

## MiniProject

**Yu-Ting TSENG (A0212195L)**

e0503474@u.nus.edu

Department of Computer Science

National University of Singapore

**Chang-Sheng SUN (A0225554E)**

e0576188@u.nus.edu

Department of Computer Science

National University of Singapore

Nov 18, 2020

The best verdict we have got is 6.716. We train 2 recurrent neural networks, respectively predict the position for each car after 1 step and 2 steps. If the agent on the top lane, move “FORWARD” as many as possible (leading no collision). If the agent is in the square, which top-lefty corner is the target position, then choose action “UP”. If the agent is not in the square, get the best prediction “UP” or “FORWARD”. We take each estimation and the history of the cars as input of RNN. Then we achieve the verdict with 6.716.

In addition, we take the approach deep Q-learning. We make the action which is predicted by the position models instead of randomizing all the possible actions when smaller than epsilon. We set the learning\_rate to be  $1e-4$ ,  $1e-5$  and  $1e-6$  in order. In this case, we suppose to get a better verdict score, however, the results aren't as good as the original model. Then, we try the method with epsilon on the same model to tune the parameters with exploration of more possible action. We assume that some hyperparameters for reinforcement learning might cause this situation. In several trials of various of combination of hyperparameters, the results eventually improved as 7.050.