

Reinforcement Learning **PENERAPAN Q-LEARNING PADA BUS ROUTE ENVIRONMENT**

LATAR BELAKANG

Masalah pada pemilihan jalur atau rute bus yang tidak optimal dapat menyebabkan dampak negatif. Dampak negatif dapat berupa waktu perjalanan yang lebih lama, biaya operasional yang dikeluarkan akan lebih tinggi, dan penumpang dapat merasa tidak nyaman karena harus menempuh jarak yang lebih jauh.

RUMUSAN MASALAH

Apakah masalah pemilihan jalur atau rute bus yang tidak optimal dapat diselesaikan dengan algoritma Q-Learning?

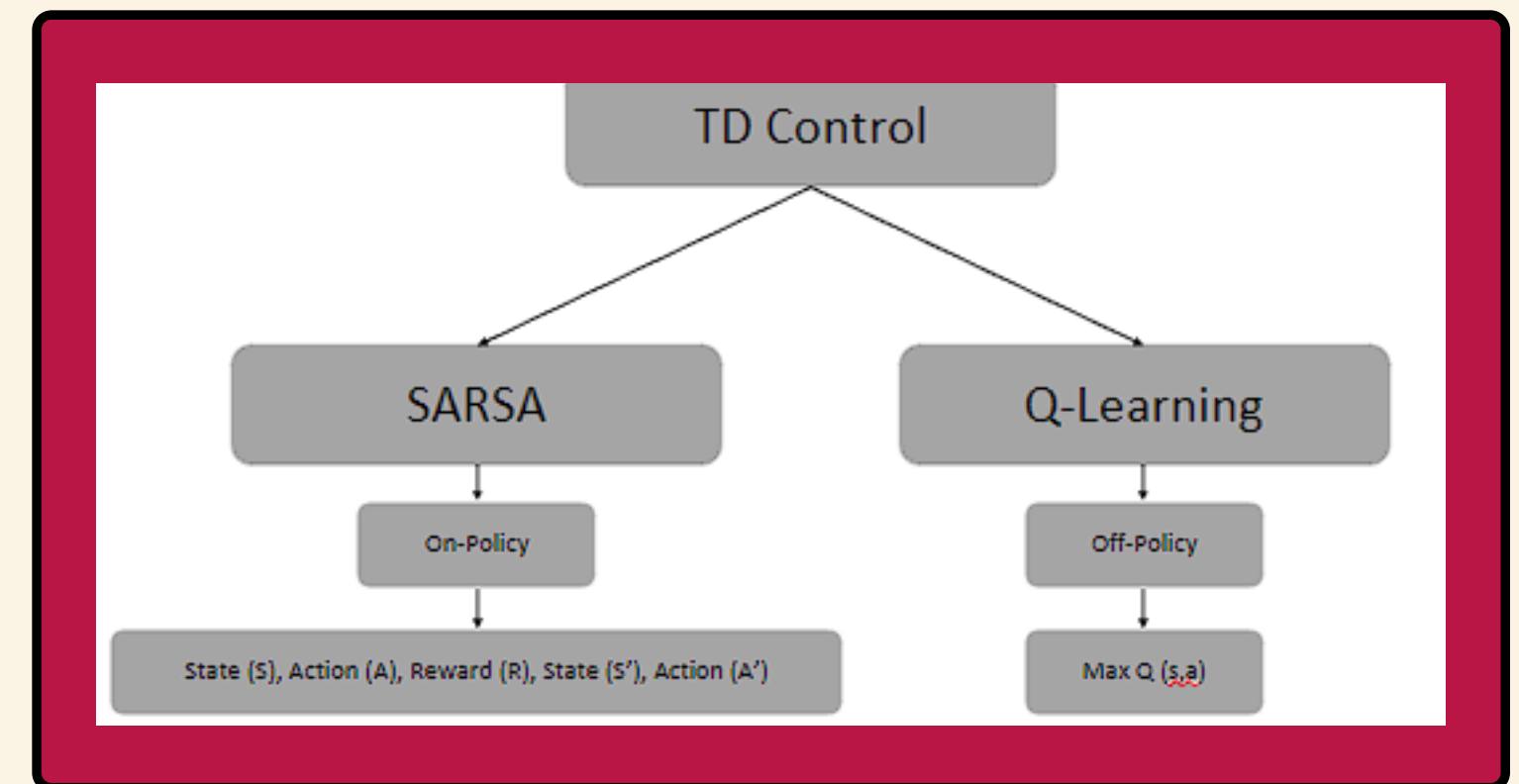
TUJUAN

Tujuannya adalah mendapatkan hasil jalur atau rute bus dari titik awal ke halte secara optimal.



