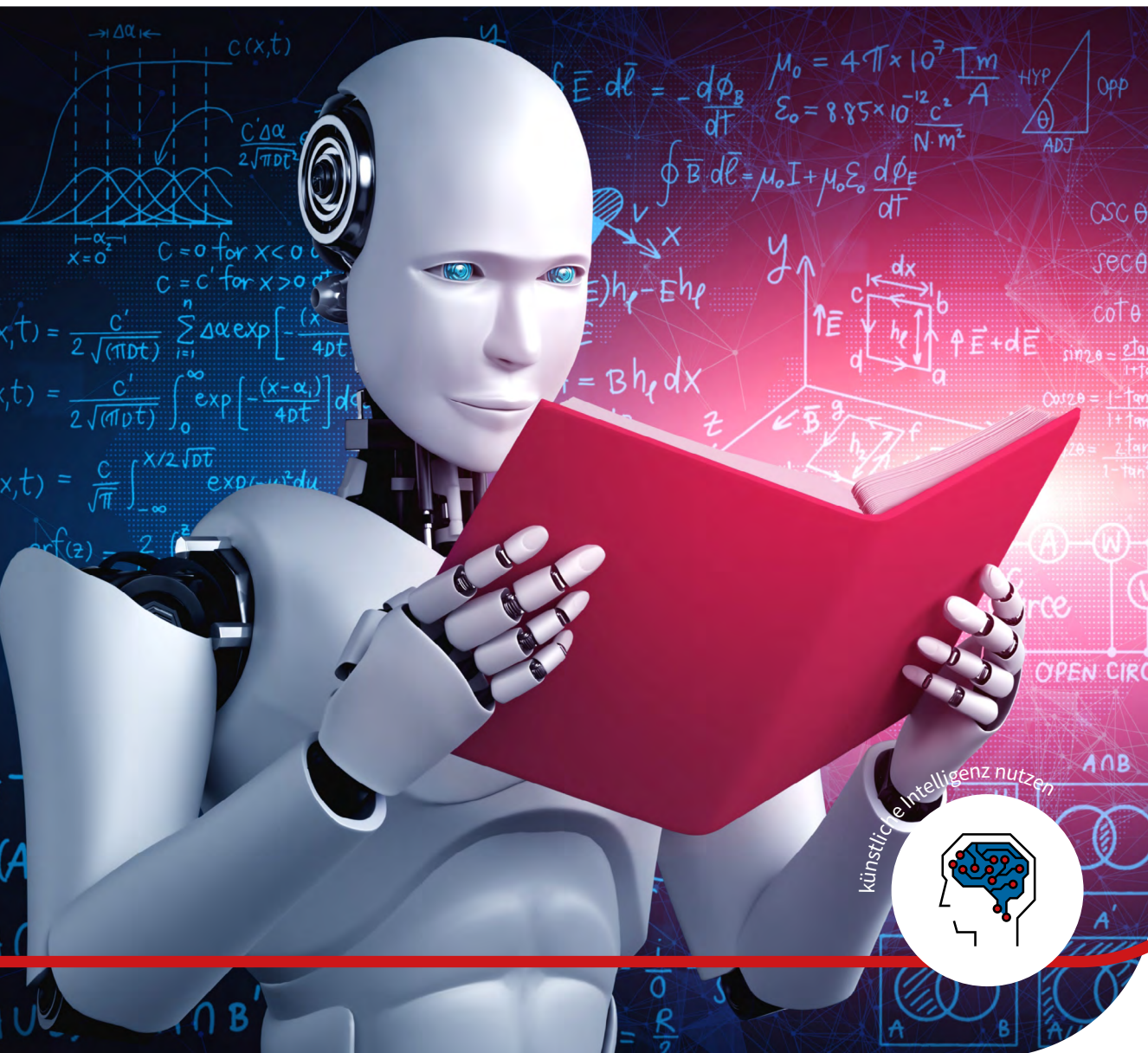




# Reinforcement Learning in Produktionssystemen

SINA NAHVI



## Impressum

### **HERAUSGEBER**

Mittelstand-Digital Zentrum Chemnitz

Tel: 0371 531 19935

Fax: 0371 531 819935

[info@digitalzentrum-chemnitz.de](mailto:info@digitalzentrum-chemnitz.de)

[www.digitalzentrum-chemnitz.de](http://www.digitalzentrum-chemnitz.de)

**REDAKTION** Anikó Lessi

### **GESTALTUNG**

PUNKT191 – Marketing und Design

[www.punkt191.de](http://www.punkt191.de)

