

Maschinelles Lernen (Müller 2017), ein Teilgebiet der **künstlichen Intelligenz**, gewinnt auch in Unternehmen zunehmend an praktischer Bedeutung. Im Rahmen des **Smart Manufacturing** können Produktionsprozesse in Industrieunternehmen mit Hilfe von Machine-Learning-Algorithmen analysiert, besser verstanden und optimiert werden. Dies wiederum ermöglicht Kosteneinsparungen, Flexibilisierungen und auch eine höhere Qualität von Produkten. Weitere mögliche Einsatzfelder im Kontext von **Industrie 4.0** sind etwa das Auffinden bzw. Vermeiden von möglichen Fehlern in Maschinen (**Predictive Maintenance**), Entdeckung unbekannter Fehlerquellen in der Qualitätskontrolle, der Einsatz von Robotern und autonomen Maschinen in Fertigung und Logistik. Mit Hilfe von im Unternehmen vorliegenden Daten können die eingesetzten Algorithmen trainiert und so Muster erkannt werden und dadurch Vorhersagen erfolgen. Auch in der Wissenschaft findet maschinelles Lernen zunehmend mehr Einsatz. So gelang es Astrophysikern kürzlich, sog. Fast Radio Bursts (FRB), also extrem schnelle Radioblitze, in schwachen radioastronomischen Signalen zu entdecken (Zhang, 2018). Im Allgemeinen werden drei Kategorien des maschinellen Lernens unterschieden: Überwachtes Lernen (**Supervised Learning**), unüberwachtes Lernen (**Unsupervised Learning**) und **Reinforcement Learning** (Abb. 1). Das vorliegende Projekt nutzt das sog. **Deep Reinforcement Learning**, d.h. maschinelles Lernen durch Feedback der Umwelt in Form von Belohnungen. Derzeit wird Deep Reinforcement Learning im Wesentlichen am Beispiel von Spielsimulationen erforscht. Im Projekt werden mit Hilfe von Spielsimulationen praktische Problemszenarien entwickelt und beschrieben, die in der Industrie eine Rolle spielen können.

Reinforcement Learning

Bei Reinforcement Learning (Lapan 2018) handelt es sich um eine Optimierungsstrategie, indem ein **Agent** (z.B. eine Drohne) während des Trainings erlernt, erhaltene und zukünftige Belohnungen (R) durch Trial-and-Error zu maximieren (**Reward Hypothesis**) (Sutton 2018):

$$E_t = \sum_{k=0}^T \gamma^k R_{t+k+1} \quad \gamma \in [0, 1)$$

Ausblick

Durch maschinelles Lernen, insbesondere Deep Reinforcement Learning, können vielfältige autonome Verhaltensweisen erlernt werden, welche bereits schon jetzt in der Praxis eingesetzt werden könnten. Ein wichtiges Ziel ist es, das am Computer erlernte Verhalten in die reale Welt zu übertragen. Im weiteren Verlauf dieses Projektes sollen u.a. Kamerabeobachtungen hinzugezogen werden, wodurch beispielsweise Roboter oder Drohnen lernen können, durch Räume bzw. Gebäude zu navigieren.

Referenzen

- Müller, A. C.; Guido, S. (2017): *Einführung in Machine Learning mit Python: Praxiswissen Data Science*, O'Reilly Media, dpunkt.verlag, Heidelberg
- Lapan, M. (2018): *Deep Reinforcement Learning Hands-On: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more*, Packt Publishing, Birmingham
- Sutton, R. S.; Barto, A. G. (2018): *Reinforcement learning: An introduction*. MIT Press, 2nd edition.
- Zhang, Y. G.; Gajjar, V.; Foster, G.; Siemion, A.; Cordes, J.; Law, C.; Wang Yu (2018), *Fast Radio Burst 121102 Pulse Detection and Periodicity: A Machine Learning Approach*, submitted, [arXiv:1809.03043](https://arxiv.org/abs/1809.03043)

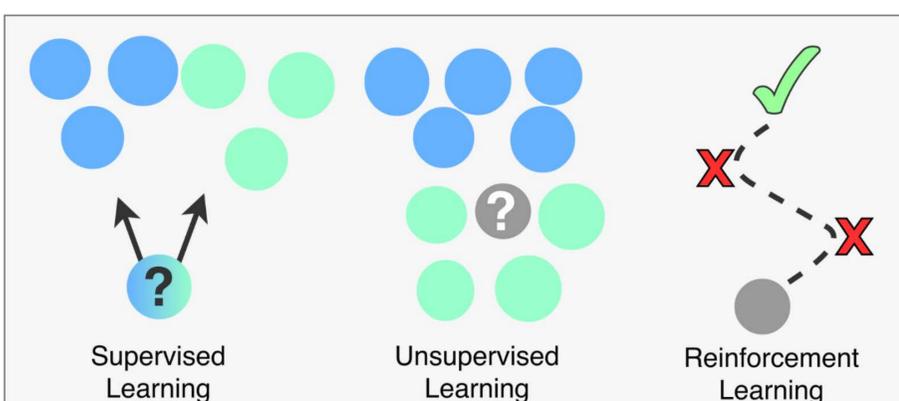


Abbildung 1: Drei Kategorien des maschinellen Lernens

Use-Case 1: Roboterarm spielt „Heißer Draht“ (Motorikschleife)

In dieser Umwelt wird ein Roboterarm gesteuert. Der Ring am Ende des Armes soll ohne Anstoßen bis zum Ende des Drahtes geführt werden.

