

# Reinforcement Learning

## Project Report

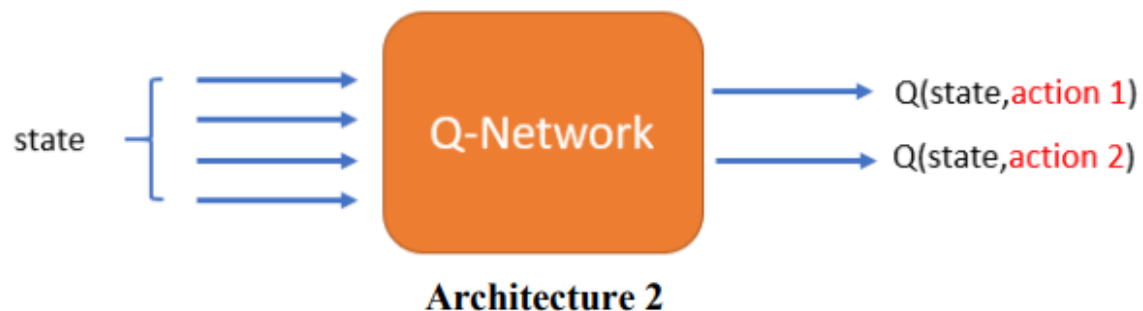
- By  
D Dhamodhar Reddy (IMT2019026)  
K S S V Naveen (IMT2019039)

# Introduction

---

Aim: Build an RL-based algorithm which can help cab drivers maximize their profits by improving their decision-making process on the field.

We used Deep Q Network of architecture type - 2.



Architecture 2 is preferable for Environments with small action space.

In the current environment, the length of the action space is 21.

## Updated Time Matrix

Time matrix contained a few zero entries in it. Those entries were replaced with the average value of the time matrix.

# Model

---

Initial State is chosen as random at the start of each episode.

## Hyperparameters used:

Learning rate = 0.05

Epsilon decay = 0.99999 (Multiplied this decay value to the epsilon for each episode)

Epsilon\_min = 0.1

Epsilon\_max = 1

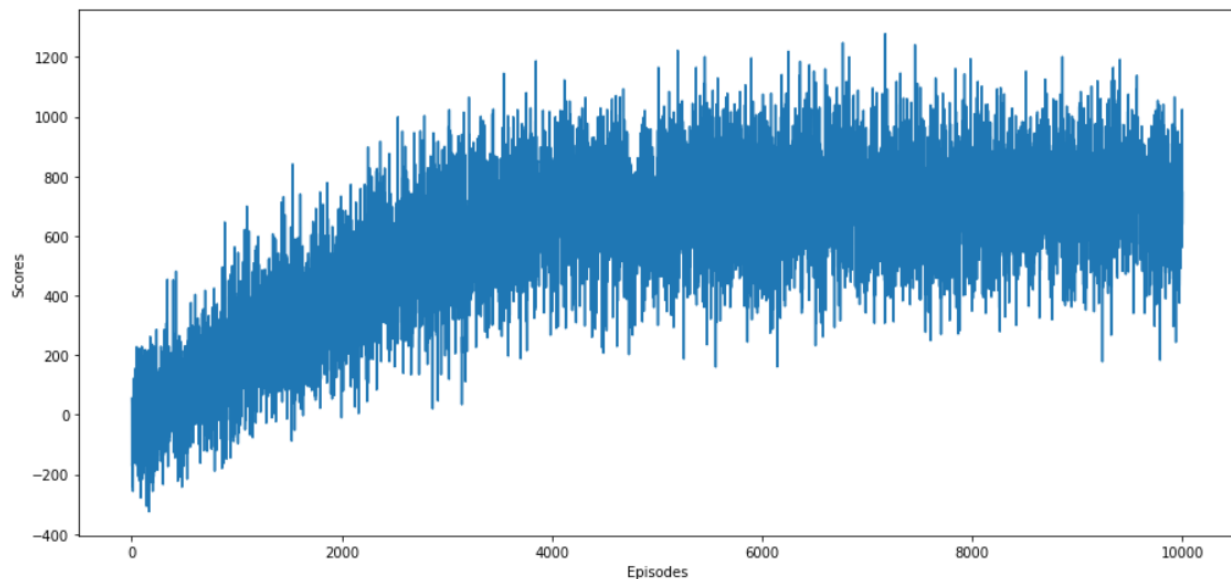
Batch\_size = 32

Queue length = 2000

## Convergence Plots

---

### Plot of Scores vs Episodes -



Average Score (for the last 2000 episodes) = 733.078

# Sample Episodes

---

Generated a sample episode with initial state = [3,6,0].

