

准备好后再次尝试

获得的成绩 66.66% 最新提交作业的评分 66.67% 通过条件 80% 或更高

再试

您正在查看此评估的自动翻译版本

您可以根据需要切换回去，以源语言查看这些内容。即使您更改语言，也不会丢失任何进度。

显示英语（English）版本

取消

1.

0 / 1个点

月球着陆器是一个连续状态的马尔可夫决策过程（MDP），因为：

- ☐ 奖励包含连续估值的数字
- ☒ 状态-动作值 $Q(s, a)$ 函数输出连续值数
- ☐ 状态包含位置和速度等连续数值。
- ☐ 状态有多个数字，而不是只有一个数字（如 x - 方向上的位置）

错误

提示： "连续 "指的是国家。

2.

1 / 1个点

在视频中描述的学习算法中，我们反复创建一个人工训练集，对其进行监督学习，其中输入 $x = (s, a)$ ，使用贝尔曼方程构建的目标为 $y =$ _____?

- ☐ $y = \max_{a'} Q(s', a')$ 其中 s' 是您在状态 s 采取行动后的状态
- ☒ $y = R(s) + \gamma \max_{a'} Q(s', a')$ 其中 s' 是您在状态 s 中采取 行动后的状态
- ☐ $y = R(s)$
- ☐ $y = R(s')$ 其中 s' 是您在状态 s 中采取 行动后的状态

正确

3.

1 / 1个点

你们已经完成了本课的最后一次练习测验！请核对所有答案，因为所有答案都是正确的！)

☒ 安德鲁向您表示衷心的祝贺！

正确

☒ 你值得庆祝！

正确

☒ DeepLearning.AI 和斯坦福在线团队希望为你们鼓掌！

正确

☒ 多么了不起的成就 -- 你做到了！

正确