gradient methods

D. Frontiers in RL (1 week)- *If time permits, an overview will be given on the following topics:*

- Hierarchical Reinforcement Learning
- Game Theory and Multi-agent Reinforcement Learning
- Decentralised POMDP

Course Project

- There will be a course project, aimed at implementing reinforcement learning theory and algorithms to practical applications. It will also give students opportunity to learn how to deploy and train reinforcement learning on AI frameworks (e.g., PyTorch, OpenAI Gym, Tensorflow). The topic would be finalised based upon detailed discussion with the Instructor. An intermediate presentation of the project topic has to be done after midterm exam. Students can possibly extend and expand on the course project later on and give it a more concrete shape leading to term project and/or masters thesis.

Tutorial on RL Platforms

Students are required to code in Python for the assignments in the course. Python codes of some important algorithms, such as value iteration, policy iteration, Q-learning etc. will be taught. Students are encouraged to go through tutorials on PyTorch, OpenAI Gym, Tensorflow. Some tutorial resources are:

- **PyTorch**: https://pytorch.org/tutorials/
- **TensorFlow**: https://www.tensorflow.org/agents/tutorials/

Two books which give hands-on tutorial are

- *Applied Reinforcement Learning* (with Python, OpenAI Gym, Tensorflow, and Keras), by Taweh Beysolow II. Apress, Berkeley, CA.
- *Hands-On Reinforcement Learning with Python*, by Sudharsan Ravichandiran © 2018 Packt Publishing

Course Policies:

- **Modality**

  This course will be conducted through LMS. Lecture notes, assignments etc will be posted in LMS Students should visit course page on LMS regularly.

- **Academic Integrity**

  1. Students are expected to work independently in the Assignments. Discussion amongst students is encouraged but offering solution to others as well as accepting solutions from others both are act of academic dishonesty. In such cases students will be penalized according to the *Academic Honesty Policy* of University.

- **Assignments/Quizes**

  1. Except under medical reasons late assignments will **not be accepted**. Late assignments will be marked down by 10%
  2. Surprise quizes may be conducted to check attentiveness, attention and attendance in the online class mode.

- **Attendance and Absence**

  Students are not supposed to miss class without strong reasons such as medical needs or emergency situations. Students must be aware that, except these cases, if a student falls short of attendance requirement specified by the Controller of Examination/Head of the Department then he may not be allowed to appear in the final examination. **A minimum of 80% attendance is required in order to appear for the finals.** Students are responsible for all missed work, regardless of the reason for absence. It is also the absentee's responsibility to get all missing notes or materials.

Grade Distribution:

Assignments	20%
Midterm Exam	20%
End term Exam	40%
Final Project	20%

Grading Policy: $> 95 : A+, 75 - 95 : A, 60 - 75 : B, 45 - 60 : C, 30 - 45 : D, < 30 : F$