Deutsche Akademie
für Management

Grundkonzepte der KI

Christian Schubert / Prof. Dr. Markus Dormann

www.akademie-management.de

Modul 15010

Impressum

Trägerin: DAM Professional School SE
Margaretenstraße 38, 12203 Berlin

mail@akademie.biz
www.akademie-management.de

Tel. 030/40508883-0
Fax. 030/40508883-9

1. Version September 2025

© 2025 Deutsche Akademie für Management, Berlin.
Alle Rechte vorbehalten.

Umweltfreundlich gedruckt auf 100% nachhaltig hergestelltem FSC®-zertifiziertem Papier

Verfasser

Christian Schubert ist Berater und Experte für KI-Analysen und Anwendungen für kleine und mittelständische Unternehmen. Er ist Pionier für die Modellierung und Spezifizierung von Large Language Modellen.

Prof. Dr. Markus Dormann ist Vizepräsident für Lehre und Professor für Gesundheitspädagogik, HealthCare Education und E-Learning an der IB Hochschule. Er studierte Soziologie und Wirtschaftspädagogik, absolvierte ein Referendariat für das Lehramt an beruflichen Schulen. Der Autor ist seit vielen Jahren als Berater und Coach und Autor in den Bereichen KI-Einsatz im Bildungs-und Gesundheitswesen, E-Learning, Digitalisierung, Kommunikation und Konfliktmanagement international tätig.

Der gesamte Inhalt des vorliegenden Moduls (Texte, Bilder, Grafiken, Design usw.) und jede Auswahl davon unterliegt dem Urheberrecht und anderen Gesetzen zum Schutze geistigen Eigentums der DAM Professional School SE oder anderer Eigentümer. Jede Verwertung außerhalb der engen Grenzen des Urheberrechts ist ohne Zustimmung des Eigentümers unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Zuwiderhandlungen werden zivil- und strafrechtlich verfolgt.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Text berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften. Sämtliche verwendete Handelsmarken oder Markenzeichen sind Eigentum der jeweiligen Rechteinhaber.

Die Deutsche Akademie für Management und ihre Autoren haben höchste Sorgfalt bei der Erstellung des vorliegenden Moduls angewandt. Dennoch übernehmen sie keinerlei Verantwortung oder Haftung für Richtigkeit oder Vollständigkeit, eventuelle Fehler oder Versäumnisse innerhalb des Moduls. Die Inhalte und Materialien werden unter Ausschluss jeglicher Gewährleistung zur Verfügung gestellt. Insbesondere erfolgt die Anwendung von im Modul dargestellten Erkenntnissen auf Gefahr der Teilnehmenden.

Gendergerechte Sprache

Gendergerechte Sprache ist ein weites Feld. Die einen sehen sie als zwingend an, die anderen lehnen sie kategorisch ab – und zwar geschlechterübergreifend. Dabei stellt sich nicht nur die Frage, ob gegendert wird oder nicht, sondern auch in welchem Ausmaß und nach welchem System.

So gibt es beispielsweise das Del-On-Se-System (Bsp.: „Del Lehreron gibt derl Schüleron das Buch."), den g-Genus (Bsp.: „Hen ist meineo beste Freundis."), das NoNa-System (Bsp.: „Hen ist meint beste Freund*in."), die SYLVAIN-Konventionen (Bsp.: „Eyn bekanntey Autor_in erzählt.", das vii-System (Bsp.: „Aufgabe einvis Lehrvis ist es, viis Schülviis das Lernen angenehm zu gestalten.") und zahlreiche andere Möglichkeiten.*

Einen Königsweg scheint es noch nicht zu geben.

Auch wir an der Deutschen Akademie für Management müssen zum Gendern Stellung beziehen. Dabei haben wir neben der Genderfairness auch pädagogische und didaktische Anforderungen zu berücksichtigen. Die in der Pädagogik aktuell stark diskutierte Sprachsensibilität von Unterrichtsmaterialien beispielsweise fordert, die problemlose kommunikative Teilhabe von Lernenden mit unterschiedlichen individuellen Kommunikationskompetenzen zu ermöglichen und zu fördern.

Die Deutsche Akademie für Management betrachtet die Gleichbehandlung aller Geschlechter als essenziell. Unsere Lernmaterialien und unser Unterricht sollen für alle Menschen ein möglichst einheitliches und einfaches Verständnis ermöglichen. Um dies zu erreichen, verwenden wir das aktuellen Kommunikationsgewohnheiten entsprechende grammatikalische Maskulinum. Dieses sprachliche Geschlecht soll dabei in keinem Zusammenhang mit dem sozialen oder biologischen Geschlecht stehen. Insofern sprechen wir mit unseren Lernmaterialien und in unserem Unterricht alle Geschlechter gleichberechtigt und gleichermaßen wertschätzend an.

