

## **Deep Reinforcement Learning Nanodegree**

### **Project 1 – Navigation**

In this project, I have trained a simple DQN for an agent which picks bananas.

According to the task, rewards for the agent are 1 if he picks yellow banana and -1 if he picks blue banana. For this task I have used plain vanilla DQN without any modifications. State space is represented by 37-dimension vectors of agent's velocity and its perception of the surroundings.

Both target neural networks and online neural network consists of input state vector, 2 hidden layers with 64 neurons and output layer which represent 4 actions.

For optimization, I have used Adam optimizer which optimized minimal squared error between target Q-value and expected Q-value. Experience replay along with epsilon greedy action selection were implemented as well.

In this project I used hyperparameters which were tuned to achieve a good performance of DQN.

**BUFFER\_SIZE** – the parameter to control the size of double ended queue for experience replay strategy. In this case I set it to 10000, so it is rather large collection of previous states and actions made by DQN.

**BATCH\_SIZE** is the minimum quantity of past experiences made by DQN agent when he is allowed to use random samples from them. I set it to 64.

Another hyperparameter is **GAMMA**=0.99, is the discount factor for the reward made in future.

**TAU**=0.001 is soft update parameter to update target network once it is needed the update. The update rule is following:  $(1 - \text{TAU}) * \text{local\_network} + \text{TAU} * \text{target\_network}$ . So in this case we have

small difference after update between local and target networks so local\_network is still allowed to continue learning.

**LR**=0.0005 is learning rate parameter for my Pytorch model.

**UPDATE EVERY**=4 is the number of timesteps when agent updates its target values after running the optimization model.

This strategy can be improved by Dueling DQN, Double DQN and prioritized replay technique. I plan to implement those changes along with pixel DQN.