OLIVER MICHAELIS

2. AUFLAGE

EINFÜHRUNG IN DIE KI

Grundlagen, Anwendungen, Recht und Verantwortung

Mit zahlreichen Beispielen, Abbildungen und Praxisfragen

WISSEN.
PRAXIS.
ZUKUNFT.

m!

Oliver Michaelis

Einführung in die KI

2. vollständig überarbeitete Auflage

2026

m!

Vorwort



Künstliche Intelligenz begegnet uns oft als fertiges Ergebnis: als flüssig formulierter Text, als erkanntes Bild oder als Vorschlag für den nächsten Arbeitsschritt. Eine überzeugende Ausgabe verrät jedoch noch nicht, wie sie entstanden ist, worauf sie beruht und wo ihre Grenzen liegen. Dieses Buch lädt dazu ein, hinter die Oberfläche zu schauen und KI nicht nur zu bedienen, sondern ihre Leistungen begründet zu beurteilen.

Dieses Buch „Die Einführung in die KI“ richtet sich an Studenten, Dozenten und Berufstätige, die technische, wirtschaftliche und rechtliche Zusammenhänge verstehen möchten, ohne ein Informatikstudium voraussetzen zu müssen. Begriffe werden schrittweise erklärt, Rechenwege an überschaubaren Beispielen entwickelt und Anwendungen in ihren tatsächlichen Arbeitsablauf eingeordnet. Im Mittelpunkt steht stets die Frage, welche Aufgabe gelöst werden soll und was ein brauchbares Ergebnis ausmacht.

Für die zweite Auflage habe ich den Stoff neu geordnet und die Verbindung zwischen Grundlagen und Anwendung vertieft. Mehr Raum erhalten das Lernen aus Daten, generative Systeme, die Prüfung von Ergebnissen sowie die organisatorischen und rechtlichen Bedingungen ihres Einsatzes. Fachbeiträge ergänzen die technischen Originalarbeiten. Bei rechtlichen Fragen werden Normtext, erläuternde Erwägungsgründe und fachliche Einordnung unterschieden; Inkrafttreten, Anwendungsbeginn und Übergangsregelungen bleiben getrennt.
Die Übersichten am Kapitelanfang zeigen den Standort im Buch und die Beziehungen zwischen den folgenden Themen. Nummerierte Fälle machen abstrakte Zusammenhänge an konkreten Situationen nachvollziehbar. Tabellen erleichtern den Vergleich; die Aufgaben am Kapitelende fordern zur eigenen Bearbeitung auf. Die ausführlichen Lösungen erläutern nicht nur das Ergebnis, sondern auch den Weg dorthin und die Bedeutung noch fehlender Informationen. Die für das Buch entwickelten Fälle und Lehrzahlen sind didaktische Beispiele, keine Berichte über tatsächlich untersuchte Betriebe.

Sie können das Buch fortlaufend lesen oder einzelne Kapitel nach Ihrem Interesse auswählen.
Querverweise, ein Glossar und ein Stichwortverzeichnis unterstützen zusätzlich die Orientierung.

Ich wünsche Ihnen viel Spaß beim Lesen und Erlernen.

