

[Página inicial](#) / [Meus cursos](#) / [CEMEAI-2020-RN](#) / [Seções](#) / [Introdução ao aprendizado por reforço](#) / [Avaliação Aula 7](#)

**Iniciado em** sábado, 7 nov 2020, 19:12

**Estado** Finalizada

**Concluída em** sábado, 7 nov 2020, 19:13

**Tempo  
empregado** 50 segundos

Questão **1**

Completo

Vale 2,00 ponto(s).

Considere a afirmação: o aprendizado por reforço é supervisionado. Em que sentido essa afirmação estaria correta?

Escolha uma:

- ☐ a. Problemas de aprendizado por reforço recebem por entrada muitos exemplos, alguns rótulos ligados a esse exemplos e executam por grande número de épocas
- ☐ b. Problemas de aprendizado por reforço recebem por entrada pares de exemplo e rótulo
- ☐ c. O aprendizado por reforço é apenas semi-supervisionado, pois um agente observa parcialmente o estado do ambiente, portanto essa afirmação está incorreta.
- ☒ d. O aprendizado por reforço é supervisionado por meio da observação do ambiente e retorno obtido na forma de uma recompensa após a realização de uma ação

Questão **2**

Completo

Vale 2,00 ponto(s).

Quais os passos fundamentais em uma execução de um algoritmo de aprendizado por reforço em um dado instante  $t$  dentro um episódio?

Escolha uma:

- ☐ a. Amostrar uma ação dentre as possíveis segundo uma política e fornecer essa ação como entrada para uma rede neural que irá classificá-la entre uma ação correta e incorreta
- ☒ b. Executar uma sequência de ações e, após atingir um estado terminal, computar a recompensa total
- ☐ c. Executar uma ação selecionada por meio de uma política atual, recebimento de uma recompensa e observação que potencialmente modifica o estado do agente
- ☐ d. Obter um exemplo de treinamento, prever uma ação com base nesse exemplo e comparar com a ação que deveria ter sido realizada

## Questão 3

Completo

Vale 2,00 ponto(s).

Quais os algoritmos de treinamento que representam as duas principais abordagens no aprendizado por reforço?

Escolha uma:

- ☐ a. Backpropagation e Stochastic Gradient Descent
- ☒ b. Policy Learning e Value Learning
- ☐ c. Actor-Critic e Feed-forward network
- ☐ d. Deep Q-Network e Entropia Cruzada

## Questão 4

Completo

Vale 2,00 ponto(s).

Na biblioteca gym carregue o ambiente Blackjack-v0 que representa o Jogo 21, no qual se pode pedir mais cartas ou parar, o ambiente Pendulum-v0 no qual o objetivo é equilibrar uma alavanca giratória num ângulo de 90 graus, o CartPole-v1, no qual o objetivo é equilibrar um poste sobre um carro, movimentando esse carro, e finalmente Acrobot-v1 que é um pêndulo bi-articulado com dois segmentos, com objetivo de levantar a ponta do segundo segmento a uma altura ao menos igual ao tamanho dos segmentos por cima da base.

Como é o espaço de ações possíveis desses ambientes?

Escolha uma:

- ☐ a. Blackjack: discreto 2 ações, Pendulum: discreto 2 ações, CartPole: contínuo entre -1 e 1, Acrobot: contínuo entre -1 e 1
- ☒ b. Blackjack: discreto 2 ações, Pendulum: contínuo entre -2 e 2, CartPole: discreto 2 ações, Acrobot: discreto 3 ações
- ☐ c. Blackjack: discreto 3 ações, Pendulum: discreto 2 ações, CartPole: contínuo entre -2 e 2, Acrobot: discreto 2 ações
- ☐ d. Blackjack: discreto 3 ações, Pendulum: contínuo entre -2 e 2, CartPole: discreto 3 ações, Acrobot: contínuo entre -1 e 1

## Questão 5

Completo

Vale 2,00 ponto(s).

Carregue o ambiente Blackjack-v0. Esse problema gera recompensa +1 para vitória, 0 para pedir uma carta, e -1 para derrota. Inicialize o ambiente, execute 100 mil episódios (cada um até o estado terminal) e calcule a média de recompensas totais (MR), a média de ações por episódio (MA), a taxa de vitórias (TV) e a taxa de derrotas (TD), sendo vitórias e derrotas medidas após o final de cada episódio. Arredonde os valores para 1 casa decimal.

Escolha uma:

- ☐ a. MR é 0.0, MP é 1.4, TV é menor do que TD
- ☒ b. MR é negativo, MP está entre 1 e 2, TV é menor do que TD
- ☐ c. MR é positivo, MP está entre 1 e 2, TV é igual a TD
- ☐ d. MR é negativo, MP está entre 2 e 3, TV é igual a TD

[◀ Exercícios - soluções](#)

Seguir para...

Avaliação ▶