### **BILDNACHWEIS TITEL**

Biancoblue - Freepik.com

**VERÖFFENTLICHUNG** Juli 2022





# Reinforcement Learning in Produktionssystemen

Immer kürzere Produktentwicklungszyklen und vollständig anpassbare Produkte stellen große Herausforderungen für Produktionssysteme dar. Diese müssen nicht nur eine größere Produktvielfalt bewältigen, sondern auch eine hohe Anpassungsfähigkeit und Robustheit gegenüber Prozessschwankungen und unvorhergesehenen Ereignissen aufweisen<sup>1</sup>. Darüber hinaus prognostizieren Experten in einigen Jahren einen Fachkräftemangel, der es Unternehmen erschweren wird, qualifizierte Arbeitskräfte zu einem budgetgerechten Preis zu finden<sup>2</sup>. Daraus ergibt sich ein Mangel an Experten und erfahrem Personal für die Überwachung, Kontrolle und Anpassung der Prozesse. Vor diesem Hintergrund nimmt die Automatisierung und Optimierung industrieller Aufgaben drastisch zu und Unternehmen stellen die derzeitigen Praktiken in Frage und ziehen Alternativen in Betracht, um die Produktions-, Betriebs- und Arbeitskosten zu senken.

Der weltweite Markt für Fabrikautomation und industrielle Steuerungssysteme wurde im Jahr 2020 auf 194,67 Mrd. USD geschätzt und wird bis 2026 voraussichtlich 339,56 Mrd. USD erreichen, wobei im Prognosezeitraum (2021-2026) eine jährliche Wachstumsrate von ca. 9 Prozent zu verzeichnen ist<sup>3</sup>. Dies zeigt, wie wichtig es für Unternehmen ist, die zuvor genannten Herausforderungen zu bewältigen.

Aufgrund der fortschreitenden Digitalisierung der Fertigungsindustrie und der daraus resultierenden verfügbaren Daten haben Methoden des Maschinellen Lernens ihren Weg in verschiedene industrielle Bereiche gefunden<sup>4</sup>. Dabei wächst in

letzter Zeit auch das Interesse, Reinforcement Learning (RL)-Algorithmen in den Kontext von Produktionssystemen zu übertragen. RL ist eine Disziplin des Maschinellen Lernens, die einen anderen Ansatz als herkömmliche Methoden bietet. Anstelle einer genauen Modellierung zur Ausführung von Aufgaben wendet RL ein menschenähnliches Lernverhalten an, das auf Trial-and-Error-Interaktionen basiert. Reinforcement Learning ermöglicht datengetriebene Online-Entscheidungen in Echtzeit und unterstützt ein reaktionsgesteuertes und adaptives Systemdesign<sup>5</sup>. Daher kann RL als datengetriebener Online-Optimierungs- und Steuerungsansatz betrachtet werden.

In diesem Nachgelesen erfahren Sie:

- Was Reinforcement Learning ist,
- wie sich Reinforcement Learning von traditionellen Methoden des Maschinellen Lernens unterscheidet und
- welche nützlichen Eigenschaften Reinforcement Learning für Produktionssysteme bietet.

# Grundlagen des Reinforcement Learning

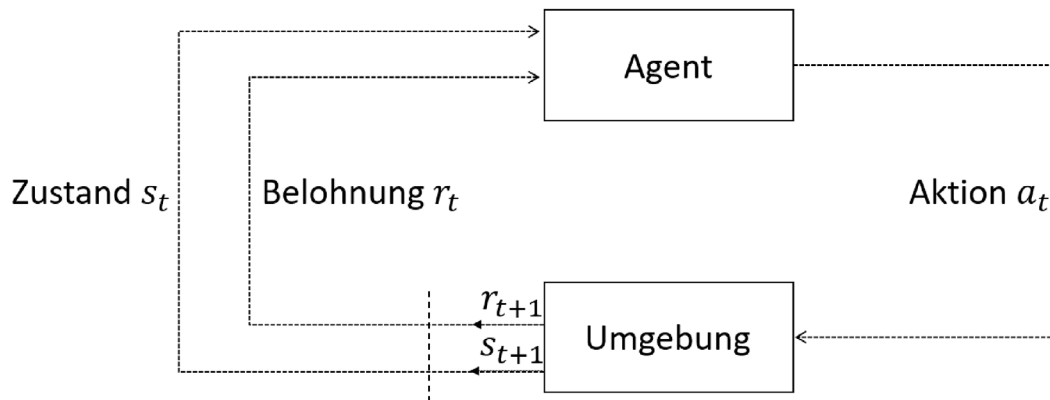
Das Schlüsselkonzept von RL beruht auf Trial-and-Error-Lernen, ohne dass zuvor gesammelte Daten oder (menschliches) Vorwissen erforderlich sind. Dabei hat es die Fähigkeit, sich flexibel an unsichere Bedingungen anzupassen<sup>6</sup>. RL besteht aus fünf wesentlichen Komponenten (Abbildung 1):

- dem Agenten
- der Umgebung
- dem Zustand
- der Aktion und
- der Belohnung.