Düsseldorf, den 25. September 2026
Oliver Michaelis

Inhaltsverzeichnis

Vorwort III

So arbeiten Sie mit diesem Buch IV

1. Was KI ist und wo ihre Grenzen liegen 1

1.1. Von einer großen Bezeichnung zu einer konkreten Aufgabe 2

1.1.1. Was mit künstlicher Intelligenz gemeint ist 2

1.1.2. Warum Definitionen verschiedene Zwecke haben 3

Fall 1: Drei Wünsche an einen Handwerksbetrieb 3

1.2. Algorithmus, Modell und System auseinanderhalten 5

1.2.1. Der Algorithmus als festgelegtes Vorgehen 5

1.2.2. Das Modell als vereinfachte Abbildung 6

1.2.3. Das System als organisierte Gesamtheit 6

Fall 2: Warum ein richtiges Modell eine falsche Rechnung nicht verhindert 7

1.3. Regeln anwenden oder Zusammenhänge lernen 8

1.3.1. Symbolische Verfahren 8

1.3.2. Maschinelles Lernen 9

1.4. KI, maschinelles Lernen und Deep Learning einordnen 10

1.4.1. Verwandte Begriffe mit unterschiedlichen Reichweiten 10

1.4.2. Generativ, multimodal und agentisch 10

Fall 3: Eine Kantine plant den nächsten Mittag 11

1.5. Was während des Lernens geschieht und was bei der Nutzung 13

1.5.1. Training und Inferenz 14

1.5.2. Verschiedene Formen von Erfahrung 14

Fall 4: Ein Fahrradladen lernt nicht bei jedem Gespräch neu 15

1.6. Aus Daten wird nicht automatisch Wissen 16

1.6.1. Daten entstehen durch Auswahl 16

1.6.2. Zusammenhang, Erklärung und Ursache 17

Fall 5: Ein voller Parkplatz erklärt noch keinen Umsatz 17

1.7. Von passenden Beispielen zu neuen Fällen 18

1.7.1. Generalisierung und Überanpassung 18

1.7.2. Warum ein guter Test unbekannte Fälle benötigt 19

Fall 6: Eine Qualitätskamera verwechselt das Teil mit seiner Umgebung 20

Fall 3a: Eine Messewoche verändert die Bedeutung der Daten 21

1.8. Was Leistungsangaben aussagen 22

1.8.1. Eine Prozentzahl braucht einen Nenner 22

1.8.2. Sicherheitssprache und kalibrierte Schätzungen 23

Fall 7: Eine Wartungswarnung braucht mehr als eine hohe Trefferquote....24
Wenn keine sinnvolle Aussage möglich ist....24
1.9. Menschliche Intelligenz als Vergleich und als Missverständnis....25
1.9.1. Leistung und inneres Erleben unterscheiden....25
1.9.2. Verstehen als unterschiedlich anspruchsvolle Aussage....25
1.10. Enge Aufgaben, breite Fähigkeiten und allgemeine Intelligenz....26
1.10.1. Warum „schwache KI“ missverständlich sein kann....26
1.10.2. Breite Sprache verdeckt ein ungleichmäßiges Können....26
Fall 8: Ein Reiseassistent kann mehr dürfen, als er zuverlässig kann....27
1.11. Informationen abrufen, Antworten erzeugen und Quellen prüfen....28
1.11.1. Drei unterschiedliche Vorgänge....28
1.11.2. Eine Quelle muss die konkrete Aussage tragen....29
1.12. Verantwortung beginnt vor der Ausgabe....30
1.12.1. Zweck und Grenzen festlegen....30
1.12.2. Menschliche Kontrolle braucht Bedingungen....30
Fall 3b: Ein Fehler wird untersucht, ohne die Verantwortung zu verwischen....31
1.13. Eine Anwendung selbst beurteilen....32
1.13.1. Ein kurzer Steckbrief als Arbeitswerkzeug....32
1.13.2. Die einfachere Vergleichslösung ernst nehmen....33
Fall 9: Ein guter Assistent muss eine Ablehnung ermöglichen....33
1.14. Zusammenfassung und Kapitelcheckliste....34
1.15. Praxisaufgaben....35
Aufgabe 1.1 Ein Postfach, zwei Anliegen....35
Aufgabe 1.2 Die scheinbare Erinnerung des Verkaufsassistenten....35
Aufgabe 1.3 Eine Warnanlage mit beeindruckender Prozentzahl....36
Aufgabe 1.4 Ein Literaturhinweis mit tadellosem Aussehen....36
Aufgabe 1.5 Vier Bildergruppen statt einer Gesamtquote....37
Aufgabe 1.6 Der schnellere Entwurf und die langsamere Erledigung....37

2. Wie sich KI entwickelt hat....39
2.1. Geschichte lesen, ohne eine Fortschrittslegende zu schreiben....40
2.1.1. Ein Ereignis hat mehrere Bedeutungen....40
2.1.2. Drei Fragen an jeden Meilenstein....40
2.2. Vor der Bezeichnung KI: Rechnen, Logik und künstliche Neuronen....41
2.2.1. Warum Rechnen eine allgemeine Bedeutung erhielt....41
2.2.2. McCulloch und Pitts: ein früher formaler Zugang....41
Fall 10: Aus einer Lagerhalle wird eine berechenbare Aufgabe....42
2.3. Turing: Eine unklare Frage wird zu einer untersuchbaren Aufgabe....43
2.3.1. Das Imitationsspiel und sein Zweck....43