* Beispiele entnommen aus dem Nichtbinär-Wiki (https://nibi.space/geschlechtsneutrale_sprache).

Inhaltsverzeichnis

Literaturempfehlungen **9**

Allgemeine Kompetenzziele **11**

1 Daten als Grundlage **12**

1.1 Der Ablauf des Trainings einer KI mit Daten 14

1.2 Zur zentralen Bedeutung von Daten 15

1.3 Datenquantität und -qualität 16

1.4 Datenvorverarbeitung 17

1.5 Feature-Extraktion: Konzentration auf das Wesentliche 20

1.6 Feature-Auswahl: Effizienz durch Fokussierung auf Relevanz 21

1.7 Datenaufbereitung und Umgang mit fehlenden Daten 23

2 Algorithmen und Modelltypen **27**

2.1 Modelle 28

2.2 Algorithmen 29

2.3 Lineare Modelle 31

2.4 Entscheidungsbäume 34

2.5 Clustering-Algorithmen 37

2.6 Ensemble-Methoden 39

3 Machine Learning **43**

3.1 Was ist Machine Learning 44

3.2 Überwachtes Lernen (Supervised Learning) 45

3.3 Unüberwachtes Lernen (Unsupervised Learning) 46

3.4 Bestärkendes Lernen (Reinforcement Learning) 47

3.5 Semi-überwachtes Lernen (Semi-Supervised Learning) 48

3.6 Selbstüberwachtes Lernen (Self-Supervised Learning) 50

4 Neuronale Netze 53

4.1 Was sind neuronale Netze? 54

4.2 Aufbau eines neuronalen Netzes 54

4.3 Stärken und Grenzen von Neuronalen Netzen 56

4.3.1 Convolutional Neural Networks (CNN) 58

4.3.2 Residual Networks (ResNet) 59

4.3.3 Recurrent Neural Networks (RNN) 59

4.3.4 Generative Adversarial Networks (GAN) 60

4.3.5 Large Language Models (LLM) 60

5 Modellvalidierung 65

5.1 Mit Bias in Daten und Modellen umgehen 66

5.2 Evaluationsmetriken: Das Handwerkszeug der Validierung 67

5.3 Trainings- und Testmethoden: Vertrauen schaffen durch solide Verfahren 69

5.4 Einfache, intrinsisch interpretierbare Modelle 71

5.5 Organisationale Transparenz und Nachvollziehbarkeit 71
5.6 Nutzen von Pre-Trained Models und Transfer Learning 72
5.7 Transfer Learning 72
6 Hardwarevoraussetzungen und KI 74
6.1 GPUs, TPUs und VRAM 75
6.2 Cloud-Computing: Flexibilität vs. Kontrolle 76
6.3 Lokale KI: Chancen und Grenzen 76
6.4 Hybride KI-Architektur: Das Beste aus beiden Welten 77
Lösungshinweise zu den Kontrollaufgaben 80
Literaturverzeichnis 84
Stichwortverzeichnis 87

Literaturempfehlungen

[1] Russell, S. & Norvig, P., 2016. Künstliche Intelligenz: Ein moderner Ansatz. 3. Aufl. München: Pearson.

Das umfassende Grundlagenwerk deckt alle klassischen KI Themen ab: Such und Optimierungsverfahren, logisches Schließen, Wissenrepräsentation, maschinelles Lernen und ethische Fragestellungen. Ideal, um einen breiten Überblick über Algorithmen und Modelltypen zu gewinnen.

[2] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer.

Ein Klassiker für statistische Lernverfahren: lineare Modelle, Entscheidungsbäume, Ensemble Methoden, Clustering, Modellvalidierung und Regularisierung. Dank zahlreicher Beispiele und mathematischer Tiefe perfekt, um die theoretischen Grundlagen von Machine Learning Algorithmen zu verstehen.

[3] Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
Fokussiert auf probabilistische Modelle und Algorithmen des überwachten und unüberwachten Lernens. Behandelt Datenvorverarbeitung, Feature Engineering, Validierungsmethoden sowie eine solide Einführung in Bayesianische Verfahren.

[4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Das Standardwerk zum Deep Learning – von der Architektur einfacher neuronaler Netze über CNNs, RNNs und GANs bis hin zu modernen Ansätzen wie Transformer Netzen. Enthält sowohl theoretische Erläuterungen als auch praktische Tipps zu Training und Regularisierung.

