**ANALISA HASIL**

|  |  |
| --- | --- |
| 1. | Bagaimana performa agen dalam mengontrol environment CartPole? |
| \* | Agen RL belajar menyeimbangkan tiang pada CartPole melalui banyak percobaan. Awalnya, sering gagal, namun seiring waktu agen menjadi lebih baik dan mampu menyeimbangkan tiang lebih lama. |
| 2. | Bagaimana perubahan parameter (misal: gamma, epsilon, learning rate) mempengaruhi kinerja agen? |
| \* | 1. Gamma (Discount Factor) Semakin tinggi nilai gamma, agen mempertimbangkan reward jangka panjang. Jika gamma terlalu rendah, agen hanya fokus pada reward langsung. 2. Epsilon (Exploration Rate) Nilai epsilon tinggi membuat agen mencoba berbagai aksi secara acak untuk menemukan strategi terbaik. Secara bertahap, nilai epsilon diturunkan agar agen lebih sering memilih aksi yang terbukti efektif. 3. Learning Rate menentukan seberapa cepat agen mengubah perkiraan nilainya. Nilai yang terlalu tinggi bisa membuat pembelajaran tidak stabil, sedangkan nilai yang terlalu rendah membuat proses belajar menjadi lambat. |
| 3. | Apa tantangan yang muncul selama pelatihan agen RL? |
| \* | Agen kadang mengalami fluktuasi reward yang menyebabkan pembelajaran tidak stabil. Menentukan kapan harus mencoba aksi baru (eksplorasi) dan kapan harus menggunakan aksi yang sudah diketahui efektif (eksploitasi) merupakan tantangan tersendiri. Proses pelatihan, terutama dengan algoritma seperti DQN, membutuhkan banyak waktu dan sumber daya komputasi. |

**DISKUSI HASIL**

|  |  |
| --- | --- |
| 1. | Apa perbedaan utama antara Reinforcement Learning dan metode supervised learning dalam sistem kendali? |
| \* | Supervised learning menggunakan data yang sudah dilabeli, sedangkan RL belajar dari interaksi langsung dengan lingkungan tanpa label. Dalam supervised learning, model belajar dari contoh yang benar. Di RL, agen belajar melalui trial-and-error dengan mendapatkan reward atau hukuman. RL berfokus pada mendapatkan reward kumulatif maksimal, sedangkan supervised learning fokus pada akurasi prediksi. |
| 2 | Bagaimana strategi eksplorasi (exploration) dan eksploitasi (exploitation) dapat dioptimalkan pada agen RL? |
| \* | Eksplorasi (Agen mencoba berbagai aksi untuk menemukan strategi terbaik) dan Eksploitasi (Agen menggunakan strategi yang sudah terbukti memberikan reward tinggi). Optimalisasi, Pengurangan nilai epsilon secara bertahap (epsilon decay) membantu menyeimbangkan antara eksplorasi (mencoba hal baru) dan eksploitasi (menggunakan strategi yang sudah diketahui efektif). |
| 3. | Potensi aplikasi lain dari RL dalam sistem kendali nyata apa saja yang dapat diimplementasikan? |
| \* | Robotika, Untuk mengontrol pergerakan robot dalam navigasi atau manipulasi objek. Kendaraan Otonom, Mengoptimalkan keputusan pada mobil self-driving. |