

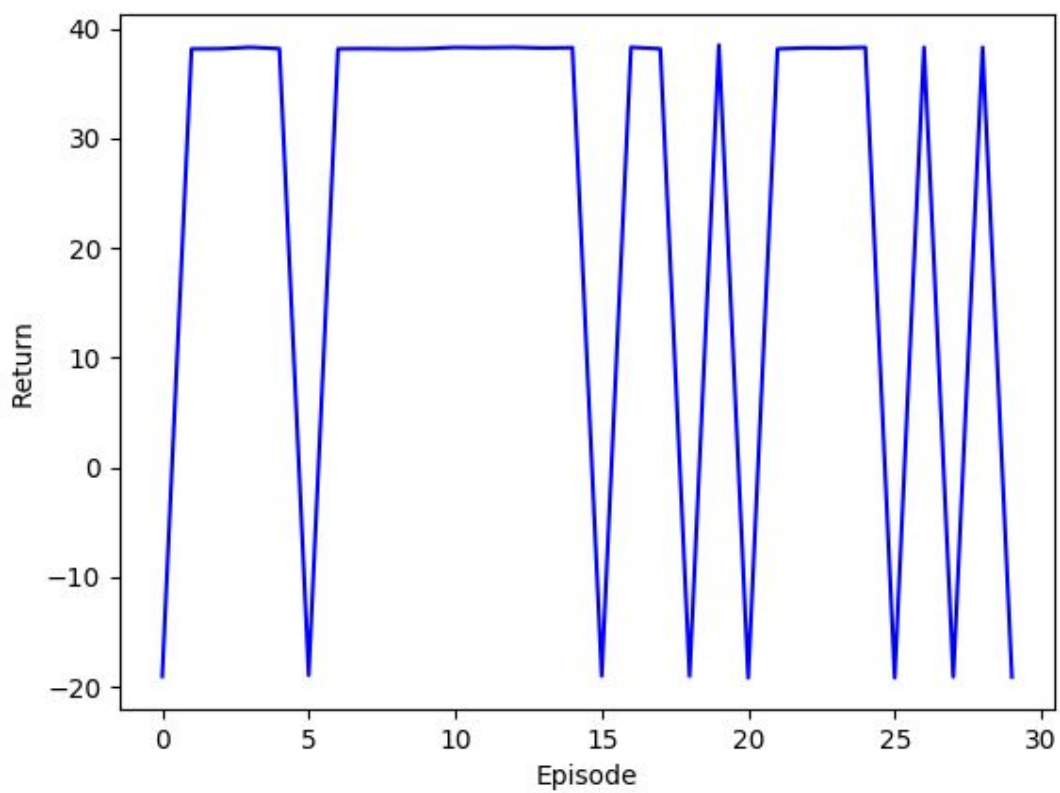
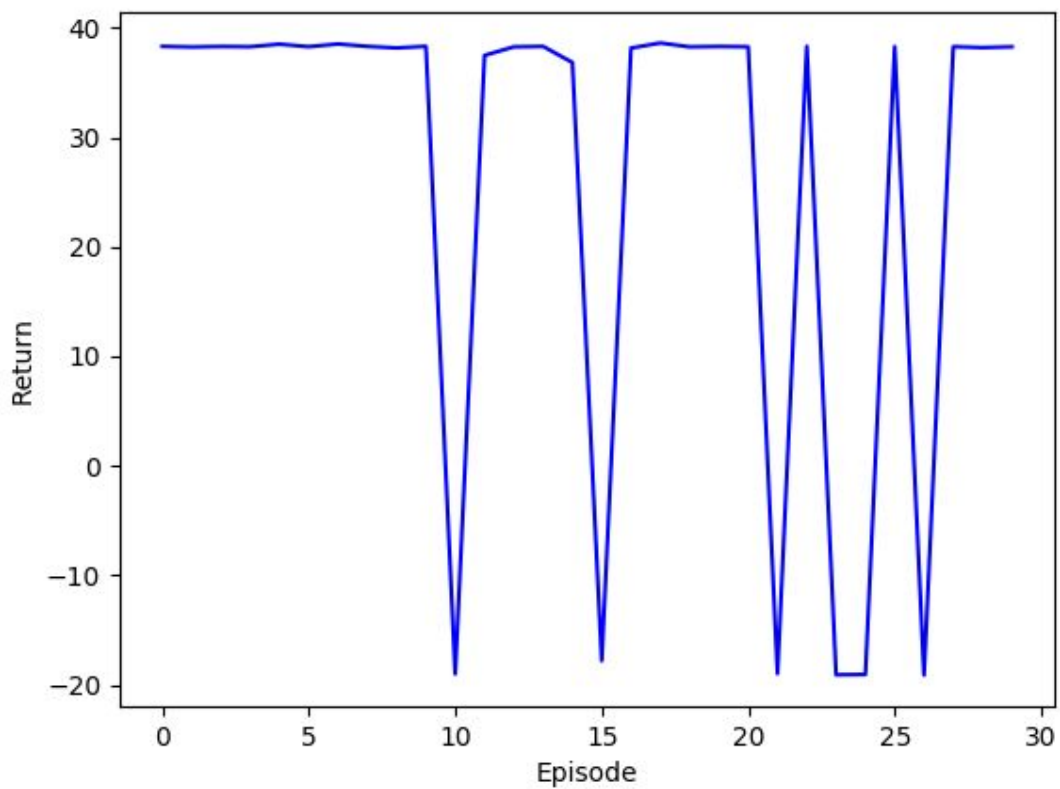
## **Relatório Lab 12 - CT213**

