Course Code: CS4122
Course Name: Reinforcement Learning
Credits: 3(3-0-0)

Course Outcomes

After successful completion of this course, students will be able to:

CO1:Tell how reinforcement learning is different from other AI paradigms like Machine/Deep learning.
CO2:Conceptually use exploration-exploitation tradeoffs to tune several reinforcement learning algorithms.
CO3:Tell the difference between bandits and complete reinforcement learning setups.
CO4:Tell the difference between algorithms that just plans (like in Markov Decision Process) and those that both learns and plans.
CO5: Given a real-world problem, how to decide whether it can be solved using reinforcement learning approach. If it can be solved using reinforcement learning then how to formally define it as a reinforcement learning setup.
CO6: Tell the working of common reinforcement learning algorithms and also code them from scratch in Python. This will later help students to modify existing algorithms and/or develop new algorithms in the future.
CO7: Use existing libraries that supports reinforcement learning like Open AI Gym and Stable Baselines.

Course Content

Introduction and Multi Armed Bandit: Differences between Multi-Armed Bandits (MAB), Markov Decision Process (MDP), and Reinforcement Learning (RL); Applications of MAB, MDP, RL; Exploration and exploitation tradeoff.

Multi-Armed Bandits, Application of MAB, MAB algorithms: ε-greedy and Upper Confidence Bounds, Contextual bandits, Bayesian Bandits, Introduction to OpenAI Gym, Implementing Bandit algorithms using OpenAI Gym and Python.

Markov Decision Process: Brief Introduction to Markov Process and Markov Reward Process, Episodes, Trajectories, formulating real-world problems as MDP; Policy evaluation, Bellman Optimality equation and Dynamic Programming; computing optimal policies using Value and Policy Iteration, Generalized Policy Iteration, Implementing Value and Policy iteration in Python.

Reinforcement Learning : Taxonomy of RL algorithms; Model based and Model free, Value based and Policy based, On-Policy and Off-Policy, Monte Carlo Methods; Temporal Difference Methods, SARSA, Q-Learning, and Double Q-Learning, Implementing Monte Carlo, SARSA, and Q- Learning in Python, Implementing Custom Open AI Gym environment.

Deep Reinforcement Learning: Motivation behind Deep RL, Deep Q-Learning: Replay memory, target network. Policy Gradient: Variance reduction and actor critic method; Proximal Policy Optimization, Implementing Deep Q-Learning and Policy Gradient using Python, Inverse Reinforcement Learning.

Localization and Mapping: Brief recap of Bayesian Filters, Localization, Mapping, Simultaneous localization and mapping (SLAM), depth from monocular cameras, traditional and deep learning based SLAM, motion planning, and exploration.

TextBook

Reinforcement Learning: An Introduction, Sutton and Barto, 2nd Edition.

References

1. Artificial Intelligence: A Modern Approach, Stuart J. Russell and Peter Norvig.
2. Deep Reinforcement Learning Hands On, Maxim Lapan, 2nd Edition. [ISBN 978-1-83882-699-4].