

Análise de Eficiência Q-learning

Gustavo Ribeiro da Mota e Pedro Lapenta Caserta

O algoritmo Q-learning é uma técnica de aprendizado artificial que consiste no registro de estados anteriores para garantir que o estado sucessor tenha a melhor taxa de escolha possível.

Em nosso trabalho, aplicamos essa técnica no famoso “snake game” e analisamos como a máquina agiu de acordo com o coeficiente de aprendizagem e o número de jogos.

1. Implementação

O Q-learning conta com 2 módulos principais. O primeiro dele é uma tabela de estados, onde é guardado a posição onde ele está e quais são os melhores movimentos para aquela posição.

O segundo módulo é o de tomada de decisões, onde há uma chance x , tal que $x \geq 1\%$. Caso a decisão seja aleatória, qualquer uma das direções é escolhida para o movimento. Caso contrário, o movimento é decidido com base na tabela Q feita anteriormente, tendo em base também os riscos da ação (colisão com o corpo ou com a parede). Caso seja um estado novo, a ação se torna aleatória.

Para a análise nesse trabalho utilizaremos 2 métricas: tamanho de amostra e tamanho do fator de aprendizado. Utilizaremos 3 tamanhos de amostras, sendo eles: 250, 500 e 1000 tentativas. Utilizaremos também 3 taxas de aprendizagem: 0.8, 0.5 e 0.2.

Faremos 3 testes para cada uma das combinações para compararmos o valor médio alcançado.

Ao final do projeto teremos a seguinte tabela preenchida:

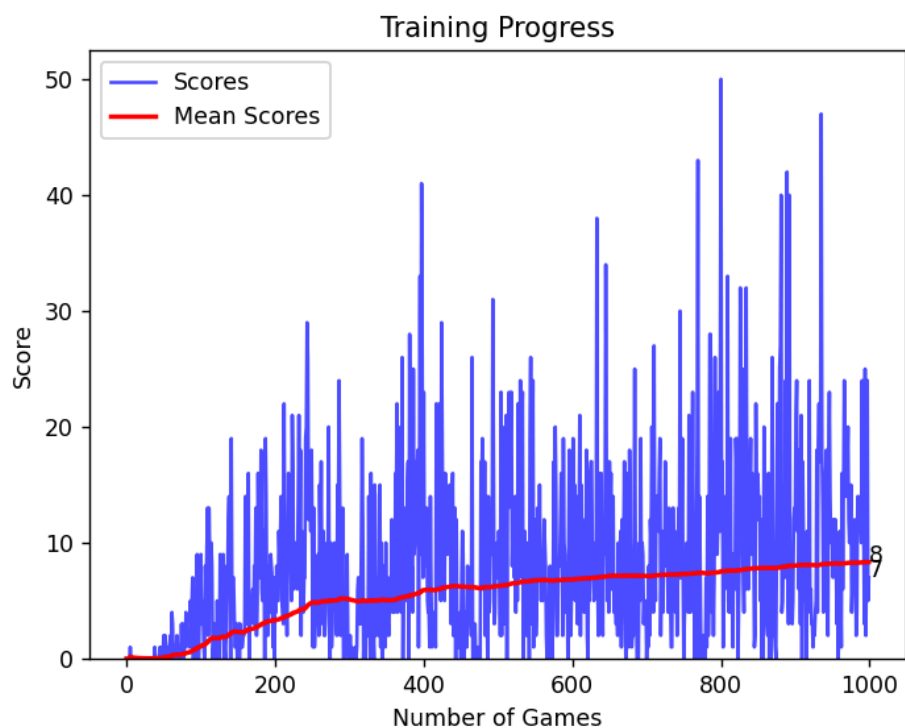
Tamanho da amostra/ Taxa de aprendizagem	0.8	Média 0.8	0.5	Média 0.5	0.2	Média 0.2
250						
500						
1000						

2. Tabela Preenchida

Tamanho da amostra/ Taxa de aprendizagem	0.8	Média 0.8	0.5	Média 0.5	0.2	Média 0.2
250	5	4	7	6	8	8,333333333
	4		6		9	
	3		5		8	
500	6	6	9	8	13	13
	6		7		14	
	6		8		12	
1000	6	6,666666667	11	10,66666667	20	17
	7		10		15	
	7		11		16	