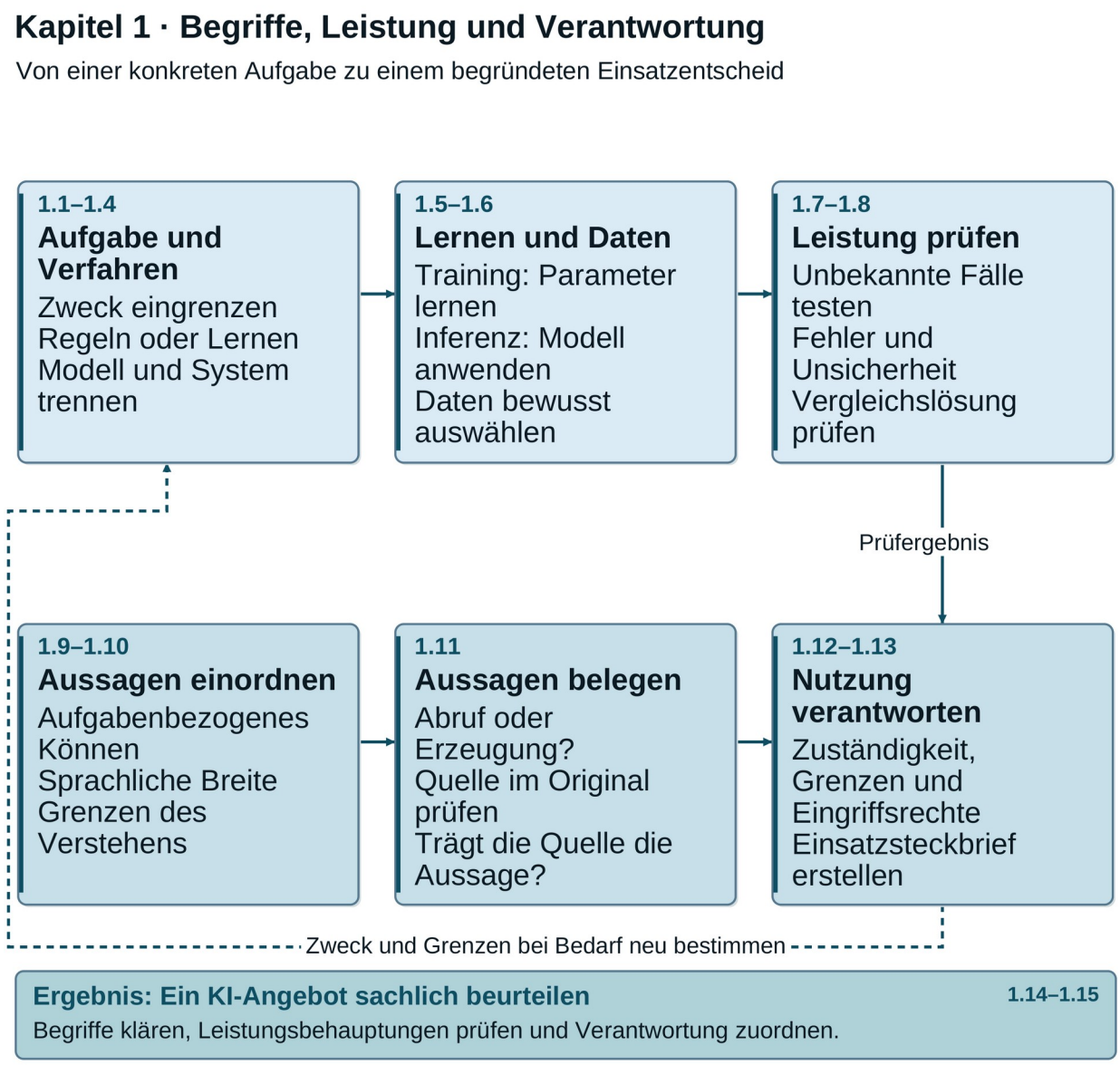
1. Was KI ist und wo ihre Grenzen liegen

Abb. 1: Kapitel 1 im Lernweg



Quelle: Michaelis, eigene Darstellung.

Abb. 2: Orientierung in Kapitel 1: Von der Aufgabe zur verantworteten Nutzung



Quelle: Michaelis, eigene Darstellung der Kapitelstruktur.

Im folgenden Fall erhält ein Student von einem Sprachmodell eine gut formulierte Erklärung mit drei Literaturhinweisen. Bei der Katalogsuche stellt sich heraus, dass einer der Aufsätze nicht existiert. Hat die Maschine die Frage verstanden, recherchiert oder gelogen? Diese Fragen betreffen unterschiedliche Eigenschaften. Ihre Vermischung führt leicht dazu, die Leistung zu überschätzen oder vollständig zu verwerfen.
Dieses Kapitel vermittelt Begriffe, um Aufgaben, technische Lösungswege und die Aussagekraft von KI-Ergebnissen zu unterscheiden. Sie lernen, ein Modell im Gesamtsystem einzuordnen und Angebote kritisch zu beurteilen. Programmierkenntnisse oder höhere Mathematik sind dafür nicht erforderlich. Technische Einzelheiten vertiefen die Kapitel 3 bis 5; Anwendungen, Risiken und Recht folgen danach.
Eine fiktive Hochschulbibliothek begleitet das Kapitel. Sie möchte Anfragen schneller bearbeiten: Öffnungszeiten, Leihfristen und schwierige Literaturrecherchen verlangen unterschiedliche Unterstützung. Erst die Aufteilung zeigt, wo feste Regeln genügen, wo ein lernendes Verfahren helfen könnte und wo ein Mitarbeiter entscheiden muss.

