

Assignment 4: Reinforcement Learning

CS486/686 – Winter 2026

Out: March 9, 2026

Due: March 20 (11:59 pm), 2026.

Submit an electronic copy of your assignment via LEARN. Late submissions incur a 2% penalty for every rounded up hour past the deadline. For example, an assignment submitted 5 hours and 15 min late will receive a penalty of $\text{ceiling}(5.25) * 2\% = 12\%$.

This assignment covers Lectures 15-17.

Submit 3 files that correspond to the 3 python notebooks with all functions filled in, as well as graphs and answers to questions included in each python notebook.

1. Part 1 (MDP.ipynb) [30%]: fill in the functions for value iteration (10%) and extract policy (10%). Run your code on a simple maze problem described in the python notebook to display the number of iterations, the value function and policy found (10%).
2. Part 2 (RL.ipynb) [40%]: fill in the functions for q-learning and model-based RL with value iteration. Run your code on the same maze problem as Part 1. Follow the instructions in the python notebook to produce graphs that show a) the impact of epsilon on q-learning (10%) and b) the performance of model-based RL with value iteration (10%). Answer the questions described in the notebook for each graph (20%).
3. Part 3 (DQN.zip) [30%]: modify the code to produce graphs that show the impact of a) the frequency of update of the target network (5%) and b) the batch size on performance (5%). Answer the questions described in the notebook for each graph (20%).