

الفرق بين خوارزمية Q-Learning و Sarsa ٣

يُعتبر الأداء أحد أهم العوامل التي تؤخذ بعين الاعتبار عند اختيار خوارزمية التعلم المعزز. خوارزميات Q-Learning و Sarsa كل منهما تتميز بنقاط قوة وضعف تؤثر على أداء الوكيل في البيئات المختلفة.

أداء Q-Learning

تميل Q-Learning إلى تقديم أداء أفضل في البيئات التي تتطلب اتخاذ قرارات جريئة أو غير متوقعة. بما أن الخوارزمية تأخذ بعين الاعتبار أفضل فعل ممكن في كل خطوة، فإنها تميل إلى تحسين السياسة بشكل أسرع، خصوصاً إذا كانت المكافآت المرتبطة بالأفعال واضحة ومباشرة. ومع ذلك، يمكن أن تؤدي هذه الخوارزمية إلى اتخاذ قرارات خطيرة إذا كانت البيئة تحتوي على حالات غير مستقرة أو غير مؤكدة.

أداء Sarsa

على الجانب الآخر، تقدم Sarsa أداءً أكثر استقراراً في البيئات التي تتطلب قرارات حذرة. من خلال التحديث بناءً على الفعل الفعلي الذي يتم اتخاذه، تتكيف Sarsa مع السياسات الحالية وتوفر أداءً أكثر تحفظاً. قد يكون هذا النهج مفيداً في البيئات التي تحتوي على مخاطر غير متوقعة أو مكافآت متغيرة.

خاتمة

يؤثر الفرق بين Q-Learning و Sarsa بشكل مباشر على أداء التعلم بناءً على طبيعة البيئة ونوع القرارات المطلوبة. Q-Learning تقدم أداءً أفضل في البيئات الثابتة والجريئة، بينما تكون Sarsa أكثر فعالية في البيئات الحذرة والمعقدة.

