AI COMP9414 ASSIGNMENT 2
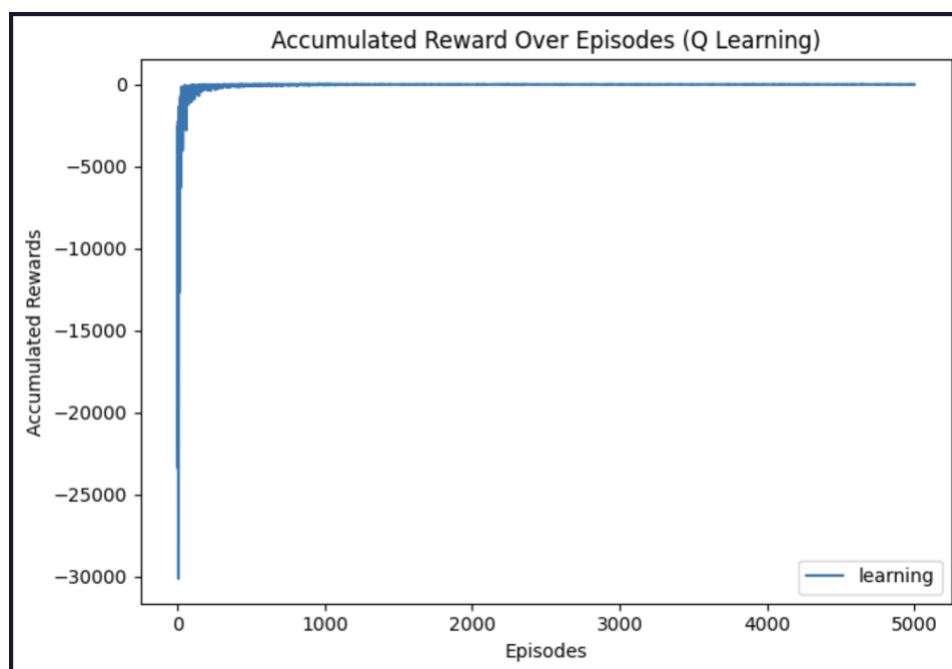# Tutor: Jingying Gao

Name - Aayushmaan Hooda
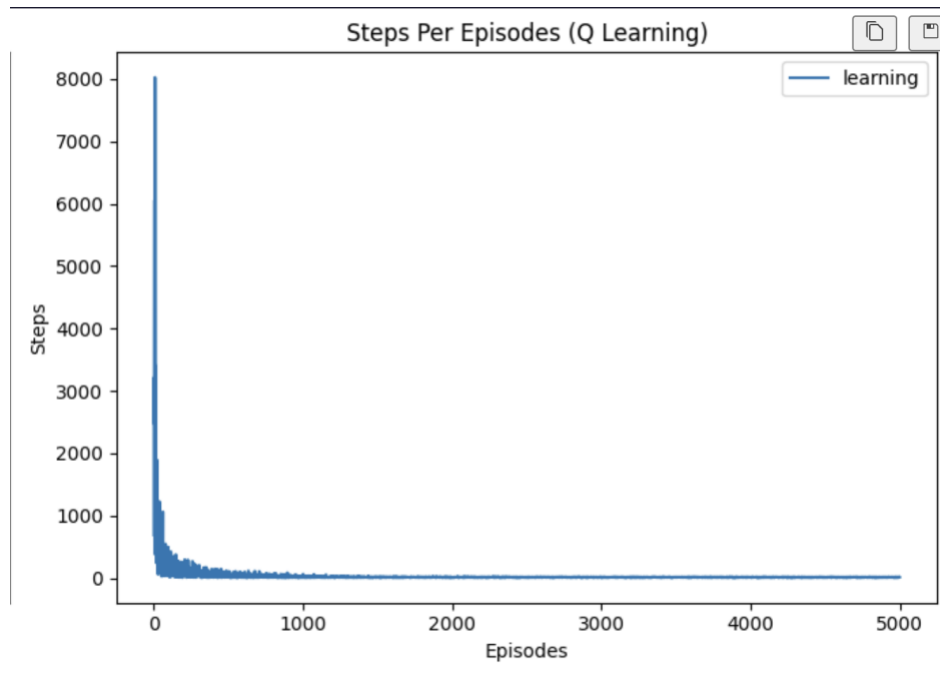**Zid - Z5518413**

