Course Plan – 2025

Name: Introduction to Reinforcement Learning

Duration - (50 hours - 25 Sessions + 3-4 assignments)

Performance analysis – Assignments (Numericals / short coding) – 90%, 10% attendance

Objectives: Animals learn to select actions based on interaction with the environment around them. Reinforcement Learning (RL) is a field learning theory, where we mimic animals using computer algorithms. Here, an RL agent takes an action based on the current environment and receives a reward. Through the reward, it estimates its optimal action to maximise rewards/increase chances of survival.

This 50–hour comprehensive course introduces the students to the foundations of Reinforcement Learning. It is targeted at students who want to learn the foundations of AI – particularly Reinforcement Learning. This is tailored for all UG students in Engineering and Sciences. It would be particularly helpful for students who want to pursue -

(i) a career in research in AI/RL/statistics
(ii) Or apply for RL positions roles in the industry (a lot of robots like Boston Robotics use RL to design their robots).

Schedule: [Start date: 1st week of March – April (pre-semester 20 hours), and End of May-June (post-semester 30 hours)]

Prerequisites: Probability and statistics, especially conditional distributions and expectations, and basics of linear algebra. You can get these contents in my previous course – Introduction to ML for GATE DA.

Phase 1: Basic RL (20 hours; 10 sessions)

(a) Recap – Conditional probability, Conditional Expectation, Vector Spaces and Linear combination *(4 hours)*

(b) What is RL? Environment, Agent, reward, action, Multiarmed Bandits, exploration vs Exploitation, UCB algorithm, Gradient Methods, Python coding *(6 hours)*

(c) Policy, action and value functions, Markov Decision Process, Bellman Equations, Optimal Policy, Policy Evaluation, Policy Iteration and Improvement, Value Iteration, Dynamic Programming, Python Coding (10 hours)

Assignment-cum-test 1 – 30%

Phase 2: Advanced RL algorithms (30 hours; 15 sessions)

(a) Monte Carlo Methods for estimating value functions. On-Policy, Off-Policy algorithms using Importance weighting, Temporal Difference (TD) Methods, SARSA, and Q-learning algorithms, and Python coding *(20 hours)*

Assignment-cum-test 2 – 40%

(b) Function Approximation-based RL methods – Tabular vs parameterized function-based algorithms, linear approximation methods for TD/Q learning, Deep RL, Intro to Actor-Critic algorithms *(10 hours)*

Assignment-cum-test 3 – 20%

Reference Books:

0) Session handwritten notes, and assignments
1) Introduction to RL, Sutton and Barto
2) Optimal Control and Dynamic Programming, Dimitri Bertsekas (for advanced proofs)

Instructor – Mainak Biswas (https://mainak-biswas1999.github.io/)
4th year PhD student, Cognition Lab, IISc.
Department of Brain, Computation and Data Science, IISc.