Teori algoritma Q-LEARNING

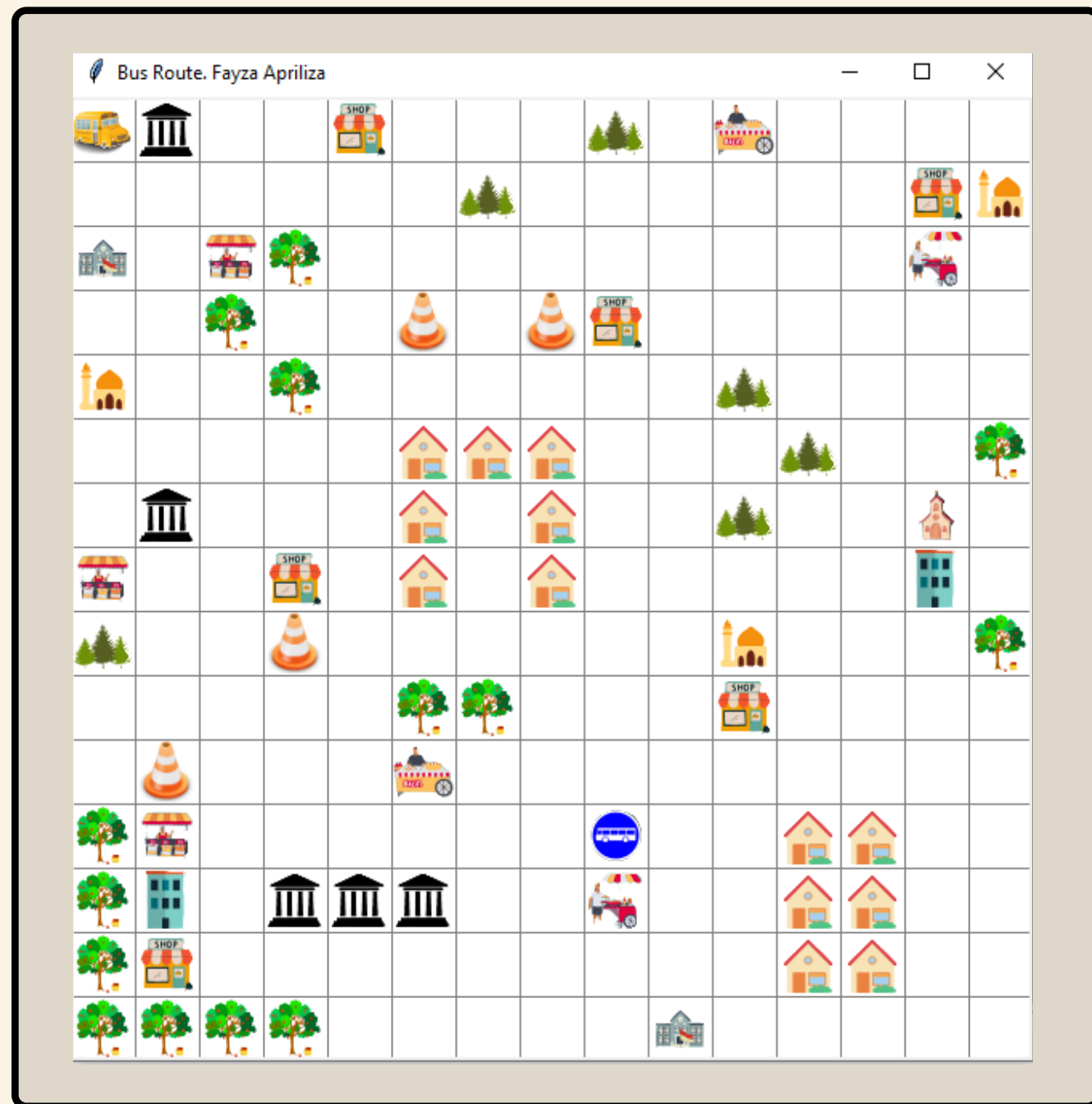
Q-learning adalah salah satu algoritma reinforcement learning yang digunakan untuk mengoptimalkan keputusan yang diambil oleh agen dalam suatu lingkungan atau situasi. Algoritma ini didasarkan pada konsep Q-value, yaitu nilai ekspektasi reward yang diperoleh jika agen melakukan aksi tertentu pada suatu state dalam lingkungan.



