

### 1) Introduction

- The goal is to get an average score of +13 over 100 consecutive episodes.

### 2) Learning Algorithm

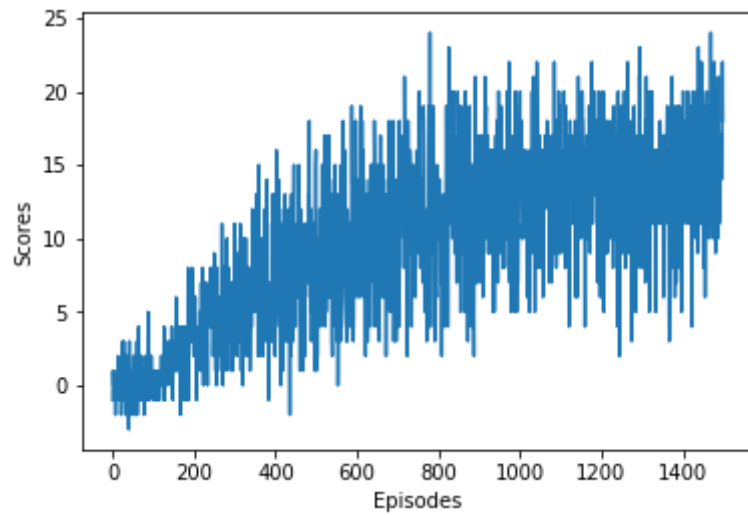
- Deep Q Network with replay memory
- Double deep q network
- Dueling network architecture

### 3) Parameters

- Common layer: 2 Hidden fully-connected layers(150, 150 units)
- Two streams: 1 Hidden fully-connected layer(75 units)
- Learning rate:  $1e-3$
- Gamma: 0.95
- Hard update frequency:  $1e4$
- Soft update( $\tau$ ):  $1e-3$
- Memory size:  $1e6$
- Batch size: 32
- Optimizer: Adam
- Exploration stop:  $1e-2$
- Exploration decay rate:  $1e-3$

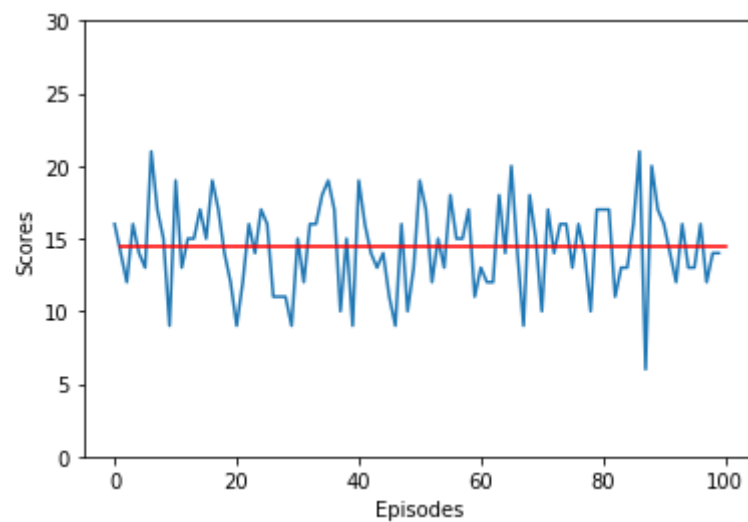
### 4) Training

- Training episodes: 1499
- The agent was trained to get an average score of +15



## 5) Results

- Average score: 14.42



## 6) Future work

- To improve speed of learning, Actor-Critic Methods will be implemented
- Also, Categorical DQN will be applied