Current State: [3, 6, 0] Possible Actions: [[4, 1], [3, 1], [2, 4], [1, 2], [2, 0], [4, 3], [1, 3], [3, 2], [1, 4], [0, 3], [0, 4], [2, 1], [1, 0], [0, 0], [0, 0]] Action Chosen: 10 [3, 6, 0] Reward: 35.0

Current State: [1, 17, 0] Possible Actions: [[3, 1], [0, 3], [4, 3], [2, 1], [4, 0], [3, 0], [3, 2], [4, 1], [0, 0], [0, 0]] Action Chosen: 13 [1, 17, 0] Reward: -12.0

Current State: [0, 23, 0] Possible Actions: [[2, 4], [3, 2], [0, 0], [0, 0]] Action Chosen: 12 [0, 23, 0] Reward: -30.0

Current State: [4, 5, 1] Possible Actions: [[3, 0], [0, 0], [0, 0]] Action Chosen: 13 [4, 5, 1] Reward: 19.0

Current State: [0, 12, 1] Possible Actions: [[0, 2], [4, 0], [2, 4], [0, 0], [0, 0]] Action Chosen: 12 [0, 12, 1] Reward: -1.0

Current State: [4, 14, 1] Possible Actions: [[3, 0], [0, 4], [2, 0], [3, 2], [3, 4], [1, 3], [4, 3], [2, 4], [0, 0], [0, 0]] Action Chosen: 16 [4, 14, 1] Reward: 24.0

Current State: [4, 20, 1] Possible Actions: [[1, 4], [4, 0], [2, 3], [3, 0], [0, 2], [4, 2], [2, 4], [4, 1], [1, 0], [1, 2], [0, 0], [0, 0]] Action Chosen: 11 [4, 20, 1] Reward: 6.0

Current State: [3, 2, 2] Possible Actions: [[4, 2], [2, 3], [1, 3], [0, 0], [0, 0]] Action Chosen: 11 [3, 2, 2] Reward: 12.0

Current State: [3, 5, 2] Possible Actions: [[3, 2], [2, 0], [4, 0], [3, 1], [4, 3], [3, 4], [0, 3], [1, 2], [2, 1], [0, 2], [0, 0], [0, 0]] Action Chosen: 10 [3, 5, 2] Reward: -5.0

Current State: [1, 15, 2] Possible Actions: [[4, 0], [0, 2], [1, 4], [2, 0], [4, 2], [0, 3], [3, 4], [1, 2], [3, 1], [4, 3], [2, 1], [1, 0], [2, 4], [0, 1], [3, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 15, 2] Reward: 4.0

Current State: [1, 16, 2] Possible Actions: [[0, 3], [3, 1], [0, 2], [2, 3], [0, 1], [4, 0], [2, 1], [1, 4], [4, 1], [2, 4], [1, 2], [4, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 16, 2] Reward: 4.0

Current State: [1, 17, 2] Possible Actions: [[2, 1], [2, 0], [0, 2], [1, 3], [2, 3], [0, 1], [3, 0], [2, 4], [3, 1], [3, 4], [4, 3], [4, 2], [4, 1], [4, 0], [0, 3], [0, 0], [0, 0]] Action Chosen: 10 [1, 17, 2] Reward: 4.0

Current State: [1, 18, 2] Possible Actions: [[3, 2], [1, 4], [1, 0], [0, 3], [0, 1], [2, 3], [3, 4], [1, 2], [0, 2], [0, 0], [0, 0]] Action Chosen: 1 [1, 18, 2] Reward: 24.0

Current State: [1, 0, 3] Possible Actions: [[1, 2], [4, 0], [2, 1], [1, 0], [3, 2], [2, 3], [0, 3], [4, 1], [3, 0], [0, 1], [0, 2], [0, 4], [1, 4], [2, 0], [3, 4], [0, 0], [0, 0]] Action Chosen: 10 [1, 0, 3] Reward: 28.0

Current State: [1, 7, 3] Possible Actions: [[1, 2], [3, 4], [1, 0], [2, 1], [0, 1], [3, 2], [0, 4], [0, 2], [4, 0], [2, 0], [2, 3], [0, 3], [4, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 7, 3] Reward: 40.0

