



Reinforcement Learning Based Systems

Course Code:	28656
Course Type:	Theoretical
Credits:	3
Course Status:	Elective Specialized Graduate Course
Prerequisite:	Supervisor Approval

Aim/Scope/Objectives: This course intends to give the required knowledge to the graduate students for designing intelligent systems based on reinforcement learning, in which the system learns by the interaction between its agents and the environment. This course cover all required bases and the students train to implement a research project by Python. This course has extensive applications in robotic, biomechanics, control and decision making in complex systems.

Course Outline:

1- An overview on Learning Based Systems

- 1-1- Introducing different learning based systems
- 1-2- Basic elements of a reinforcement learning system
- 1-3- Examples for reinforcement learning systems

2- The Official form of Reinforcement Learning Problem

- 2-1- Different structures of intelligent agents
- 2-2- Environment characteristics
- 2-3- Markov decision feature in reinforcement learning problems
- 2-4- Value functions
- 2-5- Optimal policy and Bellman optimality equations

3- Methods of Action Selection

- 3-1- Approximation of action-value
- 3-2- Direct and search methods
- 3-3- Improvement of Softmax search method
- 3-4- Upper confidence bound (UCB) action selection

4- Dynamic Programing

- 4-1- Policy evaluation
- 4-2- Policy improvement
- 4-3- Policy iteration
- 4-4- Value iteration
- 4-5- Performance of dynamic programing

5- Monte Carlo Methods

- 5-1- Monte Carlo value function prediction
- 5-2- Monte Carlo action-value function prediction
- 5-3- Monte Carlo control
- 5-4- On-policy Monte Carlo control
- 5-5- Generalized Policy Iteration (GPI)



- 5-6- Off-policy Monte Carlo control
- 6- Temporal-Difference Learning**
 - 6-1- TD prediction
 - 6-2- Batch updating in TD
 - 6-3- Control based on TD
 - 6-3-1 Sarsa learning algorithm
 - 6-3-2 Q-learning learning algorithm
 - 6-3-3 Expected Sarsa learning algorithm
 - 6-4- Evaluation methods of reinforcement learning algorithms
- 7- n- Step Temporal-Difference Learning and Eligibility Trace**
 - 7-1- n- Step bootstrapping methods
 - 7-1-1 n-Step TD prediction
 - 7-1-2 n-Step Sarsa learning algorithm
 - 7-2- Eligibility trace
 - 7-2-1 $TD(\lambda)$ prediction
 - 7-2-2 Sarsa (λ) learning algorithm
- 8- Advance Topics in Reinforcement Learning**
 - 8-1- Generalization in Reinforcement Learning
 - 8-2- Deep Reinforcement Learning and DQN
 - 8-3- Policy Gradient Methods
 - 8-4- Actor-Critic Methods
 - 8-5- Imitation Learning
 - 8-6- Maximum Entropy Reinforcement Learning
- 9- Different Structures of the Actor-Critic Networks for Control Applications**
 - 9-1- Deep Deterministic Policy Gradient (DDPG)
 - 9-2- Twin Delayed Deep Deterministic Policy Gradient (TD3)
 - 9-3- Soft Actor Critic (SAC)

Grading: 50% Final exam, 10% Homework, 40% Research Project by the Python

References:

- 1- Richard S. Sutton, Andrew G. Barto, “Reinforcement Learning: An Introduction”, second edition, MIT press, 2018.
- 2- Hao Dong, Zihan Ding, Shanghang Zhang, “Deep Reinforcement Learning, Fundamentals, Research and Applications, Springer, 2020.
- 3- Maxim Lapan, “Deep Reinforcement Learning Hands-On”, Second edition, Packt publishing, 2020.
- 4- Sudharsan Ravichandiran, “Hands on Reinforcement Learning with Python”, Packt publishing, 2018.
- 5- Andrea Lonza, “Reinforcement Learning Algorithms with Python”, Packt publishing, 2019.
- 6- Farrukh Akhtar, “Practical Reinforcement Learning, Develop Self Evolving, Intelligent Agents with Open AI Gym, Python and Java”, Packt publishing, 2017.
- 7- Mohit Sewak, “Deep Reinforcement Learning, Frontiers of Artificial intelligence”, Springer, 2019.