Q-Learning Training

1. **Accumulated Reward over Episodes (Q-learning Graph)**

2.Steps Per Episodes (q-learning)

Q-Learning Testing

```
+---------+
|R: | : :G|
| : | : : |
| : :█: : |
| | : | : |
|Y| : |B: |
+---------+


Step 0
State: 221, Action = 3, Reward = -1, Total Accumulated Reward: -1


+---------+
|R: | : :G|
| : | : : |
| :█: : : |
| | : | : |
|Y| : |B: |
+---------+
  (West)


Step 1
State: 121, Action = 1, Reward = -1, Total Accumulated Reward: -2
```

```
+---------+
|R: | : :G|
| :█| : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)

Step 2
State: 101, Action = 3, Reward = −1, Total Accumulated Reward: −3

+---------+
|R: | : :G|
|█: | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
   (West)

Step 3
State: 1, Action = 1, Reward = −1, Total Accumulated Reward: −4

+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)

Step 4
State: 17, Action = 4, Reward = −1, Total Accumulated Reward: −5

+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (Pickup)

Step 5
State: 117, Action = 0, Reward = −1, Total Accumulated Reward: −6
```

```
+---------+
|R: | : :G|
|▊: | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (South)

Step 6
State: 217, Action = 0, Reward = -1, Total Accumulated Reward: -7


+---------+
|R: | : :G|
| : | : : |
|▊: : : : |
| | : | : |
|Y| : |B: |
+---------+
  (South)

Step 7
State: 237, Action = 2, Reward = -1, Total Accumulated Reward: -8

+---------+
|R: | : :G|
| : | : : |
| :▊: : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 8
State: 257, Action = 2, Reward = -1, Total Accumulated Reward: -9


+---------+
|R: | : :G|
| : | : : |
| : :▊: : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 9
State: 157, Action = 1, Reward = -1, Total Accumulated Reward: -10
```

```
+---------+
|R: | : :G|
| : |█: : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)


Step 10
State: 177, Action = 2, Reward = -1, Total Accumulated Reward: -11


+---------+
|R: | : :G|
| : | :█: |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)


Step 11
State: 197, Action = 2, Reward = -1, Total Accumulated Reward: -12

+---------+
|R: | : :G|
| : | : :█|
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 12
State: 97, Action = 1, Reward = -1, Total Accumulated Reward: -13


+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)


Step 13
State: 85, Action = 5, Reward = 20, Total Accumulated Reward: 7


+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (Dropoff)
```
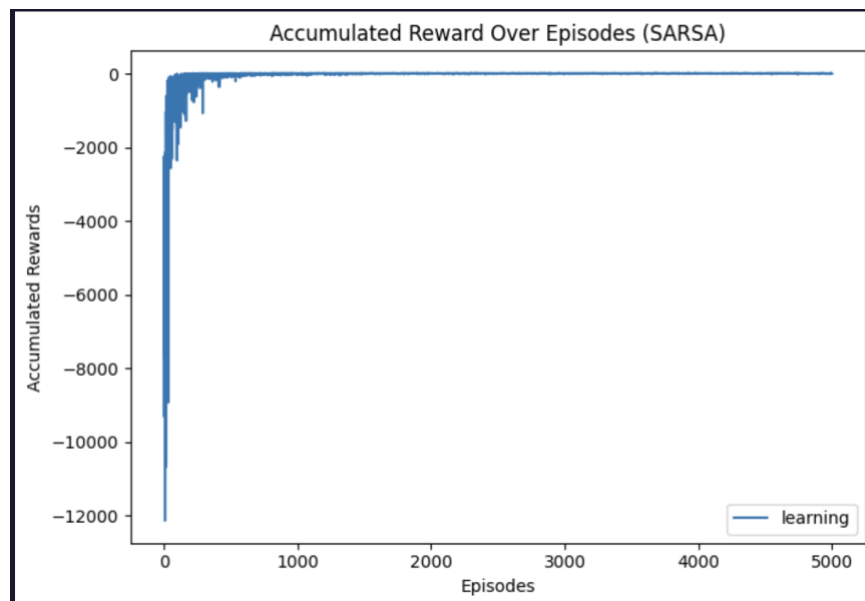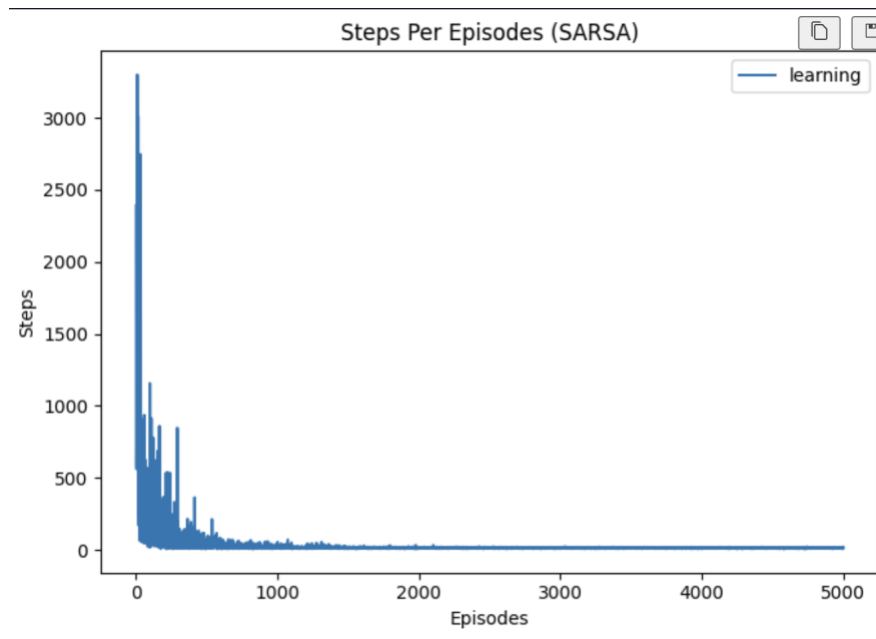
Q-Learning Averages Table