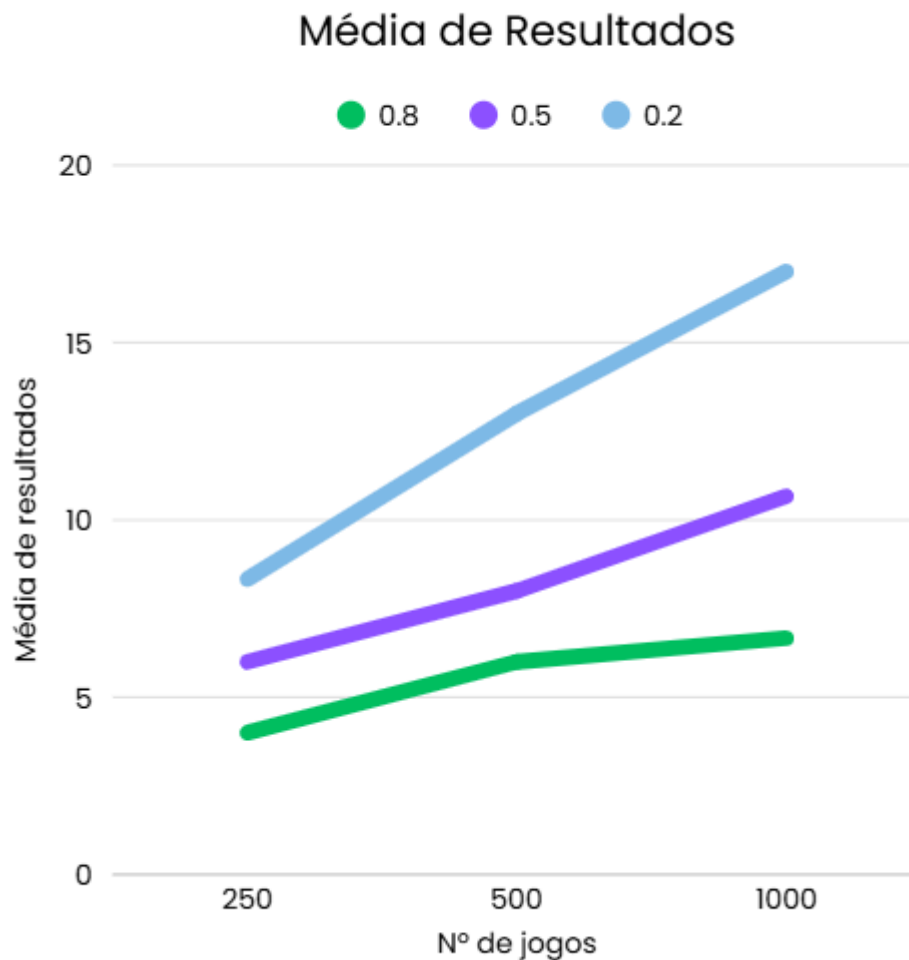
3. Conclusão

Quanto maior o alpha, menor o resultado médio apontado. Enquanto o resultado continuou progressivo, a média foi diminuindo em comparação aos alphas de menores valores.



