# DQN

Quiz, 6 questions

1 point

1.

DQN

- [ ] Is based on an on-policy method.
- [ ] Neither off-policy nor on-policy method
- [x] Is based on an off-policy method.

---

1 point

2.

How can we speed up the DQN training process?

- [ ] Use upsampling.
- [ ] Stack 4 last states together.
- [x] Parallelize learning.
- [ ] Stack more layers.
- [x] Use smaller input.
- [x] Do not use pooling.
- [x] Use strides.
- [ ] Use stacked state and action as input.
- [x] Use smaller model.

---

1

point

3.

Why do we need to stack four last frames while playing Atari games?

- [ ] Without such stacking a neural network will not learn (due to its depth).
- [ ] It stabilises training.
- [ ] At least two stacked frames is the only type of input accepted by a convolutional layer.
- [x] It makes Atari game fully observable.
- [ ] It allows to obtain additional training data.

---

1 point

4.

How can we help an agent to adapt to different scales of a reward signal?

- [x] Adjust the agent learning rate.
- [x] Give the agent more training samples with lower learning rate.
- [ ] Use moving average of last N rewards as a new reward.
- [ ] Give the agent more training samples with same learning rate.
- [x] Adjust the reward signal gamma.
- [ ] Give the agent more training samples with higher learning rate.

---

1 point

5.

Why target networks are commonly used?

- [x] It is a heuristic against the problem of correlated training data.
- [ ] Target networks use smaller amount of parameters.
- [ ] Target networks work faster.
- [ ]

- [ ] They increase training stability in tasks, where the reward scale is not known beforehand.
- [x] For training stabilization.
- [ ] They are mandatory when using an experience replay.

---

1 point

6.

What are the possible ways to improve experience replay (buffer)?

- [ ] Add more statistics to it, such as gamma, current learning rate, current step/epoch.
- [ ] Sample whole sessions from an experience replay instead of single experience (S,A,R,S') tuples.
- [ ] Use smaller buffer (up to 1 thousand).
- [x] Use samples more wisely – taking into account their contribution to the training process.
- [x] Use larger buffer (up to 1 million).

---

- [ ] I, **Jiadai Zhao**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Submit Quiz

---