Código: A implementação do código foi bem direta, a dificuldade maior foi entender como a rede neural se encaixava nas funções (entender que ela calcula o q-value de cada estado).

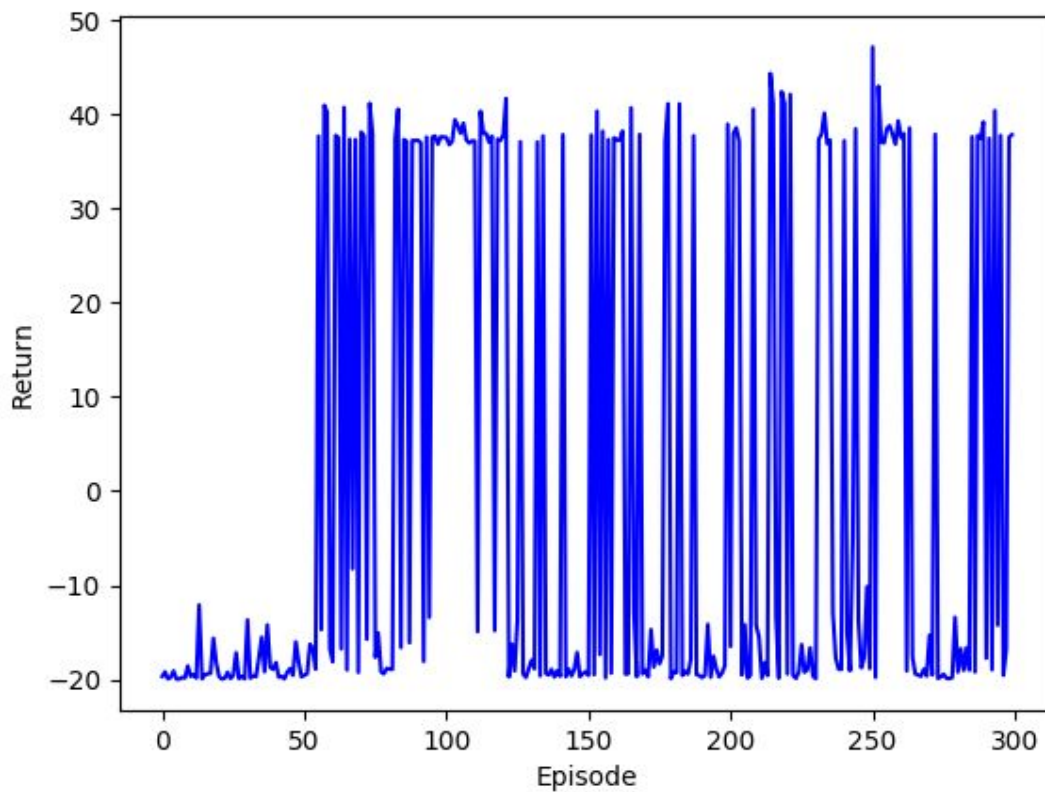
Resultados: No treinamento, o agente conseguiu completar a tarefa pela primeira vez por volta da interação 40 e, como é perceptível pelo gráfico do treinamento, a partir da interação 100, não conseguimos atingir retornos muito melhores do que os já alcançados. Talvez esse treinamento excessivo tenha prejudicado o desempenho final do modelo.

Ao avaliar a política, obtivemos resultados com taxa de objetivo concluído acima de 70% na grande maioria das vezes.

### Avaliações de política com 30 episódios:



### Treinamento do modelo com 300 épocas



### Política Aprendida pelo Agente:

