

CS 4320: Machine Learning

Assignment: Temporal Difference Q-Function (Reinforcement Learning)

Train a reinforcement agent to perform in the `MountainCar-v0` environment.

It is expected that you will use the `CartPole` example code as a starting point for your code development.

Train a Q-Function agent to obtain the highest score on 20 scoring epochs.

Consider hyper parameters such as gamma, epsilon-chance-factor, learning policies, neural network architecture, modified training rewards, serial training sessions, and any other potentially useful modifications to the Q-Function or training process.

Create a report that includes:

- A description of the states, actions, and rewards of the `MountainCar-v0` environment.
- An outline of code changes needed to work with the `MountainCar-v0` environment.
- A description of modifications attempted in the Q-Function architecture and learning process.
- A description of the best Q-Function obtained, the process used to train it, and the score obtained. This should include number of training sessions, number of epochs per session, hyper parameters used in each session, etc.
- A discussion of the effect of Q-Function and learning modifications attempted.

Required Steps

- Download the starter code.
- Verify it runs for the `CartPole` environment.
- Modify the code to run with the `MountainCar` environment.
- Attempt to train many agents with various combinations.
- Create a report with the contents mentioned above.
- Commit and push your code in the git repository.
- Submit the report (as PDF) to Canvas.