Neural Networks and Deep Learning
Academic year 2021-22

Homework 3: Deep Reinforcement Learning

General overview: In this homework you will learn how to implement and test neural network models for solving reinforcement learning problems. The basic tasks for the homework will require to implement some extensions to the code that you have seen in the Lab. The advanced tasks will require to train and test your learning agent on a different type of input (image pixels) or Gym environment. You can just choose one of the advanced tasks to get the maximum grade. If you are interested in improving your skills, feel free to try both advanced tasks. Given the higher computational complexity of RL, in this homework you don't need to tune learning hyperparameters using search procedures and cross-validation; however, you are encouraged to play with model hyperparameters to find a satisfactory configuration.

Technical notes: The homework should be implemented in Python using the PyTorch framework. The student can explore additional libraries and tools to implement the models; however, please make sure you understand the code you are writing because during the exam you might receive specific questions related to your implementation. The entire source code required to run the homework must be uploaded as a compressed archive in a Moodle section dedicated to the homework. If your code will be entirely included in a single Python notebook, just upload the notebook file.

Final report: Along with the source code, you must separately upload a PDF file containing a brief report of your homework. The report should include a brief Introduction on which you explain the homework goals and the main implementation strategies you choose, a brief Method section where you describe your model architectures and hyperparameters, and a Result section where you present the simulation results. Total length must not exceed 6 pages, though you can include additional tables and figures in a final Appendix (optional). The report must be self-contained, so please make sure that all relevant results are explicitly included in the report / notebook. Given the dynamical nature of RL problems, you can explore more sophisticated media for showing the results of your model (e.g., animated GIFs or short movies).

Grade: The maximum grade for this homework will be **8 points**. Points will be assigned based on the correct implementation of the following items:

- 3 pt: use the notebook of Lab 07 to study how the exploration profile (either using eps-greedy or softmax) impacts the learning curve. Tune a bit the model hyperparameters or tweak the reward function to speed-up learning convergence (i.e., reach the same accuracy with fewer training episodes).

- 5 pt: extend the notebook used in Lab 07, in order to learn to control the CartPole environment using *directly the screen pixels*, rather than the compact state representation used during the Lab (cart position, cart velocity, pole angle, pole angular velocity). NB: this will require to change the "observation_space" and to look for smart ways of encoding the pixels in a compact way to reduce computational complexity (e.g., crop the image around the pole, use difference of consecutive frames as input to consider temporal context, etc.).

or

- 5 pt: train a deep RL agent on a different Gym environment. You are free to choose whatever Gym environment you like from the available list, or even explore other simulation platforms.

Deadline: The complete homework (source code + report) must be submitted through Moodle at least 10 days before the chosen exam date.