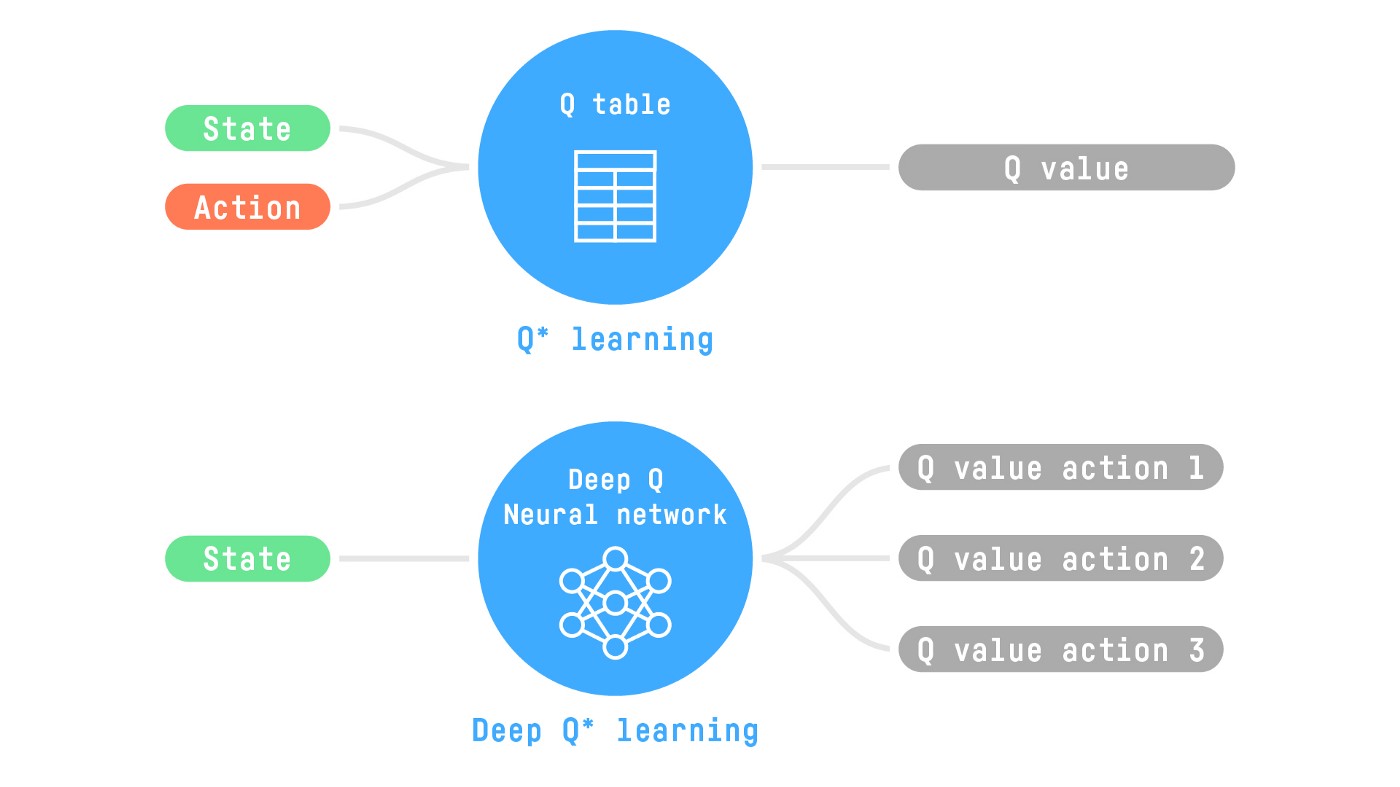
Deep Q-Learning Algorithm

Instead of using a Q-table, Deep Q-Learning uses a Neural Network that takes a state and approximates Q-values for each action based on that state.



The Deep Q-Learning training algorithm has two phases:

1. **Sampling**: we perform actions and store the observed experience tuples in a replay memory.
2. **Training**: Select a small batch of tuples randomly and learn from this batch using a gradient descent update step.

Deep Q-Learning training might suffer from instability, to help us stabilize the training, we implement three different solutions:

1. Experience Replay to make more efficient use of experiences.
2. Fixed Q-Target to stabilize the training.
3. Double Deep Q-Learning, to handle the problem of the overestimation of Q-values.