1.1. Von einer großen Bezeichnung zu einer konkreten Aufgabe

Ein sinnvoller KI-Begriff muss zur jeweiligen Frage passen. Die juristische Überblicksliteratur beschreibt KI häufig über Aufgaben, die gewöhnlich menschliche Intelligenz verlangen. Das erleichtert den Einstieg, entscheidet aber noch nicht über den Anwendungsbereich eines Gesetzes. Für die KI-Verordnung kommt es auf die Merkmale des Art. 3 Nr. 1 an. Insbesondere ist die Fähigkeit zur Ableitung von Ausgaben zu untersuchen; weder eine werbliche Bezeichnung noch die bloße Komplexität der Software genügt. Technische Arbeitsbeschreibung und rechtliche Einordnung werden deshalb im Folgenden bewusst getrennt.[1]

1.1.1. Was mit künstlicher Intelligenz gemeint ist

Künstliche Intelligenz bezeichnet ein Forschungs- und Anwendungsgebiet für technische Leistungen wie Wahrnehmen, Lernen, Sprachverarbeitung, Schlussfolgern und Planen. Diese Arbeitsbeschreibung nennt Aufgaben, keine einheitliche Methode: Eine Routensuche kann vorgegebenen Regeln folgen, eine Bilderkennung auf einem aus Beispielen entwickelten Modell beruhen. Beide gehören zum Gegenstandsbereich der KI-Forschung. Davon zu unterscheiden ist der rechtliche Systembegriff.
Art. 3 Nr. 1 der Verordnung (EU) 2024/1689 definiert „KI-System“ als „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können“.[2]
Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (englisch Organisation for Economic Co-operation and Development, OECD) verwendet eine verwandte funktionale

[1] Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz; im Folgenden: KI-VO), ABl. L, 2024/1689, 12.7.2024. Zugrunde gelegt ist die Fassung vom 27. Juli 2026; Kuhlmann, NJW 2026, 2774, Rn. 1–7, zur allgemeinen und technischen Begriffsverwendung; zur Legaldefinition Art. 3 Nr. 1 KI-VO; Erwägungsgrund 12 Sätze 2–6 KI-VO, zur Ableitungsfähigkeit und zur Abgrenzung einfacher Regelverarbeitung.

[2] Art. 3 Nr. 1 KI-VO.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

Beschreibung. Ein menschenähnlicher Körper oder nachgewiesenes Bewusstsein ist auch dort kein Definitionsmerkmal.[3]
„Intelligenz“ ist kein Gütesiegel: Ein KI-System kann unzuverlässig arbeiten, eine gewöhnliche Datenbank eine Aufgabe hervorragend erfüllen. Auch Komplexität entscheidet nicht; eine komplizierte Abrechnung kann festen Regeln folgen, ein kleines statistisches Modell sinnvoll klassifizieren. Aus der Bezeichnung folgen weder Qualität noch rechtliche Zulässigkeit.
Bei der Bibliothek könnte eine einfache Regel lauten: Ist das gewünschte Rückgabedatum ein Schließtag, wird der nächste Öffnungstag angezeigt. Dafür ist kein Sprachmodell erforderlich. Soll eine frei formulierte Nachricht einer Sachgruppe zugeordnet werden, entstehen andere Schwierigkeiten: Studenten verwenden unterschiedliche Wörter, machen Tippfehler oder stellen mehrere Fragen gleichzeitig. Hier kann ein lernendes Verfahren nützlich sein. Es muss jedoch gesondert geprüft werden, ob es die relevanten Unterschiede tatsächlich erfasst.

1.1.2. Warum Definitionen verschiedene Zwecke haben

Die Legaldefinition bestimmt den Anwendungsbereich der KI-Verordnung, keine vollständige technische Taxonomie. Erwägungsgrund 12 Sätze 2–6 KI-VO grenzt einfache, ausschließlich von Menschen vorgegebene Regeln zum automatischen Ausführen von Operationen ab, erfasst aber auch logik- und wissensgestützte Ableitung. Daher sind weder alle historisch als KI untersuchten Verfahren automatisch KI-Systeme im Rechtssinn noch nur lernende Verfahren erfasst. Kapitel 9 erläutert die daran anknüpfenden Pflichten.[4]
Auch im Rechtssinn muss ein KI-System nicht ständig weiterlernen oder beliebige Aufgaben übernehmen können. Die Produktbezeichnung „Assistent“ sagt ebenso wenig über seine Befugnisse aus: Dahinter kann eine Suchfunktion, ein Sprachmodell mit Dokumentenzugriff oder eine Anwendung mit mehreren Werkzeugen stehen.
Fragen Sie daher nach Eingabe, Ausgabe, Verwendung und erkennbarem Fehler. Bei der Bibliothek: E-Mail, vorgeschlagene Sachgruppe, Weiterleitung durch einen Mitarbeiter und falsch gewählte Zuständigkeit. Diese Aufgabenbeschreibung ist bereits prüfbar, bevor das technische Verfahren feststeht.



