Reinforcement Learning

Deep Q-Network

Dr. Alireza Aghamohammadi

Deep Q-Network (DQN)

- The goal of the Deep Q-Network (DQN) is to approximate the optimal action-value function $q(s, a)$.
- The DQN architecture uses a neural network with:
  - An input layer with the same size as the state space.
  - An output layer with the same size as the action space.
- The Bellman optimality equation for $q(s, a)$ is given as:

$$q(s_t, a_t) = r(s_t, a_t) + \gamma \cdot \max_a q(s_{t+1}, a),$$

  where:
  - $r(s_t, a_t)$ is the reward for taking action a_t in state s_t,
  - γ is the discount factor,
  - and $\max_a q(s_{t+1}, a)$ is the maximum expected future reward.
- The Bellman equation assumes that the policy π has converged to the optimal policy π^*.
- In practice, the difference between the predicted $q(s_t, a_t)$ and the target value is known as the **temporal difference error** δ.
- To minimize this error, we optimize the following loss function:

$$\mathcal{L}(\phi) = \left[r(s_t, a_t) + \gamma \cdot \max_a q(s_{t+1}, a; \phi) - q(s_t, a_t; \phi) \right]^2,$$

  where ϕ represents the parameters of the neural network.