

XAI613 Fall 2024
Assignment 1
Due: September 30 at 9:00 am (KST)

1 Frozen Lake MDP [25 pts]

You will implement value iteration and policy iteration for the Frozen Lake environment from [OpenAI Gym](#). We have provided custom versions of this environment in the starter code.

- (a) **(coding)** Read through `vi_and_pi.py` and implement `policy_evaluation`, `policy_improvement` and `policy_iteration`. The stopping tolerance (defined as $\max_s |V_{old}(s) - V_{new}(s)|$) is $\text{tol} = 10^{-3}$. Use $\gamma = 0.9$. Return the optimal value function and the optimal policy. [10pts]
- (b) **(coding)** Implement `value_iteration` in `vi_and_pi.py`. The stopping tolerance is $\text{tol} = 10^{-3}$. Use $\gamma = 0.9$. Return the optimal value function and the optimal policy. [10 pts]
- (c) **(written)** Run both methods on the Deterministic-4x4-FrozenLake-v0 and Stochastic-4x4-FrozenLake-v0 environments. In the second environment, the dynamics of the world are stochastic. How does stochasticity affect the number of iterations required, and the resulting policy? [5 pts]

How to submit For the written question write your answers with any software that can export a pdf file. Zip `vi_and_pi.py` file and your pdf file into the zip file named `StudentID_YourName.zip`. Submit the zip file through Blackboard.