Konsep & Cara kerja Q-LEARNING

- Terinspirasi dari value iteration
- Sample an action
- Observe the reward and the next state
- Take the action with the highest Q (Max Q)





Studi Kasus PEMODELAN RL

Membantu robot bus untuk menemukan halte bus dengan step yang sedikit dan menghindari obstacles yang ada. Pada kasus terdapat 4 action, yakni ke atas, bawah, kanan, atau kiri.

State	Reward
Empty	0
Obstacles	-1
Halte	1



Reinforcement Learning PEMODELAN

0.01

Learning Rate

0.9

Discount Factor

0.9

Epsilon Greedy

1500

Episodes

```
# Summed costs for all episodes in re
all_costs = []

for episode in range(1500):
    # Initial Observation
    observation = env.reset()
```

```
# Creating class for the Q-learning table
class QLearningTable:
    def __init__(self, actions, learning_rate=0.01, reward_decay=0.9, e_greedy=0.9):
        # List of actions
        self.actions = actions
        # Learning rate
```



ALFREDO TORRES

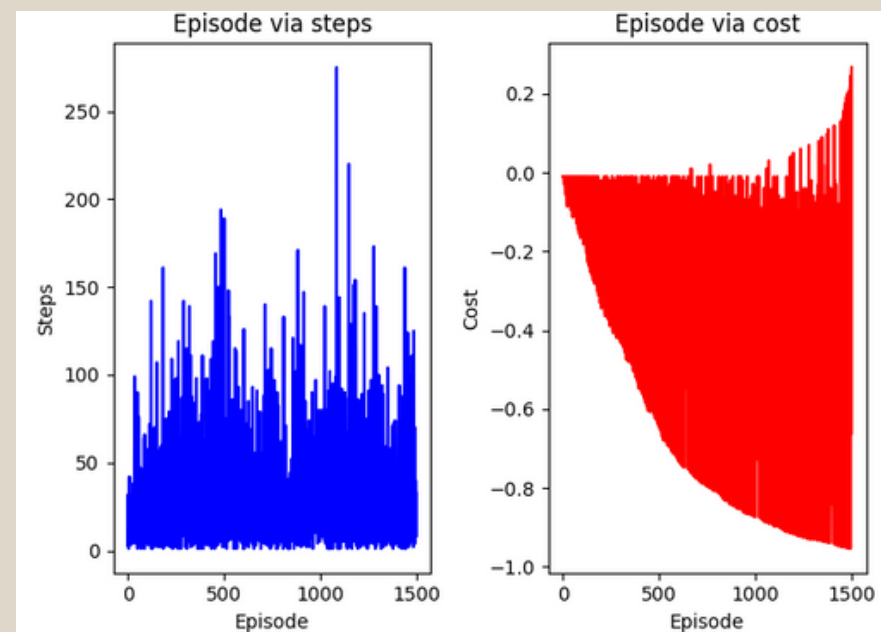
Hasil PERFORMA

Berdasarkan hasil jalur atau rute yang diperoleh, step yang dibutuhkan untuk mencapai goals sudah efektif. Artinya, secara keseluruhan agent telah belajar dengan baik dan jalur atau rute yang dihasilkan sudah optimal.