Fall 1: Drei Wünsche an einen Handwerksbetrieb

Ein Heizungsbetrieb möchte Anfragen sortieren, den Ersatzteilbedarf schätzen und Kunden über Bearbeitungsstände informieren. Die drei Funktionen könnten dieselbe Benutzeroberfläche nutzen, benötigen aber unterschiedliche Daten und Prüfmaßstäbe. Der Fall zeigt, weshalb aus dem Wunsch nach weniger Büroarbeit zunächst getrennte, überprüfbare Aufträge entstehen müssen.
Bei der Postsortierung gehört ein tropfender Wasserhahn zum Kundendienst, eine Bewerbung zur Personalbearbeitung. Fragt ein Kunde jedoch nach einer Rechnung und meldet zugleich den erneuten Heizungsausfall, wäre eine alleinige Weiterleitung an die Buchhaltung unbrauchbar. Vor der Modellauswahl muss der Betrieb Mehrfachzuordnungen und die Rangfolge dringlicher Anliegen festlegen.
Die Ersatzteilplanung schätzt zukünftigen Bedarf, etwa an Dichtungen. Frühere Bestellungen bilden ihn nur unvollständig ab: Nicht lieferbare Teile können zu ausgefallenen Aufträgen geführt haben, die in der Bestellhistorie fehlen. Eine darauf beschränkte Prognose könnte den Mangel fortschreiben. Der Betrieb muss deshalb

[3] OECD, Explanatory Memorandum on the Updated OECD Definition of an AI System, OECD Artificial Intelligence Papers Nr. 8, Paris 2024, S. 4 (Definition), S. 6 (Autonomie und Anpassungsfähigkeit) sowie S. 8–9 (Eingaben, Ableitung und Ausgaben).

[4] Art. 3 Nr. 1 KI-VO; Erwägungsgrund 12 Sätze 2–6 KI-VO, zur Abgrenzung einfacher Regelverarbeitung und zu maschinellem Lernen sowie logik- und wissensgestützten Konzepten. Erwägungsgründe erläutern die Auslegung; sie ersetzen nicht die Tatbestandsmerkmale des Artikels.

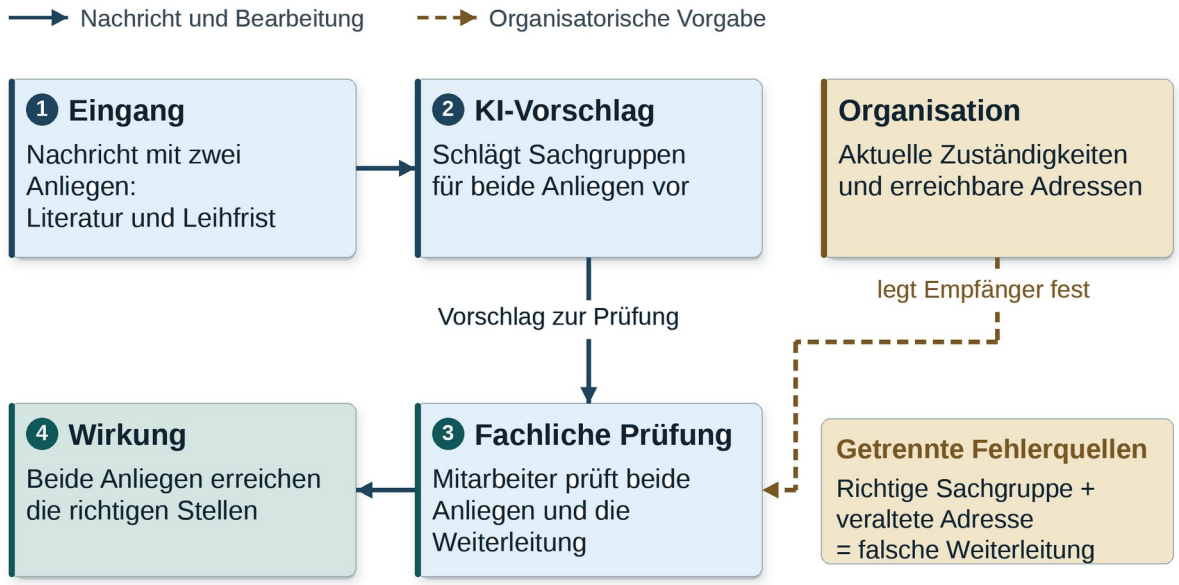
zwischen Bestellungen, Reparaturen, Kundenanfragen und tatsächlich benötigten Teilen als Zielgröße unterscheiden.
Eine Statusauskunft benötigt aktuellen, berechtigten Zugriff auf den richtigen Auftrag. Ein Sprachmodell kann einen Statuscode verständlich formulieren; den vereinbarten Termin muss es aus dem freigegebenen Auftragsbestand erhalten. Nennt der Kunde eine falsche Auftragsnummer, braucht es einen Klärungsweg. Eine sprachlich plausible Zuordnung ersetzt die fehlende Vorgangsinformation nicht.
Der betriebliche Prüfpunkt unterscheidet sich entsprechend: Erreichen dringliche Anliegen rechtzeitig die zuständige Stelle? Verringert die Planung fehlende Teile und unnötige Bestände? Erhält der richtige Kunde den aktuellen Status seines Auftrags? Die technische Ausgabe ist jeweils danach zu beurteilen, welche Wirkung sie im Arbeitsablauf erzielt.