Current State: [1, 17, 3] Possible Actions: [[3, 1], [1, 4], [2, 1], [0, 2], [2, 0], [0, 3], [4, 3], [1, 2], [2, 4], [2, 3], [3, 2], [3, 0], [4, 1], [1, 3], [4, 0], [0, 0], [0, 0]] Action Chosen: 7 [1, 17, 3] Reward: -4.0

Current State: [3, 1, 4] Possible Actions: [[0, 2], [2, 0], [0, 1], [2, 1], [0, 0], [0, 0]] Action Chosen: 10 [3, 1, 4] Reward: 3.0

Current State: [1, 13, 4] Possible Actions: [[1, 3], [0, 1], [1, 2], [3, 0], [3, 2], [4, 3], [0, 3], [2, 3], [1, 4], [3, 4], [4, 2], [1, 0], [0, 4], [3, 1], [2, 4], [0, 0], [0, 0]] Action Chosen: 11 [1, 13, 4] Reward: 27.0

Current State: [3, 22, 4] Possible Actions: [[0, 1], [1, 0], [3, 0], [2, 3], [3, 2], [2, 1], [0, 2], [1, 3], [0, 0], [0, 0]] Action Chosen: 10 [3, 22, 4] Reward: 25.0

Current State: [1, 11, 5] Possible Actions: [[0, 2], [1, 2], [3, 1], [1, 0], [1, 3], [2, 0], [3, 2], [0, 3], [1, 4], [4, 3], [4, 0], [0, 1], [4, 2], [0, 4], [3, 0], [0, 0], [0, 0]] Action Chosen: 1 [1, 11, 5] Reward: 20.0

Current State: [1, 16, 5] Possible Actions: [[2, 3], [2, 1], [3, 0], [2, 4], [3, 2], [1, 3], [4, 3], [3, 1], [4, 2], [3, 4], [0, 3], [0, 2], [1, 0], [1, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 16, 5] Reward: 8.0

Current State: [1, 18, 5] Possible Actions: [[4, 0], [2, 1], [3, 4], [2, 3], [0, 4], [2, 0], [1, 2], [0, 3], [3, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 18, 5] Reward: 24.0

Current State: [1, 0, 6] Possible Actions: [[1, 0], [2, 4], [2, 3], [2, 0], [3, 4], [0, 4], [1, 2], [4, 0], [1, 4], [3, 0], [3, 2], [4, 3], [0, 3], [1, 3], [2, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 0, 6] Reward: 28.0

Current State: [1, 7, 6] Possible Actions: [[2, 3], [3, 4], [3, 2], [3, 0], [3, 1], [2, 4], [0, 4], [4, 2], [4, 1], [1, 4], [1, 2], [2, 0], [4, 3], [1, 0], [0, 0], [0, 0]] Action Chosen: 11 [1, 7, 6] Reward: 4.0

Current State: [3, 17, 6] Possible Actions: [[2, 0], [4, 2], [0, 1], [0, 0], [0, 0]] Action Chosen: 1 [3, 17, 6] Reward: 8.0

Current State: [1, 19, 6] Possible Actions: [[0, 1], [3, 4], [2, 3], [3, 1], [0, 3], [0, 4], [0, 0], [0, 0]] Action Chosen: 11 [1, 19, 6] Reward: 8.0

Current State: [3, 21, 6] Possible Actions: [[4, 2], [3, 1], [1, 3], [0, 0], [0, 0]] Action Chosen: 7 [3, 21, 6] Reward: 0.0

Current State: [3, 21, 6] Possible Actions: [[4, 2], [0, 1], [1, 4], [1, 0], [4, 1], [3, 1], [0, 3], [4, 3], [3, 0], [3, 4], [2, 4], [0, 0], [0, 0]] Action Chosen: 1 [3, 21, 6] Reward: 8.0

Current State: [1, 23, 6] Possible Actions: [[1, 2], [1, 0], [3, 2], [3, 0], [2, 3], [4, 1], [4, 0], [1, 4], [2, 1], [4, 2], [4, 3], [0, 3], [3, 4], [2, 0], [0, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 23, 6] Reward: 32.0

Current State: [1, 7, 0] Possible Actions: [[0, 2], [1, 2], [2, 0], [0, 3], [4, 0], [4, 2], [0, 4], [0, 1], [3, 1], [2, 3], [3, 2], [1, 0], [3, 0], [1, 4], [3, 4], [0, 0], [0, 0]] Action Chosen: 6 [1, 7, 0] Reward: -10.0