Algorithm	Mean Rewards	Mean Steps
Q-learning	8.03	12.97

SARSA Training

1. Accumulated Reward over Episodes (SARSA Graph)

2. Steps Per Episodes (q-learning)

Steps Per Episodes (SARSA)

SARSA Testing

```
+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+


Step 1
State: 1, Action = 4, Reward = -1, Total Accumulated Reward: -1

+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (Pickup)

Step 2
State: 17, Action = 0, Reward = -1, Total Accumulated Reward: -2
```

```
+---------+
|R: | : :G|
|█: | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (South)

Step 3
State: 117, Action = 2, Reward = -1, Total Accumulated Reward: -3

+---------+
|R: | : :G|
| :█| : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 4
State: 137, Action = 0, Reward = -1, Total Accumulated Reward: -4

+---------+
|R: | : :G|
| : | : : |
| :█: : : |
| | : | : |
|Y| : |B: |
+---------+
  (South)

Step 5
State: 237, Action = 2, Reward = -1, Total Accumulated Reward: -5

+---------+
|R: | : :G|
| : | : : |
| : :█: : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 6
State: 257, Action = 1, Reward = -1, Total Accumulated Reward: -6
```

```
+---------+
|R: | : :G|
| : |█: : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)

Step 7
State: 157, Action = 1, Reward = -1, Total Accumulated Reward: -7


+---------+
|R: |█: :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (North)

Step 8
State: 57, Action = 2, Reward = -1, Total Accumulated Reward: -8


+---------+
|R: | :█:G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 9
State: 77, Action = 2, Reward = -1, Total Accumulated Reward: -9


+---------+
|R: | : :█|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (East)

Step 10
State: 97. Action = 5. Reward = 20. Total Accumulated Reward: 11

+---------+
|R: | : :G|
| : | : : |
| : : : : |
| | : | : |
|Y| : |B: |
+---------+
  (Dropoff)
```

SARSA Averages Table

Algorithm	Mean Rewards	Mean Steps
SARSA	7.78	13.22