Tab. 2: Drei betriebliche Aufgaben und ihre unterschiedlichen Daten- und Prüfanforderungen

Aufgabe	Notwendige Datengrundlage	Gewünschte Ausgabe	Betrieblicher Prüfpunkt
Anfragen sortieren	Nachrichtentext; Regeln zu Zuständigkeit, Mehrfachanliegen und Dringlichkeit.	Sachgruppe oder mehrere Sachgruppen; Hinweis auf Dringlichkeit; Vorschlag zur Weiterleitung.	Erreichen alle Anliegen rechtzeitig die zuständige Stelle? Wird etwa eine Störungsmeldung neben einer Rechnungsfrage erkannt?
Teilebedarf abschätzen	Frühere Bestellungen und Hinweise auf nicht erfüllten Teilebedarf; klar festgelegter Planungszeitraum.	Geschätzter Teilebedarf für den festgelegten Zeitraum.	Werden Bestellungen und tatsächlicher Bedarf auseinandergehalten? Verringert die Planung fehlende Teile und unnötige Bestände?
Auskunft zum Status vorbereiten	Eindeutig zugeordneter Auftrag; aktueller freigegebener Status; Berechtigung zur Auskunft.	Aus dem Auftragsbestand belegbare Statusauskunft in verständlicher Sprache.	Betrifft die Information den richtigen Kunden und Auftrag? Bei unklarer Zuordnung nachfragen; fehlende Angaben nicht ergänzen.

Quelle: Michaelis, eigene Systematisierung des didaktischen Falls „Drei Wünsche an einen Handwerksbetrieb“; keine empirischen Forschungsdaten.

Abb. 3: Eine KI-Aufgabe vom Eingang bis zur Wirkung beschreiben



Die Modellleistung ist nur ein Teil des Ergebnisses. Datenpflege, Prüfung und Weiterleitung bestimmen die Wirkung mit.

Quelle: Michaelis, eigene Darstellung (didaktisches Beispiel).

Zwischen Ausgabe und Wirkung liegen Zuständigkeiten, Entscheidungen und Kontrollen. Eine richtig erkannte Sachgruppe kann an eine veraltete Adresse weitergeleitet werden. Dann muss die Organisation korrigiert werden, nicht unbedingt die Texterkennung. Die aufgeteilte Darstellung macht den Fehlerort sichtbar.

1.2. Algorithmus, Modell und System auseinanderhalten

Eine Fehleranalyse darf nicht am Sprachmodell stehen bleiben. Der Prüfungsansatz von Siddiq unterscheidet Governance, Daten, Modell, Anwendung und IT-Infrastruktur. Das erklärt, weshalb eine richtig berechnete Ausgabe im falschen Arbeitsablauf trotzdem unbrauchbar sein kann. Für den Einstieg werden zunächst Algorithmus, Modell und System unterschieden; die organisatorischen Verantwortlichkeiten treten hinzu. Auch die KI-Verordnung trennt Modelle mit allgemeinem Verwendungszweck von den daraus aufgebauten Systemen. Eine Modellprüfung deckt daher weder automatisch die gesamte Anwendung noch die Arbeit des Nutzers ab.[5]

1.2.1. Der Algorithmus als festgelegtes Vorgehen

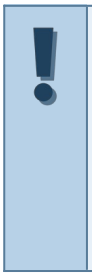
Ein Algorithmus ist eine hinreichend genaue Vorschrift zur Bearbeitung einer Aufgabe. Kochrezept und schriftliche Division veranschaulichen die Verbindung einzelner Schritte. Der Vergleich mit dem Rezept hat allerdings Grenzen: „Nach Gefühl würzen“ ist ohne weitere Übersetzung keine eindeutige Rechenoperation.
Auch lernende Systeme arbeiten nach programmierten Verfahren. Entwickler legen etwa Datenverarbeitung, Fehlermessung und Parameteranpassung fest, ohne für jeden Eingabefall das spätere Ergebnis vorzuschreiben. „Normal programmiert oder KI“ ist deshalb ein irreführender Gegensatz: Programmierung und menschliche Annahmen bleiben beteiligt.

[5] Siddiq, CCZ 2026, 140 (142 f.), Abschn. C.II.1, zur ergänzenden Prüfung der fünf KI-Systemelemente. Art. 3 Nr. 63 und 66 KI-VO; Erwägungsgrund 97 Sätze 1–2 und 6–8 KI-VO zur Abgrenzung von Modell und System. Die vereinfachte Dreiteilung im Buch ist eine eigene didaktische Ordnung, kein wörtlich übernommener Prüfungsstandard.

Bücher nach Erscheinungsjahr zu sortieren erfordert keine Erfahrung mit Leserinteressen. Die Nachfrage im nächsten Monat vorherzusagen ist dagegen eine andere Aufgabe: Ausleihen, Lehrveranstaltungen und Zeitpunkte dienen als Anhaltspunkte. Ein Fehler betrifft hier die Prognose oder Datengrundlage, nicht die Sortierregel.