Current State: [2, 3, 1] Possible Actions: [[2, 0], [0, 1], [0, 0], [0, 0]] Action Chosen: 1 [2, 3, 1] Reward: -18.0

Current State: [1, 12, 1] Possible Actions: [[3, 0], [2, 1], [2, 0], [0, 3], [4, 3], [1, 3], [4, 0], [3, 1], [0, 4], [2, 3], [0, 0], [0, 0]] Action Chosen: 10 [1, 12, 1] Reward: 12.0

Current State: [1, 15, 1] Possible Actions: [[3, 1], [1, 4], [1, 3], [4, 3], [2, 0], [0, 1], [4, 1], [2, 3], [4, 0], [0, 4], [3, 2], [2, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 15, 1] Reward: 12.0

Current State: [1, 18, 1] Possible Actions: [[4, 1], [2, 1], [1, 4], [4, 3], [1, 2], [1, 3], [4, 0], [4, 2], [2, 3], [0, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 18, 1] Reward: 44.0

Current State: [1, 5, 2] Possible Actions: [[1, 3], [4, 1], [0, 3], [0, 4], [4, 3], [2, 4], [2, 3], [4, 2], [3, 0], [1, 4], [1, 2], [2, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 5, 2] Reward: 20.0

Current State: [1, 10, 2] Possible Actions: [[1, 3], [0, 1], [2, 4], [1, 4], [2, 0], [3, 2], [4, 3], [4, 2], [1, 2], [0, 3], [3, 1], [0, 0], [0, 0]] Action Chosen: 1 [1, 10, 2] Reward: 36.0

Current State: [1, 19, 2] Possible Actions: [[3, 0], [2, 0], [0, 1], [2, 1], [1, 2], [3, 2], [1, 3], [3, 1], [4, 0], [1, 4], [4, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 19, 2] Reward: 28.0

Current State: [1, 2, 3] Possible Actions: [[4, 1], [0, 1], [2, 3], [4, 3], [0, 4], [4, 0], [2, 4], [1, 4], [1, 0], [3, 2], [2, 1], [0, 0], [0, 0]] Action Chosen: 18 [1, 2, 3] Reward: 8.0

Current State: [1, 4, 3] Possible Actions: [[3, 1], [4, 1], [3, 4], [2, 4], [4, 3], [0, 2], [3, 2], [0, 4], [2, 3], [1, 3], [1, 2], [3, 0], [0, 0], [0, 0]] Action Chosen: 11 [1, 4, 3] Reward: 5.0

Current State: [3, 12, 3] Possible Actions: [[2, 4], [2, 1], [0, 3], [1, 2], [4, 1], [3, 4], [4, 0], [3, 1], [0, 0], [0, 0]] Action Chosen: 14 [3, 12, 3] Reward: -4.0

Current State: [1, 20, 3] Possible Actions: [[1, 0], [3, 1], [2, 1], [4, 0], [4, 2], [0, 1], [1, 3], [2, 0], [0, 4], [1, 4], [4, 3], [4, 1], [0, 3], [3, 2], [3, 0], [0, 0], [0, 0]] Action Chosen: 10 [1, 20, 3] Reward: 32.0

Current State: [1, 4, 4] Possible Actions: [[0, 1], [2, 1], [0, 3], [2, 3], [2, 0], [1, 4], [1, 0], [1, 2], [3, 0], [3, 1], [0, 2], [4, 3], [1, 3], [4, 1], [2, 4], [0, 0], [0, 0]] Action Chosen: 18 [1, 4, 4] Reward: 12.0

Current State: [1, 7, 4] Possible Actions: [[0, 2], [3, 2], [1, 3], [1, 4], [0, 4], [3, 4], [3, 0], [1, 0], [4, 0], [2, 3], [2, 1], [0, 0], [0, 0]]

Action Chosen: 10 [1, 7, 4] Reward: 28.0

Current State: [1, 14, 4] Possible Actions: [[4, 3], [2, 4], [0, 4], [4, 0], [3, 1], [4, 2], [2, 0], [1, 0], [3, 4], [0, 2], [0, 1], [1, 3], [4, 1],

[0, 0], [0, 0]] Action Chosen: 1 [1, 14, 4] Reward: 12.0

Current State: [1, 17, 4] Possible Actions: [[2, 3], [3, 1], [4, 3], [1, 2], [0, 4], [0, 0], [0, 0]] Action Chosen: 11 [1, 17, 4]