Aktionen des Agenten (Abb. 2):

- Rotieren von 4 Gelenken des Roboterarms
- Verschiebung der horizontalen Position des Roboterarms

Observationen des Agenten (Abb. 2):

- Abtasten der Distanzen des Ringes zum Draht
- Rotation der Gelenke
- Horizontale Position

Belohnungen und Bestrafungen:

- Erreichen des Endes +1
- Berühren des Drahtes -1
- Jeder benötigte Schritt -0,0001
- Höhe und Richtung der Geschwindigkeit $\text{velocity} \times 0,01$

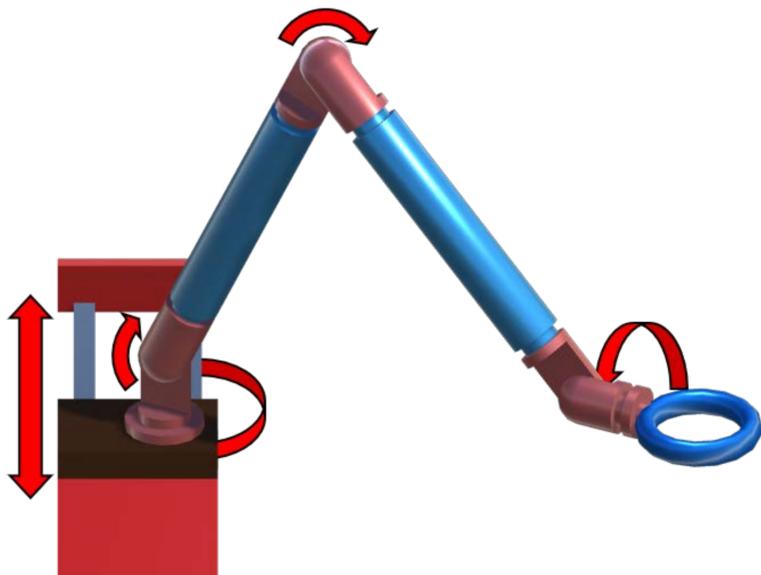


Abbildung 2: Illustration der Aktionen und der Observationen des Agenten (Roboterarm)

Use-Case 2: Sortiermaschine auf dem Fließband

Ein Agent (graue Box) sortiert Pakete auf einem Fließband.

Aktionen des Agenten (Abb. 3):

- Horizontale Bewegung

Observationen des Agenten (Abb. 3):

- Abtasten der Distanzen zu Objekten auf dem Fließband

Belohnungen und Bestrafungen:

- Richtig sortierter Gegenstand +1
- Falsch sortierter Gegenstand -1

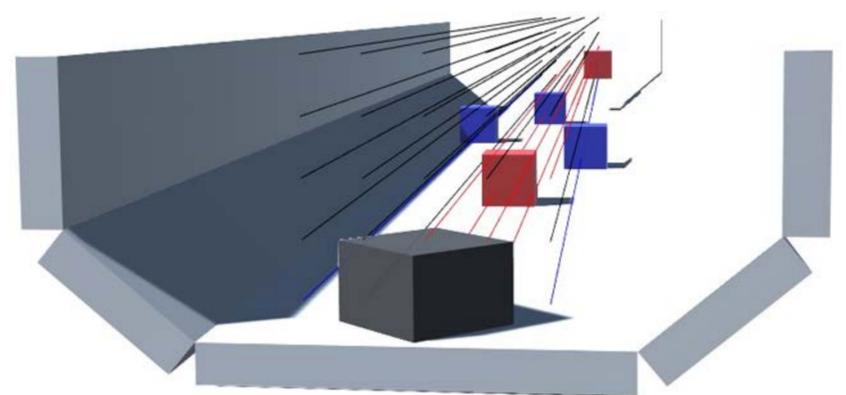
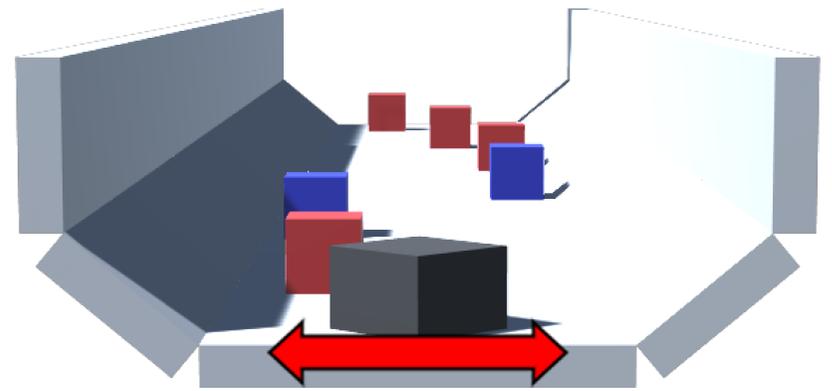


Abbildung 3: Illustration der Aktionen und Observationen des Agenten (graue Box)