Der Agent ist der Entscheidungsträger und interagiert mit der Umgebung. Die Umgebung beobachtet ihre aktuelle Situation in einem Zustand. Basierend auf dem beobachteten Zustand entscheidet der Agent über Aktionen, die zu einem neuen Zustand führen. Des Weiteren erhält der Agent eine Rückmeldung in Form einer Belohnung. Dies wird als (rein-

che Maßnahmen es ergreifen soll. Stattdessen muss es verschiedene Aktionen ausprobieren und auf der Grundlage der erhaltenen Rückmeldungen entscheiden, welche die beste ist. Das Ziel des Agenten ist es, durch die Auswahl seiner Aktionen die Belohnungen über die Zeit zu maximieren. Die Belohnungsfunktion muss in Bezug auf das Ziel des Agenten und die Systemdynamik definiert werden und muss sowohl kurz- als auch langfristige Ergebnisse berücksichtigen können<sup>1</sup>.

Betrachten wir als Beispiel einen Agenten, der durch Betätigung eines Roboterarms mittels Versuch und Irrtum lernt, direkt beliebige Zielpositionen zu erreichen. Die Belohnungsfunktion soll so konstruiert sein, dass sie die euklidische Distanz zur Position berücksichtigt. Bewegt sich der Arm direkt auf die Position zu, wird der Agent mit einem positiven Wert belohnt. Bei einer Abweichung von der direkten Linie wird er mit einem negativen Wert bestraft. Andere Parameter der Gelenke, wie Position und Geschwindigkeit, können ebenfalls in die Belohnungsfunktion einbezogen werden, um optimale Bewegungen zu finden. Eine Belohnung führt dazu, dass diese Verhaltensweisen zukünftig häufiger durch den Agenten ausprobiert werden. Im Falle einer Bestrafung wird die verwendete Strategie in Zukunft seltener erprobt.



↑ Abbildung 1: Interaktion zwischen Agent und Umgebung

forcement) Verstärkungssignal bezeichnet. Der Agent kann viele verschiedene Zustände erreichen, je nachdem, welche Aktionen er durchführt. Wenn der Agent positive Fortschritte gemacht hat, erhält er eine positive Verstärkung (Belohnung). Wenn er eine schlechte Entscheidung getroffen hat, erhält er eine negative Verstärkung (Strafe). Die Belohnung erfolgt nicht immer sofort. Sie kann auch erst nach einigen Entscheidungen erfolgen. Das System wird auf der Grundlage seiner Gesamtleistung und nicht für jeden einzelnen Schritt belohnt. Es bedarf erheblicher Trial-and-Error-Interaktionen, bevor das System die besten Entscheidungen treffen kann. Zu keinem Zeitpunkt wird dem System gesagt, wel-

Durch Reinforcement Learning in Kombination mit Funktions-Approximatoren wie künstlichen neuronalen Netzen, dem sogenannten Deep Reinforcement Learning (DRL), lassen sich bereits eine Vielzahl komplexer Steuerungsaufgaben lösen. Die Motivation für die Anwendung von DRL ist häufig eine ungenaue Abbildung konventioneller Methoden, die mit Nichtlinearitäten nicht angemessen umgehen kann oder zu sehr auf fehleranfälliges Expertenwissen angewiesen ist<sup>1</sup>. Im Gegensatz zu statischen und menschenbasierten Prozessmodellen kann DRL aufgrund seines adaptiven und kontinuierlichen Aktionsraums selbst sensible Prozesse stabil halten.

Das Anwendungsspektrum von Deep Reinforcement Learning beinhaltet alle wichtigsten Gebiete in einem Produktionssystem und zeigt die Fähigkeit von DRL, eine Vielzahl von Aufgaben zu erledigen, wie z. B. in der Prozesssteuerung, Produktionsdisposition, Logistik, Montage, Robotik, Wartung, Energiemanagement, Prozessgestaltung und Qualitätskontrolle.

## Herausforderungen des Reinforcement Learning

Einer der wichtigsten Aspekte bei der Implementierung von Reinforcement Learning ist die Frage, wie der Agent dazu gebracht werden kann, die gewünschte Fähigkeit zu üben. RL-Algorithmen benötigen in der Regel viele Daten, um komplexe Verhaltensweisen zu erlernen. Die Daten müssen durch Interaktionen zwischen dem Agenten und seiner Umgebung erzeugt werden. Es gibt zwei Möglichkeiten zur Datengewinnung:

- 1 durch Interaktion mit der realen Umgebung oder
- 2 durch Simulation.

