**HW 3 Feedback Group 8 from Group 22**

Functionality:

* Maybe there is a bug/room for improvement: Output layer activation is relu but linear could be better for a Q Value? But maybe your solution works even better 😀
* Why do you use create\_trajectory for both the model and model\_target in the beginning?
* Why did you limit the trajectory in the training process to 100 steps? It might be not even possible to solve the problem in 100 frames(???) 😀
* We ranyour code for 600 episodes and had rewards of around –2 – this seems to be inlausable, when you look at the reward range of this environment. Why is that? Is it because you take the mean of one whole trajectory instead of taking the means of many trajectories? You would need the sum of one Trajectory and the mean of sums of trajectories.
* Your epsilon gets bigger than one quite fast and therefore stops exploring relatively quickly, maybe fix that?

Visualisation:

* Awesome to give the option for rendering and also doing it from time to time!

Code:

* Nice that you worked with tf functions instead of basic python functions whenever possible
  + This made the code short and understandable
* You could have used more expressive variable names (but personal preference)
* Hyperparameters are hardcoded most of the time – maybe defining them as variables at a central place would help!
* You wrote your whole code in a very procedural way, maybe use classes next time :)