Impact of Parameter Choices in RL

Graded Assignment • 40 min

← Back

English Due Jun 27, 11:59 PM +07

Your grade: 83.33%

Your latest: **83.33%** • Your highest: **83.33%**

To pass you need at least 80%. We keep your highest score.

Next item →

1. Which of the following meta-parameters can be tuned to improve performance of the agent? Performance refers to the cumulative reward the agent would receive *in expectation* across different runs. (Select all that apply) 1 point

- [x] Exploration parameter (e.g., epsilon in e-greedy or the temperature tau in the softmax policy)

  Correct
  Correct. We have to try different levels of exploration that the agent begins with, because different problems may require different extents of exploration. We do not know this beforehand.

- [x] The step size in the update rule of the learning algorithm (e.g., alpha in Q-learning)

  Correct
  Correct. If the step size is too low, learning might be very slow. But if it is too high, there might be a lot of variance in the learning behaviour.

- [x] Random seed (for the random number generator)

  This should not be selected
  Incorrect. Good scientific practice involves trying the same configuration of meta-parameters of an algorithm over a number of different random seeds for drawing statistically-sound inferences. The random seed is not a meta-parameter to be tuned for 'good performance'.

- [x] Number of hidden-layer units in a neural network approximating the value function

  Correct
  Correct. If the number of hidden units is too small, the representational capacity may be insufficient for learning good behavioural policies. On the other hand, a large number of hidden units could help to learn a good representation, but learning progress might be very slow due to the sheer number of parameters.

2. Suppose a problem that you have formulated as an MDP has k continuous input dimensions. You are considering using tile coding as a function approximator. With T tilings and t tiles per dimension in each tiling, which of the following represent the resultant number of features? (Assume each tiling covers all k dimensions.) 1 / 1 point

- (•) T·t^k
- ( ) T · t / k
- ( ) k·T^t