

1.

1 / 1 punto

El módulo de aterrizaje lunar es un proceso de decisión de Markov (MDP) de estado continuo porque:

- ☐ El estado tiene varios números en lugar de un solo número (como la posición en el X -dirección)
- ☐ El valor de la acción estatal $Q(s, a)$ la función genera números de valor continuo
- ☒ El estado contiene números como la posición y la velocidad que tienen valores continuos.
- ☐ La recompensa contiene números que tienen un valor continuo

✓ Correcto
¡Así es!

2.

1 / 1 punto

En el algoritmo de aprendizaje descrito en los videos, creamos repetidamente un conjunto de entrenamiento artificial al que aplicamos aprendizaje supervisado donde la entrada $X = (s, a)$ y el objetivo, construido usando las ecuaciones de Bellman, es $y =$ _____?

- ☐ $y = R(s')$ dónde s' es el estado al que llegas después de tomar acción a en estado s
- ☒ $y = R(s) + C \max_a q(s', a)$ dónde s' es el estado al que llegas después de tomar acción a en estado s
- ☐ $y = \max_a q(s', a)$ dónde s' es el estado al que llegas después de tomar acción a en estado s
- ☐ $y = R(s)$

✓ Correcto

3.

1 / 1 punto

¡Has llegado al examen de práctica final de esta clase! ¿Qué significa eso? (Por favor, compruebe todas las respuestas, ¡porque todas son correctas!)

✓ ¡Qué logro, lo lograste!

✓ Correcto

✓ ¡Andrew te envía sus más sinceras felicitaciones!

✓ Correcto

✓ ¡Los equipos de DeepLearning.AI y Stanford Online quisieran darle un aplauso!

✓ Correcto

✓ ¡Te mereces celebrar!

✓ Correcto