

Terminology

Agent	Pelaku
Reward	Hadiah
Action	Tindakan
State	Keadaan
Environment	Lingkungan
Eksplorasi	Aksi algoritme RL dalam menggunakan aksi sebelumnya yang mendapat reward positif.
Ekplorasi	Aksi algoritme RL dalam melakukan aksi berbeda untuk mencari reward positif yang lain.
Policy	Rules (aturan) atau strategi yang digunakan oleh agent untuk melakukan action (A) selanjutnya, berdasarkan state (S) saat ini.
Value function (V)	Total nilai jangka panjang yang diharapkan (expected longterm return without discount) dari state saat ini di bawah policy π . Value ini adalah kebalikan dari short-term reward (R).
Optimal Value Function	Sebuah fungsi yang memiliki nilai tertinggi untuk semua state dibandingkan dengan fungsi nilai lainnya.
Optimal Policy	Policy (kebijakan) yang memiliki fungsi nilai optimal.
Model Environment	Segala sesuatu yang meniru perilaku environment. Atau, secara umumnya, sebuah kesimpulan tentang bagaimana perilaku dari environment.
Model Based RL	Model dasar atau paling sederhana dan memiliki perencanaan terhadap tindakannya, atau dapat juga dimaknai sebagai agent mengeksploitasi informasi yang dipelajari sebelumnya untuk menyelesaikan tugasnya.
Model Free RL	Kebalikan dari model Model Based RL, karena agent belajar dari lingkungan melalui metode trial and error untuk memperoleh pengalamannya.

Tipe-Tipe Environment RL:

- Deterministic environment
- Stochastic environment
- Fully observable environment
- Partially observable environment
- Discrete environment
- Continuous environment
- Episodic and non-episodic environment
- Single and multi-agent environment