In vielen Situationen kann die Interaktion mit einer physischen Umgebung, in der ein RL-Algorithmus lernen kann, sehr zeitaufwendig, kostspielig und sogar gefährlich sein. Für viele große Systeme ist daher die Simulation die einzige Möglichkeit, die beste Abfolge der Aktionen zu finden. In diesen Fällen muss ein digitales Modell des physischen Systems erstellt werden, um die Daten zu generieren, die das RL benötigt. Simulationen ermöglichen es, die Abfolge von Aktionen in einer digitalen Welt durchzuführen und ein KI-System so zu trainieren, dass es sein volles Potential ausschöpft und gleichzeitig Zeit und Geld spart. Die Lücke zwischen Simulation und Realität sollte so gering sein, dass die in der Simulation gelernten Strategien auf die reale Welt übertragbar sind.

## Fazit

Im Gegensatz zu konventionellen Methoden, die nur bedingt auf sich ändernde Bedingungen reagieren können, sind Reinforcement Learning-Algorithmen in der Lage, die Robustheit und Anpassungsfähigkeit von Produktionen zu erhöhen. Durch RL können Unternehmen die Abhängigkeit von immer knapper werdendem Humankapital begrenzen und datengesteuerte Abläufe proaktiv nutzen, um kostenintensive manuelle und expertenbasierte Prozesse zu reduzieren<sup>1</sup>. Die Folge dieser Gewinne werden Effizienzsteigerungen und Kosteneinsparungen in den Produktionssystemen sein, was u. a. zu einer Verringerung der CO<sub>2</sub>-Emissionen führt. Zusammengefasst ist RL ein vielversprechendes Werkzeug, um die Leistung moderner Produktionssysteme zu verbessern und den Übergang zu Industrie 4.0 zu ermöglichen<sup>1</sup>.



## Anmerkungen/Quellen

- 1** Panzer, Marcel & Bender, Benedict. (2021). Deep reinforcement learning in production systems: a systematic literature review. International Journal of Production Research. 1-26. 10.1080/00207543.2021.1973138.
- 2** Employee Turnover Statistics: 2022 Causes, Cost & Prevention Data. Online unter: <https://financesonline.com/employee-turnover-statistics> (abgerufen am 10.06.2022).
- 3** Global Factory Automation and Industrial Controls Market – Growth, Trends, COVID-19 Impacts and Forecasts (2022-2027). Online unter: [https://www.mordorintelligence.com/industry-reports/global-factory-automation-and-industrial-controls-market-industry#:~:text=Market%20Overview,period%20\(2021%2D2026\)](https://www.mordorintelligence.com/industry-reports/global-factory-automation-and-industrial-controls-market-industry#:~:text=Market%20Overview,period%20(2021%2D2026)) (abgerufen am 10.06.2022).
- 4** D. Weichert, P. Link, A. Stoll, S. Rüping, S. Ihlenfeldt, and S. Wrobel, “A review of machine learning for the optimization of production processes,” Int J Adv Manuf Technol, vol. 104, 5-8, pp. 1889–1902, 2019, DOI: 10.1007/s00170-019-03988-5
- 5** Han, Bao-An, and Jian-Jun Yang. 2020. “Research on Adaptive Job Shop Scheduling Problems Based on Dueling Double DQN.” IEEEAccess 8: 186474–186495. doi:10.1109/ACCESS.2020.3029868.
- 6** Sutton, Richard S., and Andrew G. Barto. 2017. Reinforcement Learning: An Introduction. 2nd ed. Adaptive computation and machine learning series. Cambridge: The MIT Press.

## Autor

**SINA NAHVI** ist wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik (IWU). Im Mittelstand-Digital Zentrum Chemnitz unterstützt er als KI-Trainer kleine und mittlere Unternehmen bei der Anwendung von künstlicher Intelligenz in der Produktion.

[sina.nahvi@digitalzentrum-chemnitz.de](mailto:sina.nahvi@digitalzentrum-chemnitz.de)

## Weitere Informationen

Das Mittelstand-Digital Zentrum Chemnitz gehört zu Mittelstand-Digital. Mit dem Mittelstand-Digital Netzwerk unterstützt das Bundesministerium für Wirtschaft und Klimaschutz die Digitalisierung in kleinen und mittleren Unternehmen und dem Handwerk.

### **WAS IST MITTELSTAND-DIGITAL?**

Das Mittelstand-Digital Netzwerk bietet mit den Mittelstand-Digital Zentren, der Initiative IT-Sicherheit in der Wirtschaft und Digital Jetzt umfassende Unterstützung bei der Digitalisierung. Kleine und mittlere Unternehmen profitieren von konkreten Praxisbeispielen und passgenauen, anbieterneutralen Angeboten zur Qualifikation und IT-Sicherheit. Das Bundesministerium für Wirtschaft und Klimaschutz ermöglicht die kostenfreie Nutzung und stellt finanzielle Zuschüsse bereit. Weitere Informationen finden Sie unter [www.mittelstand-digital.de](http://www.mittelstand-digital.de).





Mittelstand-Digital  
Zentrum  
Chemnitz

