ELEC-E8125 - Reinforcement learning D, Lecture, 4.9.2023-29.11.2023: View: User report

User report

Nguyen Binh

Grade item	Grade	Feedback
ELEC-E8125 - Reinforcement learning D, Lecture, 4.9.2023-29.11.2023		
QUIZ Quiz 1	10.00	
QUIZ Quiz 2	10.00	
QUIZ Quiz 3	10.00	
QUIZ Quiz 4	10.00	
QUIZ Quiz 5	10.00	
QUIZ Quiz 6	8.00	
QUIZ Quiz 7	10.00	
QUIZ Quiz 8	8.00	
ASSIGNMENT Exercise 1	100.00	Email address: / T1: 10/10 Q1.1: 10/10 T2: 10/10 Q2.1: 15/15 Q2.2: 10/10 T3: 20/20 T4: 10/10 Q4.1: 5/5 Q4.2: 10/10 Total: 100/100
ASSIGNMENT Exercise 2	95.00	Student number (from Oodi): / Email address: / T1: 30/30 Q1.1: 5/5 Q1.2: 5/5 Q1.3: 5/5 T2: Generally the policy may converge earlier than value function because the optimal policy does not need the full state-value function to be optimal. - 10/15 T3: 5/5 T4: 10/10 Q4.1: 10/10 Q4.2: 15/15 Total: 95/100
ASSIGNMENT Exercise 3	97.50	Student number (from Oodi): / Email address: / T1 (25): 25/25 T2 (10): 10/10 Q2.1 (15): After the first episode only the states that were in that episode will have non-zero values and those values are alpha * 1, since all rewards are 1 for CartPole. - 12.5/15 T3 (10): 10/10 Q3.1 (5): 5/5 Q3.2 (15): 15/15 T4 (5): 5/5 Q4.1 (15): 15/15 Total: 97.5/100
ASSIGNMENT Exercise 4	94.00	DQN_Q3:stability RBF T1: 20/20 RBF Q1.1: 7/10 RBF Q1.2: 10/10 RBF Q1.3: 10/10 RBF T2: 9/10 Total: 56.0/60 RBF Task 1 (20 points) If plots are all correct: 20 points Else If Handcrafted features are correct: +2.5 points Else If States featurized correctly: +2.5 points Else If Terminal state handled correctly: +5 points Else If Target is correct: +5 points Else If one plot is missing: -1 point RBF Question 1.1 (10 points) Argue relation of state and value function is non-linear and therefore it cannot be learned directly with a linear approximator: 10/10 Discuss deeply about the features being non-linear but misses or doesn't mention the non-linear relation between action values and states: 8/10 Vague answer, e.g., dynamics non-linear, etc, bases observation on task 1 (we use an SGD regressor for the Q-values which is actually non-linear): 5/10 Says no without any argumentation: 1/10 Yes: 0/10 RBF Question 1.2 (10 points) Full answer: reuse previous data, randomness in sampling from buffer creates less correlated training data: 10/10 Arguing it uses previous data: 5/10 Just say we use experience replay because it makes learning faster: 2/10 RBF Question 1.3 (10 points) If say grid-based methods aren't sample efficient because they don't take similarities between states into account: +7 points You need to discretize continuous state spaces, which means there are a lot of states that must be explored during training: +3 points If say grid-based method may sometimes be sample-efficient, and provide a good explanation: 5/10 points If say no, because the plot shows the grid-based method performs poorly: 2/10 points If say no, without justifications: 1/10 points If say yes, with poor or incorrect justification: 0/10 points RBF Task 2 (10 points) If Good plot: 10/10 points If Incorrect axis labels, no legend, wrong axes: 5/10 points If plotting the wrong thing (e.g., wrong dimension or value function): 0/10 points
ASSIGNMENT Exercise 5	98.00	Student number (from Oodi): / Email address: / T1a (15): 15/15 T1b (5): 5/5 T1c (5): 5/5 Q1.1 (15): Elaborating on how a well-chosen baseline increases the probability of actions leading to higher returns would strengthen the answer. - 13/15 T2 (10): 10/10 Q2.1 (5): 5/5 Q2.2 (5): 5/5 Q3 (15): 15/15 Q4.1 (5): 5/5 Q4.2 (10): 10/10 Q5 (10): 10/10 Total: 98/100
ASSIGNMENT Exercise 6	100.00	Email address: / T1: 20/20 Q1.1: 10/10 Q1.2: 5/5 Q1.3: 10/10 Q1.4: 10/10 T2: 25/25 Q2.1: 10/10 Q2.2: 10/10 other: / Total: 100/100
ASSIGNMENT Exercise 7	100.00	Email address: / T1: 30/30 Q1.1: 5/5 Q1.2: 20/20 Q2.1: 10/10 T2: 30/30 Q2.2: 5/5 other: / Total: 100/100
ASSIGNMENT Final RL project	100.00	
COURSE TOTAL **Course total include empty grades.**	**860.50**	