1.2.2. Das Modell als vereinfachte Abbildung

Ein Modell bildet ausgewählte Beziehungen vereinfacht ab. Ein Stadtplan zeigt etwa Straßen, aber nicht sämtliche Eigenschaften der Stadt. Ebenso kann ein statistisches Modell den Zusammenhang zwischen Buchmerkmalen und Nachfrage darstellen, ohne alle Gründe einzelner Ausleihen zu erfassen.
Bei vielen lernenden Modellen besteht das Ergebnis des Trainings aus Zahlenwerten, die das Verhalten des Modells bestimmen. Solche veränderbaren Größen heißen Parameter. Ein einfaches Modell könnte den erwarteten Bedarf aus drei gewichteten Merkmalen berechnen: Zahl der Einschreibungen, bisherige Ausleihen und Aufnahme in eine Literaturliste. Die Gewichte legen fest, welchen Einfluss die Merkmale in der Berechnung erhalten. Bei großen neuronalen Netzen ist die Beziehung erheblich komplexer; der Grundgedanke veränderbarer Parameter bleibt jedoch hilfreich.



Hinweis: Ein Modell ist kein Quellenkatalog

Eine Datenbank hält gezielt abrufbare Einträge wie Titel, ISBN und Standort bereit. Ein trainiertes Modell enthält gelernte Strukturen, die im Regelfall keinen verlässlichen Einzelabruf ursprünglicher Quellen gewährleisten. Ein genannter Buchtitel kann deshalb aus einer gelernten Verbindung, einem tatsächlich abgerufenen Datensatz oder einer falschen Kombination stammen.

1.2.3. Das System als organisierte Gesamtheit

Ein nutzbares **KI-System** umfasst regelmäßig mehr als das Modell. Dazu gehören etwa Eingabemasken, Datenaufbereitung, Suchwerkzeuge, Rechteverwaltung, Protokollierung und eine Benutzeroberfläche. Hinzu kommen Menschen und Arbeitsabläufe, die bestimmen, welche Daten eingegeben und welche Ergebnisse übernommen werden. Die Forschung zu technischen Folgelasten im maschinellen Lernen weist gerade auf diese Verflechtungen hin: Ein gut funktionierender Modellbaustein kann in einem schlecht beherrschten Gesamtsystem erhebliche Probleme verursachen.[6]
Selbst zuverlässig formulierte Antworten genügen nicht, wenn das Bibliothekssystem fremde Ausleihdaten offenlegt oder eine überholte Benutzungsordnung verwendet. Datenstand und Berechtigungen sind Eigenschaften des Gesamtsystems; die Wahl eines leistungsfähigeren Modells behebt solche Fehler nicht automatisch.

[6] Sculley u. a., Advances in Neural Information Processing Systems 28, 2015, S. 2503–2511, insbesondere Abschn. 2–5.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

Tab. 3: Vier Betrachtungsebenen einer KI-Anwendung

Ebene	Leitfrage	Bibliotheksbeispiel	Typischer Prüfpunkt
Algorithmus	Wie wird vorgegangen?	Nachrichten werden nach einem Verfahren verarbeitet.	Ist das Verfahren für die Aufgabe geeignet?
Modell	Welche gelernte Abbildung wird genutzt?	Ein Modell ordnet Nachrichten Sachgruppen zu.	Wie gut arbeitet es bei neuen Nachrichten?
System	Wie wird die Funktion eingesetzt?	Oberfläche, Modell, Katalog und Mitarbeiter wirken zusammen.	Stimmen Daten, Rechte und Zuständigkeiten?
Arbeitsprozess	Welche Handlung folgt?	Der Mitarbeiter beantwortet oder leitet weiter.	Ist die Entscheidung überprüfbar und korrigierbar?

Quelle: Michaelis, eigene didaktische Systematisierung des Bibliotheksfalls; zum technischen Systemkontext vgl. Sculley et al. (2015), Abschn. 2–5.



Fall 2: Warum ein richtiges Modell eine falsche Rechnung nicht verhindert

Ein Großhändler lässt im Fall Rechnungsnummer, Betrag und Lieferant aus PDF-Dateien erkennen und Buchungen vorschlagen. In der Vorführung liest die Anwendung die ausgewählten Beträge richtig. Damit ist jedoch erst ein Teil des Vorgangs geprüft: Auch fachliche Zuordnung, Stammdatenabgleich und Freigabe können Fehler verursachen.

