# project summary

|  |  |  |
| --- | --- | --- |
| Tanggal | Project | Nama Peneliti |
| **22-08-22 s/d 26-08-22** | **Hardware Architecture for Intelligent Traffic Light Based on Reinforcement Learning** | **Zulfikar Nima Arifuzzaki** |

# Status Summary

* Membuat pemodelan hardware

# Project Overview

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Hari | Pekerjaan | Persentase | Kendala | Rencana Kedepannya | Catatan |
| Senin | Membuat model hardware v1 pada python | 50% | Nilai Q-value hanya terisi pada state dengan level kemacetan tinggi. | * Melakukan modifikasi mekanisme perubahan state terhadap pemilihan action. * Mengecilkan range reward yang diberikan pada agent. * Mengubah mekanisme perubahan action. * Menambahkan state goal | Karakteristik model v1:   * State (256): + 1 level jika merah, -1 level jika hijau * Action (4): memilih ruas jalan denga kemacetan tertinggi. **Action diambil berdasarkan state.** * Reward (3): negative jika hijau diberikan pada ruas jalan lengang, positif jika lampu hijau diberikan pada ruas jalan macet, 0 jika bukan kedua case diatas. * Reward based on action * Tidak ada state goal |
| Selasa | 100% | Nilai cumulatif reward tidak converge. | Studi literatur lanjut untuk dataflow algoritma Q-learning |
| Rabu | Membuat model hardware v2.1 pada python. | 50% | Nilai Q-value hanya terisi pada state dengan level kemacetan tinggi. | * Melakukan modifikasi mekanisme perubahan state terhadap pemilihan action. * Mengubah reward policy * Studi literatur lanjut untuk dataflow algoritma Q-learning | Karakteristik model v2.1:   * **State (256): + 1 level jika merah, level jika hijau.** * Action (4): memilih ruas jalan dengan kemacetan tertinggi. **Action diambil berdasarkan Q-value.** * Reward (3): negative jika hijau diberikan pada ruas jalan lengang, positif jika lampu hijau diberikan pada ruas jalan macet, 0 jika bukan kedua case diatas. * Reward based on action * **Ada 4 state goal** |
| 100% | - Nilai kumulatif reward sangat linear (tidak sesuai ekspektasi). |
| Agent tidak pernah mencapai state goal. |
| Kamis | Membuat model hardware v2.2 pada python | 50% | Nilai Q-value hanya terisi pada state dengan level kemacetan tinggi (namun sudah lebih baik daripada v2.1) | * Melakukan modifikasi mekanisme perubahan state terhadap pemilihan action. * Mengubah reward policy * Studi literatur lanjut untuk dataflow algoritma Q-learning | Karakteristik model v2.2:   * **State (256): + 1 level jika merah, level jika hijau.** * Action (4): memilih ruas jalan denga kemacetan tertinggi. **Action diambil berdasarkan Q-value.** * Reward (3): negative jika hijau diberikan pada ruas jalan lengang, positif jika lampu hijau diberikan pada ruas jalan macet, 0 jika bukan kedua case diatas. * Reward based on action * **Ada 4 state goal** |
| 100% | Agent tidak pernah mencapai state goal. |
| Jumat | Membuat model hardware v2.3 pada python | 50% | Nilai Q-value hanya terisi pada state dengan level kemacetan tinggi (namun sudah lebih baik daripada v2.1) | * Melakukan modifikasi mekanisme perubahan state terhadap pemilihan action. * Mengubah reward policy * Studi literatur lanjut untuk dataflow algoritma Q-learning | Karakteristik model v2.2:   * **State (256): + 1 level jika merah, level jika hijau.** * Action (4): memilih ruas jalan denga kemacetan tertinggi. **Action diambil berdasarkan Q-value.** * Reward (3): negative jika next state max traffic, positive jika next state low traffic. * **Reward based on next state** * **Ada 4 state goal** |
| 100% | Agent tidak pernah mencapai state goal. |