


Deep Q-Learning

Janot George, Maurice Borries

Hochschule für Technik und Wirtschaft Berlin
Studiengang Wirtschaftsmathematik (Bachelor)
Seminar

02.12.2021

Inhaltsübersicht

- 
- 1 Konzept
 - 2 Bestandteile
 - 3 Abbildungs- und Quellenverzeichnis

Inhaltsübersicht

- 
- 1 Konzept
 - 2 Bestandteile
 - 3 Abbildungs- und Quellenverzeichnis

Was ist (Deep) Q-Learning



- Algorithmus trainiert eine KI
- Zuckerbrot-und-Peitschen-Prinzip
- KI startet ohne Daten
- Konzept: Daten erzeugen und Funktionen live optimieren
- Ziel: Annäherung an optimale Lösung des Problems

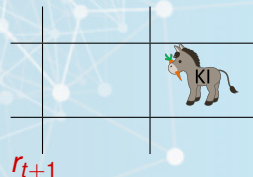
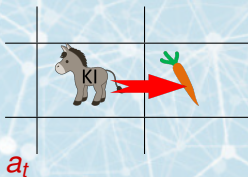
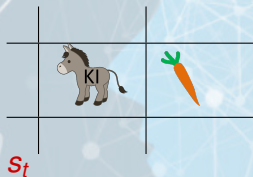
Inhaltsübersicht

- 
- 1 Konzept
 - 2 Bestandteile**
 - 3 Abbildungs- und Quellenverzeichnis

Markov-Entscheidungsprozess

Definitionen:

- Zustand der Umgebung zum Zeitpunkt t : s_t
- Aktion der KI (des Agenten) zum Zeitpunkt t : a_t
- Belohnung der Aktion a_t : r_{t+1}



Exploration versus Exploitation

Exploration

- Erkundung des Umfeldes
- zufällige Aktionen ausführen
- neue Daten sammeln

Exploitation

- Ausnutzung gesammelter Daten
- Daten überprüfen
- Daten aktualisieren

Optimale Anwendung der Erkundungsrate ϵ erfordert:

- ϵ_{max} : Erkundungsrate zu Beginn
- ϵ_{red} : Reduktionswert der Erkundungsrate
- ϵ_{min} : Minimale Erkundungsrate

$$\epsilon = \epsilon_{min} + \frac{\epsilon_{max} - \epsilon_{min}}{e^{\epsilon_{red} \cdot \text{Episodennummer}}}$$

Bellman-Gleichung

Aktionsbewertungsfunktion

$$Q(s_t, a_t) = \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \right]$$

Bellman-Gleichung

$$Q^*(s_t, a_t) = \mathbb{E} \left[r_{t+1} + \gamma \max_{a'} [Q^*(s_{t+1}, a')] \right]$$

Annäherung mit Lernrate α :

$$Q^\#(s_t, a_t) = (1 - \alpha) \cdot Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_{a'} [Q^*(s_{t+1}, a')] \right]$$

Inhaltsübersicht

- 
- 1 Konzept
 - 2 Bestandteile
 - 3 Abbildungs- und Quellenverzeichnis**

Abbildungsverzeichnis

- Algorithmus:
<https://www.allaboutlean.com/employee-motivation-1/carrot-and-stick/>; Zuletzt aufgerufen: 23.11.2021
- KI:
<https://de.cleanpng.com/png-b1ndvx/>; Zuletzt aufgerufen: 22.11.2021
- Hintergrund:
<https://it-talents.de/it-wissen/machine-learning-accuracy-und-precision/>; Zuletzt aufgerufen: 7.11.2021

Quellenverzeichnis

- Richter, S. (2019) Statistisches und maschinelles Lernen, Berlin, Springer.
- K.-L. Du and M. N. S. Swamy, Neural Networks and Statistical Learning.
- Ilyas, Agakishiev: METIS: Reinforcement Learning, Humboldt-Universität zu Berlin, IRTG1792.HU-Berlin.de
- <https://www.learndatasci.com/tutorials/reinforcement-q-learning-scratch-python-openai-gym/> , Satwik Kansal, Brendan Martin, zuletzt abgerufen: 7.12.2021
- <https://ichi.pro/de/einfuehrung-in-das-reinforcement-learning-markov-entscheidungsprozess-75613770348762> , ICHI.PRO, Laurenz Wuttke, zuletzt abgerufen: 7.12.2021
- <https://hci.iwr.uni-heidelberg.de/system/files/private/downloads/541645681/dammann-reinforcement-learning-report.pdf> , Patrick Dammann, zuletzt abgerufen: 7.12.2021
- <http://www.informatik.uni-ulm.de/ni/Lehre/SS05/RL/vorlesung/rl03.pdf> , F. Schwenker, zuletzt abgerufen: 7.12.2021
- <https://deeplizard.com/> , Chris and Mandy, zuletzt abgerufen: 7.12.2021
- <https://datasolut.com/neuronale-netzwerke-einfuehrung/> , Laurenz Wuttke, zuletzt abgerufen: 7.12.2021
- <https://www.samyzaf.com/ML/rl/qmaze.html> , zuletzt abgerufen: 7.12.2021