Feedback from Group 20 on HW3 for Group 8

Hey Group 8,

We read through your code for HW3 and would like to share our feedback with you. We really enjoyed how concise, well structured and easy to understand your code was. The comments you made in your code were the perfect amount in order to further understand the code. We especially liked the simple implementation of the environment visualization.

We do however have some suggestions on how you could improve it even further:

- The result of the training is not that great. In the 2 runs we did, the space ship ended up not doing much at all and just floated to the bottom. Our results surely were not better and this is a bit out of scope for this homework but since your implementation is so well done this is one of the only things we could criticize so improving the performance would only be the cherry on top.
- You used a sequential model instead of implementing one with a functional forward step
- Q_target (Q_max in your code) should not be considered in the the loss calculation if the datapoint from which it originates was the last one in the trajectory and thus done=true. One way to implement this would be something like that: Q_targets = rewards + (gamma * Q_max * (1-done))
- We don’t get the use of the seed in the beginning