Reward: 27.0

Current State: [3, 2, 5] Possible Actions: [[1, 3], [0, 3], [4, 0], [1, 2], [2, 3], [0, 0], [0, 0]] Action Chosen: 11 [3, 2, 5] Reward: 44.0

Current State: [3, 13, 5] Possible Actions: [[3, 1], [1, 4], [1, 2], [3, 4], [1, 0], [2, 0], [3, 0], [3, 2], [0, 4], [1, 3], [4, 1], [0, 1], [4, 2],

[0, 3], [0, 0], [0, 0]] Action Chosen: 1 [3, 13, 5] Reward: 6.0

Current State: [1, 19, 5] Possible Actions: [[4, 2], [2, 3], [0, 4], [4, 0], [2, 1], [0, 3], [3, 4], [4, 1], [1, 4], [0, 1], [3, 2], [0, 0], [0, 0]]

Action Chosen: 10 [1, 19, 5] Reward: 24.0

Current State: [1, 1, 6] Possible Actions: [[3, 4], [4, 2], [0, 1], [4, 0], [3, 1], [1, 4], [4, 1], [2, 4], [3, 2], [4, 3], [1, 2], [0, 3], [1, 3], [0,

0], [0, 0]] Action Chosen: 1 [1, 1, 6] Reward: 24.0

Current State: [1, 7, 6] Possible Actions: [[1, 4], [2, 1], [2, 0], [0, 3], [3, 2], [3, 1], [3, 4], [4, 3], [0, 4], [1, 0], [0, 2], [4, 2], [0, 1], [4,

0], [0, 0], [0, 0]] Action Chosen: 10 [1, 7, 6] Reward: 16.0

Current State: [1, 11, 6] Possible Actions: [[2, 0], [4, 3], [1, 0], [3, 4], [0, 1], [1, 3], [4, 2], [1, 2], [0, 3], [0, 4], [1, 4], [2, 1], [3, 1],

[2, 3], [4, 1], [0, 0], [0, 0]] Action Chosen: 10 [1, 11, 6] Reward: 16.0

Current State: [1, 15, 6] Possible Actions: [[0, 2], [2, 3], [2, 1], [4, 0], [2, 4], [4, 3], [0, 3], [1, 4], [4, 1], [1, 3], [0, 0], [0, 0]] Action

Chosen: 10 [1, 15, 6] Reward: 8.0

Current State: [1, 17, 6] Possible Actions: [[2, 3], [0, 4], [1, 2], [4, 3], [0, 1], [3, 4], [2, 0], [0, 2], [2, 4], [3, 1], [1, 0], [4, 2], [0, 0],

[0, 0]] Action Chosen: 11 [1, 17, 6] Reward: 32.0

Current State: [3, 1, 0] Possible Actions: [[2, 1], [3, 4], [0, 1], [0, 0], [0, 0]] Action Chosen: 10 [3, 1, 0] Reward: -5.0

Current State: [1, 11, 0] Possible Actions: [[2, 3], [3, 4], [1, 2], [0, 3], [2, 0], [4, 1], [0, 1], [1, 4], [3, 2], [0, 4], [4, 3], [3, 1], [0, 0],

[0, 0]] Action Chosen: 14 [1, 11, 0] Reward: 4.0

Current State: [1, 12, 0] Possible Actions: [[0, 1], [3, 2], [3, 1], [4, 1], [0, 2], [1, 4], [4, 2], [4, 0], [2, 4], [3, 0], [1, 3], [4, 3], [2, 1],

[1, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 12, 0] Reward: 4.0

Current State: [1, 13, 0] Possible Actions: [[2, 3], [4, 3], [2, 0], [1, 4], [3, 0], [3, 1], [4, 0], [1, 0], [0, 3], [2, 4], [0, 1], [4, 2], [4, 1],

[0, 0], [0, 0]] Action Chosen: 11 [1, 13, 0] Reward: 31.0

Current State: [3, 23, 0] Possible Actions: [[2, 0], [0, 2], [0, 4], [4, 1], [0, 1], [1, 0], [0, 3], [1, 2], [0, 0], [0, 0]] Action Chosen: 1

[3, 23, 0] Reward: 3.0

Current State: [1, 2, 1] Possible Actions: [[0, 1], [2, 1], [4, 0], [1, 3], [0, 2], [1, 0], [4, 2], [0, 0], [0, 0]] Action Chosen: 10 [1, 2,

1] Reward: 24.0

**Total Reward = 818.0**

## Trajectory Observed for the above Episode -

(state -> action\_index -> reward -> next\_state)

