Q-Learning Agent

Reward System:

- 0.01 – For surviving
- +5 if previous distance between head of snake and fruit is less that current distance between head of snake and fruit and -1 if opposite
- -1000 for hitting in wall and -800 for hitting on itself
- 200 + 100 * score, for getting fruit (score will be incremented by one if fruit is acquired)

State:

- Parameter 1 : Danger Ahead (1 if body or wall is right ahead(just next cell), 0 otherwise)
- Parameter 2 : Danger Ahead (1 if body or wall is in right(next cell), 0 otherwise)
- Parameter 3 : Danger Ahead (1 if body or wall is in left(next cell), 0 otherwise)
- Parameter 4 : 0 if fruit is in -x dirn, 2 if fruit is in x direction and 1 otherwise
- Parameter 5 : Similarly for y

So total number of states = 2 * 2 * 2 * 3 * 3 = 72

Actions :

4 Actions – (1,0) direction, (0,1) direction, (-1,0) direction and (0, -1) direction. And can't go -180 degree direction.

DEEP Q Learning Agent

Reward System and Actions were defined same as Q Learning Agent.

Implemented a Linear Neural Network. The input neurons were a grid of [cell number][cell number] elements in which the value is 10 if snake's head is in that cell 10 to 2 in a linear decrement if snake's body is in that cell, 1 for fruit and 0 otherwise (I did not use negative values as I used ReLU as activation function). This input is flattened and then fed into Linear Neural Network, with layers 900(input) → 512 → 128 → 4(action space)