1. Gráfico do segundo jogo de 1000 tentativas com $\alpha = 0.8$

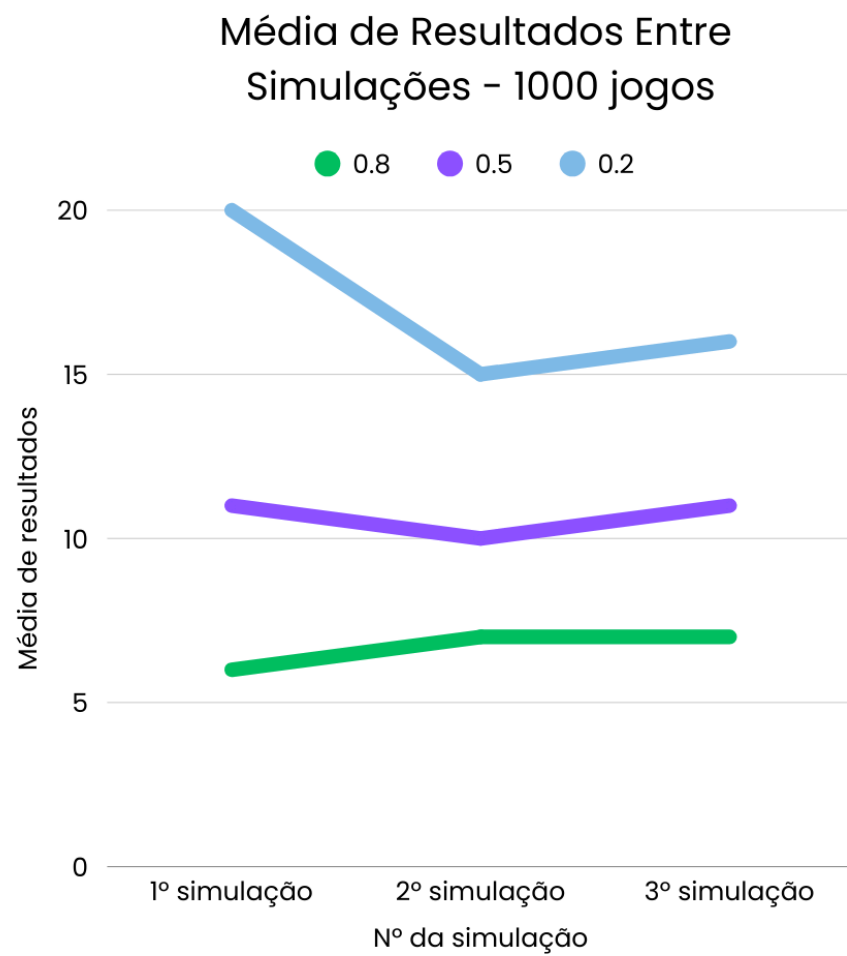
A média de resultados das simulações de diferentes tamanhos pode ser observada a partir do gráfico