[5] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

Die maßgebliche Einführung in bestärkendes Lernen. Erklärt Agent Umwelt Interaktionen, Policy und Value Basierte Verfahren, Monte Carlo Methoden und Temporal Difference Learning. Unerlässlich, um den dritten großen Lernparadigmen Block (Reinforcement Learning) abzudecken.

[6] Hennessy, J. L., & Patterson, D. A. (2017). Computer architecture: A quantitative approach (6. Aufl.). Morgan Kaufmann

Der internationale Standard¬text zur Rechnerarchitektur erklärt CPUs, GPUs und beschleunigende Spezial¬prozessoren (TPUs, FPGAs) in technischer Tiefe. Enthält Vergleiche von Leistung, Energieeffizienz und Speicherhierarchien und zeigt, wie moderne Hardware-Designs KI Workloads beschleunigen. Empfehlenswert, um die Hardware-Seite Ihres Kapitels 2.6 fundiert abzurunden.

allgemeine kompetenzziele

Allgemeine Kompetenzziele:

Nach Bearbeitung des Studienbriefs können die Lernenden

- den kompletten Ablauf von der Datenerhebung über Vorverarbeitung, Modelltraining bis zur Validierung und Inferenz nachvollziehen und dessen Bedeutung für reale Anwendungen reflektieren.
- beurteilen wie Datenqualität, -quantität und -struktur die Leistung von KI-Systemen beeinflussen, und geeignete Methoden zur Bereinigung, Feature-Engineering und zum Umgang mit fehlenden Werten anwenden.
- zentrale KI Modelltypen (lineare Modelle, Entscheidungsbäume, Clustering Algorithmen, Ensemble Methoden, neuronale Netze) sowie die fünf Lernparadigmen (supervised, unsupervised, reinforcement, semi supervised, self supervised) fachgerecht erklären und anhand von Praxisbeispielen unterscheiden.
- den Aufbau verschiedener neuronaler Netze (z. B. CNN, RNN, ResNet, GAN, LLM) und die wesentlichen Trainingsschritte (Vorwärtsdurchlauf, Rückpropagation, Anpassung der Verbindungsgewichte) in einfachen Worten erklären und typische Anwendungsgebiete für jede Architektur nennen.
- die Lernenden gängige Validierungsverfahren (Train/Test Split, K Fold-Cross Validation) und Performance Metriken (Precision, Recall, F1 Score, AUC) anwenden, Bias in Daten und Modellen erkennen und mit Erklärmethoden (Feature Importance, LIME, SHAP) die Nachvollziehbarkeit von Modellergebnissen sichern.
- die wesentlichen Hardware Komponenten (GPUs, TPUs, VRAM) sowie die Vor und Nachteile von Cloud vs. Edge KI Architekturen erläutern und eine hybride Strategie zur Optimierung von Performance, Kosten und Datenschutz skizzieren.

1 Daten als Grundlage

kompetenzziele

Kompetenzziele:

Nachdem Sie dieses Kapitel studiert haben, sollen Sie in der Lage sein:

- die Bedeutung von Daten für Künstliche Intelligenz (KI) zu verstehen.
- die wesentlichen Schritte der Datenaufbereitung und -verarbeitung zu identifizieren und zu verstehen.
- Konzepte von Datenqualität und -quantität anhand spezifischer Kriterien zu bewerten.
- Kriterien und Methoden für die Auswahl und Extraktion von Features zu beschreiben.
- Strategien für den Umgang mit fehlenden Datenpunkten zu benennen und zu verstehen.

Einleitung

Stellen Sie sich vor, Sie möchten ein langfristig erfolgreiches Restaurant führen. Damit Ihre Gäste immer wiederkommen und sich wohlfühlen, brauchen Sie eine klare Entscheidungsgrundlage. Dazu sammeln Sie Feedback von Ihren Gästen, analysieren Verkaufszahlen und beobachten, welche Gerichte besonders beliebt sind und welche weniger häufig bestellt werden. Diese Daten helfen Ihnen nicht nur, die Menügestaltung laufend zu verbessern, sondern auch, passgenauere Werbeaktionen zu planen und gezielter auf die Wünsche Ihrer Gäste einzugehen. Je mehr relevante und präzise Informationen Sie zur Verfügung haben, desto fundierter und nachhaltiger können Sie handeln. Genauso wie in diesem Beispiel sind Daten das Fundament erfolgreicher Entscheidungsfindungen in der Künstlichen Intelligenz (KI).

Daten sind der Rohstoff der modernen KI-Technologien – in Qualität und Quantität entscheidend. Ohne Daten ist die Weiterentwicklung und optimale Nutzung von KI nicht denkbar. Anwendungen wie Sprachassistenten, Empfehlungssysteme oder autonomes Fahren wären ohne geeignete und qualitativ hochwertige Daten