

Laporan Tugas 3

*Machine Learning*

***Q-Learning***

Disusun oleh :

Odia Pratama 1301154405

IF-39-13

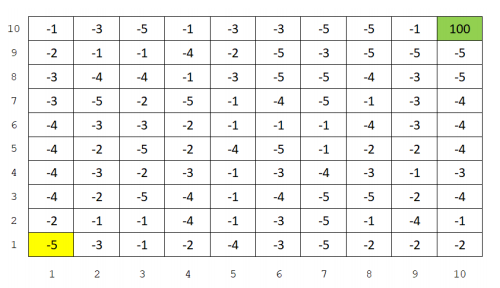
Universitas Telkom

Bandung

2018

Analisis Masalah

Membangun sebuah sistem Q-Learning untuk menemukan *optimum policy* sehingga Agent yang berada pada posisi (1, 1) dapat menuju ke *goal state* (10, 10) dengan mendapatkan total reward yang maksimum. *Agent* dapat melakukan berbagai aksi diantaranya : N (*North*) atau bergerak ke atas, E (*East*) atau bergerak ke kanan, W (*West*) atau bergerak ke kiri, dan S (*South*) atau bergerak ke bawah. Dengan *grid world* sebegai berikut:



Desain

**Inisialisasi Q**

**Pilih *action* dari Q**

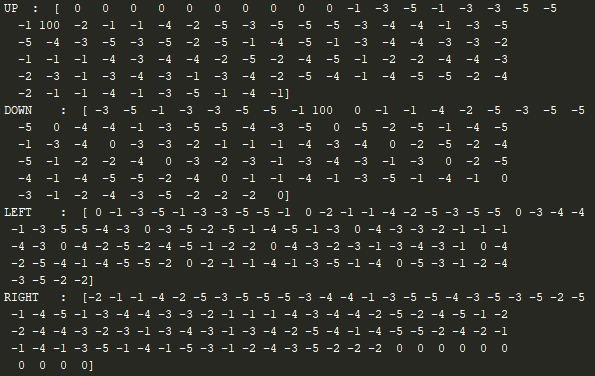
**Melakukan *action***

**Mengukur/menghitung *reward***

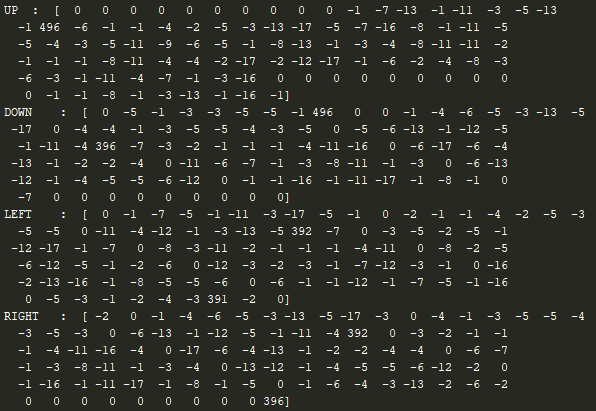
**Update Q**

Evaluasi Hasil Eksperimen

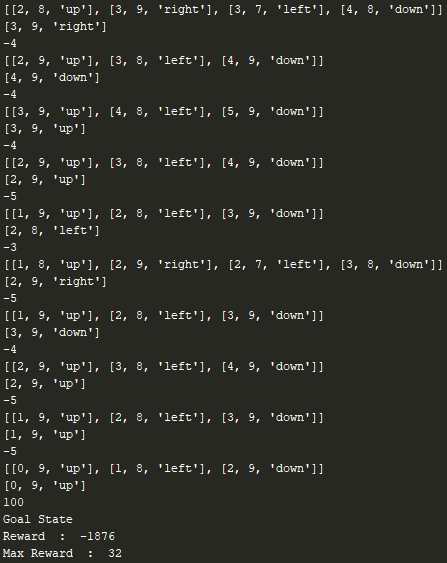
Tabel R



Tabel Q



Pergerakan *Agent* dan Reward maksimal yang didapat



Pada percobaan running program menggunakan jumlah *episode* sebanyak 1000. Dapat disimpulkan bahwa jumlah *episode* dapat memengaruhi jumlah *reward* yang didapat, namun bukan berarti jika semakin banyak *episode* maka akan semakin bagus *reward* yang didapat. Jika semakin banyak *episode* maka komputasi program akan semakin membutuhkan waktu yang lama.

Dari program yang dijalankan akan menghasilkan tabel Q yang berisikan nilai *reward* yang telah dihitung untuk menentukan step selanjutnya dengan formula:



Pada program yang dibuat dengan menggunakan nilai gamma = 0,8. Sehingga hasil maksimal *reward* yang didapat adalah 32. Nilai tersebut didapat dengan pergerakkan dari *initial state*(1,1) hingga *goal state*(10,10).