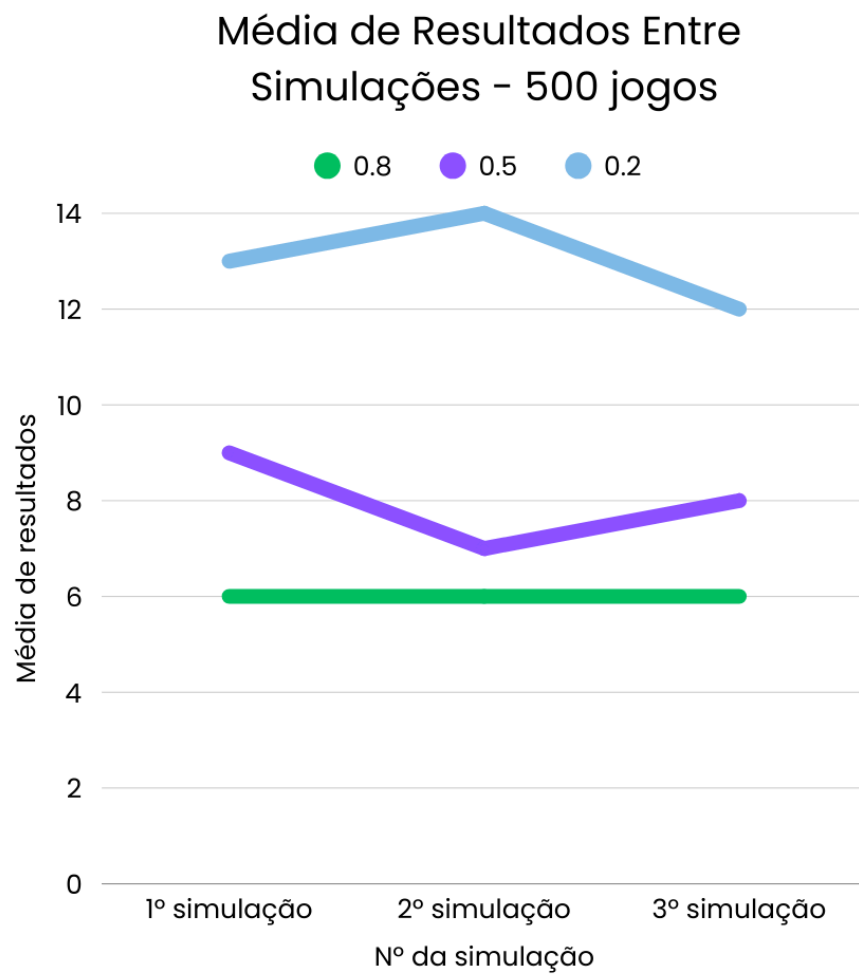
2. Gráfico da média dos resultados

Como podemos observar, quando α foi 0.8, a média dos resultados foi a menor independentemente de quantas simulações foram realizadas. Já quando estava em 0.2, os resultados foram muito mais altos, ficando até 204% maior quando é feita a simulação com 1000 jogos.

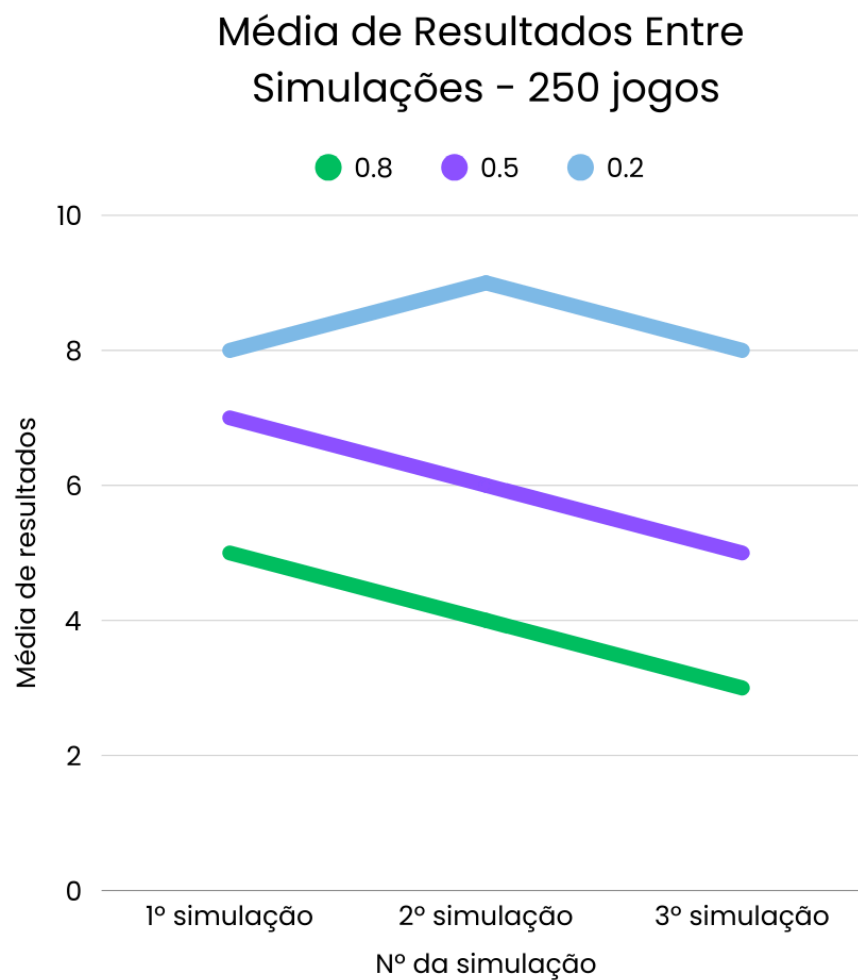
Quando criamos o gráfico em relação a 1000 jogos, percebemos que o resultado nos 3 casos de simulação apresentaram o melhor resultado quando α está em 0.2. O resultado é o mesmo com 500 e 250 jogos.



3. Gráfico da média de resultados entre simulações - 1000 jogos



4.Gráfico da média de resultados entre simulações - 500 jogos



5. Gráfico da média de resultados entre simulações - 250 jogos