

Aufgabenblatt 15

Abgabe: 15.03.2021

Aufgabe 1.

Implementiere einen Deep-Q-Algorithmus für Taxi-v3. Auf <https://www.learndatasci.com/tutorials/reinforcement-q-learning-scratch-python-openai-gym/> findest du nochmal ausführlich wie Q-Learning bei Taxi-v3 funktioniert (für die Intuition ist die Erklärung auf <https://en.wikipedia.org/wiki/Q-learning> hilfreich). Jedoch wird da die Funktion Q über eine Tabelle dargestellt (da es diskret ist) [den Code findest du hier <https://pastebin.com/BwB5UQNh>]. Implementiere also nun ein neuronales Netz, welches die Funktion

$$Q^* : \text{State} \times \text{Action} \rightarrow \mathbb{R}$$

approximiert. Minimiere dafür den Fehler

$$\delta = Q(s_t, a_t) - (r_t + \gamma \max_a Q(s_{t+1}, a)).$$

Schau dir z.B. https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html an, dort wird es noch einmal etwas genauer erklärt. Verwende hierfür auch ein Gedächtnis basiertes Training, d.h. minimiere nicht nach jedem Schritt, sondern trainiere zufällig auf bereits passiertten Aktionen. (Merke dir dafür den Ursprungs-State, die Aktion, den Reward und den resultierenden State.)