Dear Group 22,

Task 1 looks really good, we like that you mention how the reward function can be modeled as a biased- or unbiased version.
Regarding the policy, do you mean that the state-action pair has to be defined manually? That sounds like a lot of work.

The same holds for the policy in Task 2.
If all of the factors for the state transition function are unknown maybe It would be needed to be modeled as probabilistic in lack of better options? Otherwise nicely done.

There's not much to remark in Task 3 either. In the self-driving car example, you write that the reward function can partly be designed by a user or engineer, doesn't it always has to completely be designed by someone, or where would the other part be obtained from?

Best,
Group 20