

DA-IICT, IT 567, Winter 2024-2025
Lab Exercise 2
Date: 16/01/2025, Expected by: 24/01/2025
Prepared by: Dr. Abhishek Jindal

References for perusal:

- [1] *Deep Reinforcement Learning Hands-On*, 2nd edition, M. Lapan, Packt Publishing Ltd., 2020.
- [2] *Reinforcement Learning Algorithms with Python*, A. Lonza, Packt Publishing Ltd., 2019.
- [3] *Reinforcement Learning: An Introduction*, 2nd edition, Richard S. Sutton, and Andrew G. Barto, The MIT Press, 2018.

References (online) for perusal:

- {a} *Policy and Value Iteration*, CIS 522 - Deep Learning, available on: <https://www.youtube.com/watch?v=187rgLg90HI>
- {b} *Policy Evaluation in GridWorld*, E. Sayar, available on: <https://www.youtube.com/watch?v=jt3qygQFt7Q>
- {c} *How to use Bellman Equation Reinforcement Learning*, M. Huddar, available on: <https://www.youtube.com/watch?v=3Rx2x2traxw>

Disclaimer: For the video links above, I, or DA-IICT, do not endorse any of the online learning platforms and the video creators. Links are provided since the content is freely available and I assume that it might help in learning.

- The lab exercise comprises of 4 problems. First two are analytical and the remaining two are to be done using Python.
 - All the problems need to be solved individually. However, collaboration is encouraged for building an initial understanding of the problem, and subsequently, for creating the solution strategy.
 - All the required soft copies of the texts referred to in the exercises are available in the lecture folder of the instructor.
-

1. From [3], solve exercises 3.7, 3.8 and 3.9.
2. Go through example 3.4. Then observe cartpole example in OpenAI Gym. For the same, refer to pgs. 28 - 32 from [2]. For an additional reference, you can see pgs. 37 - 40 from [1]. Subsequently, solve exercise 3.6.
3. From [3], read example 3.5 (Gridworld). Obtain the value function v_π for the policy of equal probability for all four actions in all states, i.e., reproduce (using Python code) the matrix shown in the right of figure 3.2. Before solving for v_π matrix, you need to refer to {b}. The solution to this problem will help you understand the application of Bellman equation.
4. From [3], based on the knowledge gained from solving the above problem, solve exercises 3.14, 3.15 and 3.16.
5. **[Study exercise]** Go through {a} and {c} for building an initial understanding of policy and value iteration. We will discuss it in subsequent lectures.

Instructions for Preparing Lab Report:

- For analytical problems, either create a digital copy of your solution (through latex/Word/etc.), or write it on a paper and scan, whichever choice you make, send it to the instructor through email.
- For problems solved in simulations environments (python/Visual studio/etc.), send the code file along with the snapshot of the result.

General Instructions:

- The lab is intentionally made from the references given above so that you have ample resources to refer to and learn.
- For the final evaluation, we may have a quiz/lab test which will test if you have read through the references and solved the problems yourself.