[3, 6, 0], 10, 35.0, [1, 17, 0]-----> [1, 17, 0], 13, -12.0, [0, 23, 0]-----> [0, 23, 0], 12, -30.0, [4, 5, 1]----->  
[4, 5, 1], 13, 19.0, [0, 12, 1]-----> [0, 12, 1], 12, -1.0, [4, 14, 1]-----> [4, 14, 1], 16, 24.0, [4, 20, 1]----->  
[4, 20, 1], 11, 6.0, [3, 2, 2]-----> [3, 2, 2], 11, 12.0, [3, 5, 2]-----> [3, 5, 2], 10, -5.0, [1, 15, 2]----->  
[1, 15, 2], 10, 4.0, [1, 16, 2]-----> [1, 16, 2], 10, 4.0, [1, 17, 2]-----> [1, 17, 2], 10, 4.0, [1, 18, 2]----->  
[1, 18, 2], 1, 24.0, [1, 0, 3]-----> [1, 0, 3], 10, 28.0, [1, 7, 3]-----> [1, 7, 3], 10, 40.0, [1, 17, 3]----->  
[1, 17, 3], 7, -4.0, [3, 1, 4]-----> [3, 1, 4], 10, 3.0, [1, 13, 4]-----> [1, 13, 4], 11, 27.0, [3, 22, 4]----->  
[3, 22, 4], 10, 25.0, [1, 11, 5]-----> [1, 11, 5], 1, 20.0, [1, 16, 5]-----> [1, 16, 5], 10, 8.0, [1, 18, 5]----->  
[1, 18, 5], 10, 24.0, [1, 0, 6]-----> [1, 0, 6], 10, 28.0, [1, 7, 6]-----> [1, 7, 6], 11, 4.0, [3, 17, 6]----->  
[3, 17, 6], 1, 8.0, [1, 19, 6]-----> [1, 19, 6], 11, 8.0, [3, 21, 6]-----> [3, 21, 6], 7, 0.0, [3, 21, 6]----->  
[3, 21, 6], 1, 8.0, [1, 23, 6]-----> [1, 23, 6], 10, 32.0, [1, 7, 0]-----> [1, 7, 0], 6, -10.0, [2, 3, 1]----->  
[2, 3, 1], 1, -18.0, [1, 12, 1]-----> [1, 12, 1], 10, 12.0, [1, 15, 1]-----> [1, 15, 1], 10, 12.0, [1, 18, 1]----->  
[1, 18, 1], 10, 44.0, [1, 5, 2]-----> [1, 5, 2], 10, 20.0, [1, 10, 2]-----> [1, 10, 2], 1, 36.0, [1, 19, 2]----->  
[1, 19, 2], 10, 28.0, [1, 2, 3]-----> [1, 2, 3], 18, 8.0, [1, 4, 3]-----> [1, 4, 3], 11, 5.0, [3, 12, 3]----->  
[3, 12, 3], 14, -4.0, [1, 20, 3]-----> [1, 20, 3], 10, 32.0, [1, 4, 4]-----> [1, 4, 4], 18, 12.0, [1, 7, 4]----->  
[1, 7, 4], 10, 28.0, [1, 14, 4]-----> [1, 14, 4], 1, 12.0, [1, 17, 4]-----> [1, 17, 4], 11, 27.0, [3, 2, 5]----->  
[3, 2, 5], 11, 44.0, [3, 13, 5]-----> [3, 13, 5], 1, 6.0, [1, 19, 5]-----> [1, 19, 5], 10, 24.0, [1, 1, 6]----->  
[1, 1, 6], 1, 24.0, [1, 7, 6]-----> [1, 7, 6], 10, 16.0, [1, 11, 6]-----> [1, 11, 6], 10, 16.0, [1, 15, 6]----->  
[1, 15, 6], 10, 8.0, [1, 17, 6]-----> [1, 17, 6], 11, 32.0, [3, 1, 0]-----> [3, 1, 0], 10, -5.0, [1, 11, 0]-----> [  
1, 11, 0], 14, 4.0, [1, 12, 0]-----> [1, 12, 0], 10, 4.0, [1, 13, 0]-----> [1, 13, 0], 11, 31.0, [3, 23, 0]----->  
[3, 23, 0], 1, 3.0, [1, 2, 1]-----> [1, 2, 1], 10, 24.0, [1, 8, 1]-----> [1, 8, 1]