

Q-Learning adalah algoritma reinforcement learning. Tujuan Q-Learning adalah untuk membentuk kebijakan/policy berdasarkan lingkungannya (reward atau punishment).

Sebuah agen Q-learning terdiri atas set *States* dan set *Actions*. Dengan melakukan action, agen mengalami transisi ke state lain. Tujuan agen adalah memaksimalkan total reward kedepannya dengan cara melihat dan mengakumulasi reward pada state selanjutnya.

Q-Learning menggunakan formula berikut untuk melakukan transisi:

$$Q(state, action) \leftarrow (1 - \alpha)Q(state, action) + \alpha \left(reward + \gamma \max_a Q(next\ state, all\ actions) \right)$$

Nilai α mempengaruhi learning rate agen. Nilai yang besar dapat mempercepat proses learning, namun akan menghasilkan osilasi pada skenario tertentu.

Nilai γ mempengaruhi agen dalam mengambil reward sekarang vs masa depan (nilai mendekati 0, agen akan memprioritaskan reward sekarang)

Berikut adalah algoritma Q-Learning:

1. Inisialisasi tabel Q
2. Untuk setiap episode
 - a. Inisialisasi state s
 - b. Untuk setiap langkah pada episode tersebut

Lakukan action random pada state tersebut secara random (eksploratif)

Saat mengambil aksi action, cek nilai rewardnya pada state' (state masa depan)

Update: $Q(state, action) \leftarrow (1 - \alpha)Q(state, action) + \alpha \left(reward + \gamma \max_a Q(next\ state, all\ actions) \right)$

Update: State \leftarrow state'

-1.0	-2.0	-3.0	-4.0	-5.0	-6.0	-7.0	-8.0	-9.0	-10.0	-11.0	-12.0	-13.0	-14.0	500.0
-1.0	-3.0	-1.0	2.0	-4.0	-1.0	-3.0	-1.0	-4.0	-2.0	-4.0	-2.0	-3.0	-2.0	-1.0
-4.0	-2.0	-1.0	-4.0	-2.0	-1.0	-2.0	-4.0	-2.0	-3.0	-2.0	-1.0	-2.0	-4.0	-4.0
-4.0	-2.0	-4.0	-1.0	-3.0	-2.0	-3.0	-2.0	-4.0	-2.0	-4.0	-1.0	-2.0	-4.0	-2.0
-4.0	-2.0	-2.0	-3.0	-2.0	-3.0	-1.0	-1.0	-4.0	-2.0	-1.0	-3.0	-4.0	-2.0	-4.0
-4.0	-3.0	-3.0	-4.0	-2.0	-3.0	-4.0	-2.0	-2.0	-3.0	-1.0	-2.0	-1.0	-2.0	-1.0
-2.0	-3.0	-2.0	-1.0	-1.0	-3.0	-2.0	-1.0	-4.0	-3.0	-1.0	-1.0	-2.0	-3.0	-3.0
-3.0	-1.0	-1.0	-4.0	-4.0	-3.0	-1.0	-2.0	-3.0	-1.0	-1.0	-4.0	-4.0	-3.0	-3.0
-3.0	-1.0	-4.0	-2.0	-3.0	-3.0	-1.0	-4.0	-4.0	-4.0	-2.0	-2.0	-2.0	-2.0	-1.0
-3.0	-4.0	-4.0	-2.0	-3.0	-4.0	-3.0	-3.0	-2.0	-2.0	-3.0	-4.0	-3.0	-4.0	-1.0
-3.0	-4.0	-1.0	-1.0	-1.0	-4.0	-4.0	-4.0	-4.0	-1.0	-2.0	-4.0	-2.0	-2.0	-1.0
-1.0	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0	-4.0	-1.0	-2.0	-4.0	-1.0	-2.0	-4.0
-2.0	-2.0	-1.0	-2.0	-2.0	-2.0	-2.0	-4.0	-3.0	-1.0	-4.0	-1.0	-4.0	-2.0	-2.0
-2.0	-1.0	-3.0	-1.0	-4.0	-4.0	-1.0	-3.0	-3.0	-1.0	-1.0	-2.0	-3.0	-4.0	-3.0
-2.0	-2.0	-1.0	-4.0	-4.0	-4.0	-2.0	-2.0	-3.0	-1.0	-2.0	-2.0	-1.0	-1.0	-3.0