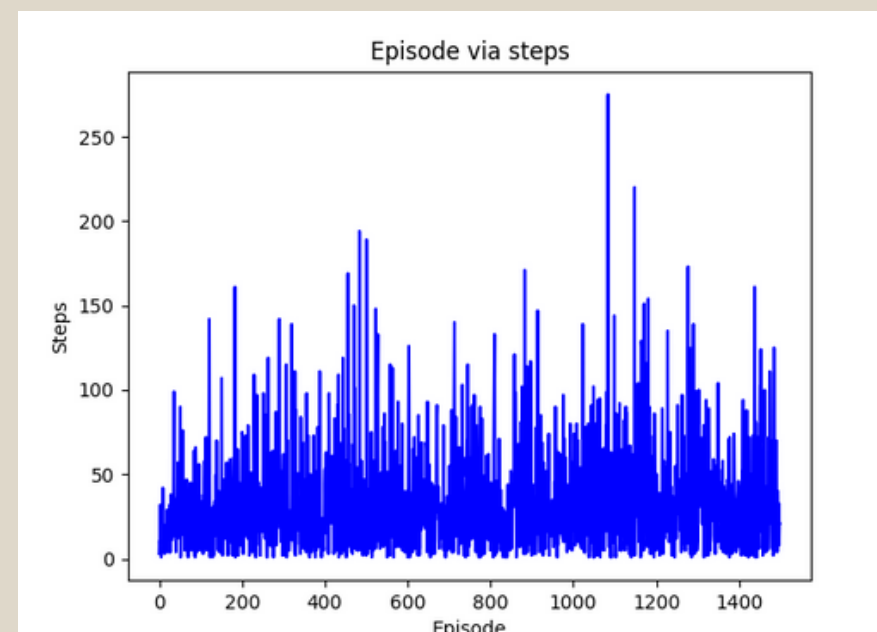


HASIL

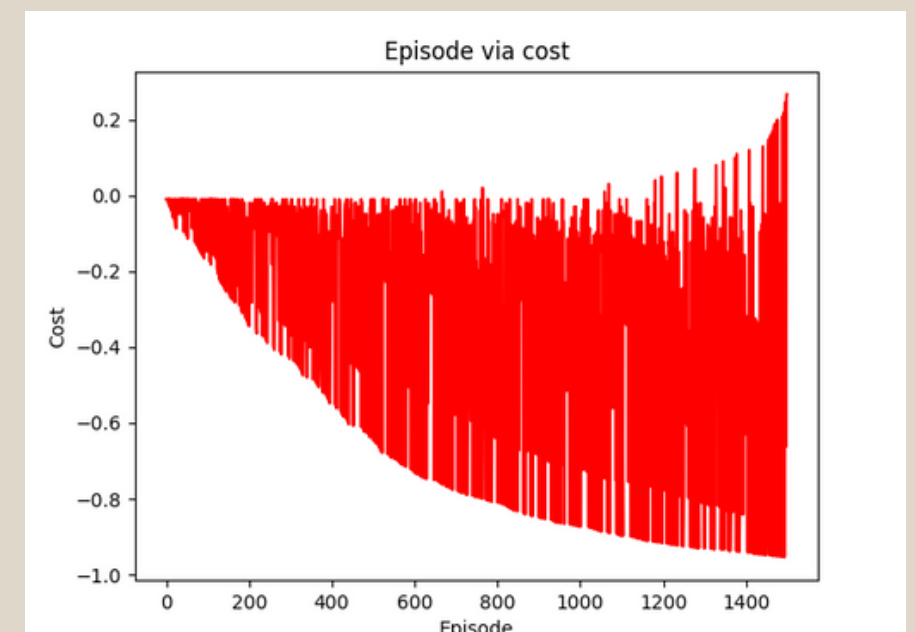
PERFORMA MODEL



OVERALL



EPS VS STEP



EPS VS COST



Analisa KESIMPULAN

1. Algoritma Q-Learning cocok diterapkan pada environment route bus ini yang memiliki ukuran 15x15. Akan tetapi, jika semakin besar ukuran environment maka algoritma ini tidak terlalu cocok karena akan membuat ukuran Q-Table semakin besar.
 2. Hasil jalur atau rute yang dihasilkan sudah optimal. Hal ini dibuktikan dengan jumlah step yang sedikit dan tidak melewati obstacle dari titik awal agent hingga mencapai goals.
 3. Jumlah episodes tidak terlalu mempengaruhi jumlah steps yang dibutuhkan
 4. Semakin banyak episode maka cost yang dihasilkan minus atau negatif. Akan tetapi, saat episode-episode akhir cost yang dihasilkan semakin naik.
-

it's the end
THANK YOU