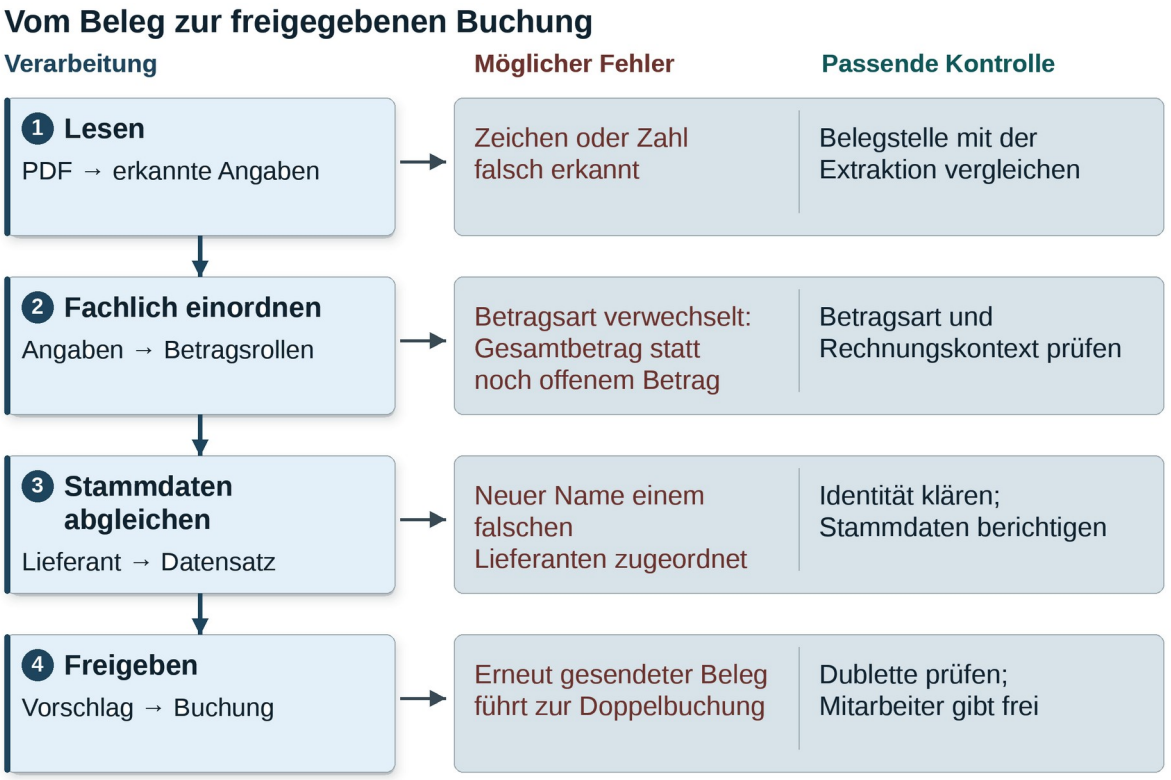
Eine Rechnung kann Netto-, Steuer-, Gesamt- und nach Vorauszahlung noch offene Beträge enthalten. Das Modell mag alle Zahlen richtig lesen; wählt die nachfolgende Zuordnung die falsche Betragsart, bleibt der Buchungsvorschlag falsch. Ein Zeichenerkennungstest erfasst die fachliche Rolle der Zahl noch nicht.

Hat der Lieferant seinen Namen geändert, kann die Zuordnungslogik den korrekt erkannten neuen Namen mit einem ähnlich benannten Geschäftspartner verwechseln. Mehr Training der Erkennung berichtigt veraltete Stammdaten nicht. Benötigt werden Datenpflege und ein Verfahren für unklare Identitäten.

Wird derselbe Beleg nochmals per E-Mail versandt, können zwei korrekte Extraktionen zu einer doppelten Buchung führen. Der Prozess braucht deshalb eine Dublettenprüfung. Die Rechnungsnummer allein muss nicht eindeutig sein; Lieferant, Datum und Betrag können die Zuordnung absichern. Welche Kombination genügt, ist fachlich zu prüfen.

Zur Fehlerklärung werden Beleginhalt, extrahierte Angaben, zugeordneter Datensatz und vorgesehene Handlung getrennt dokumentiert. So wird erkennbar, ob Lesen, fachliche Einordnung, Datenabgleich oder Freigabe fehlerhaft waren. Eine veraltete Lieferantenadresse verlangt Datenpflege; regelmäßig verwechselte Betragsarten können eine Änderung am Erkennungsverfahren erfordern. Der Eingriff sollte den festgestellten Fehlerort treffen.

Abb. 4: Vier Fehlerorte zwischen Rechnungsdatei und Buchung



Fehlerprotokoll: Originalangabe, Extraktion, Zuordnung und beabsichtigte Handlung festhalten. Die Korrektur setzt am jeweiligen Fehlerort an.

Quelle: Michaelis, eigene Darstellung (didaktischer Fall; keine empirischen Forschungsdaten).

1.3. Regeln anwenden oder Zusammenhänge lernen

Feste Regeln legen Bedingungen ausdrücklich fest; lernende Verfahren passen Zusammenhänge anhand von Erfahrung an. Eine Leihanfrage zeigt, warum sich beide Wege auch kombinieren lassen.

1.3.1. Symbolische Verfahren

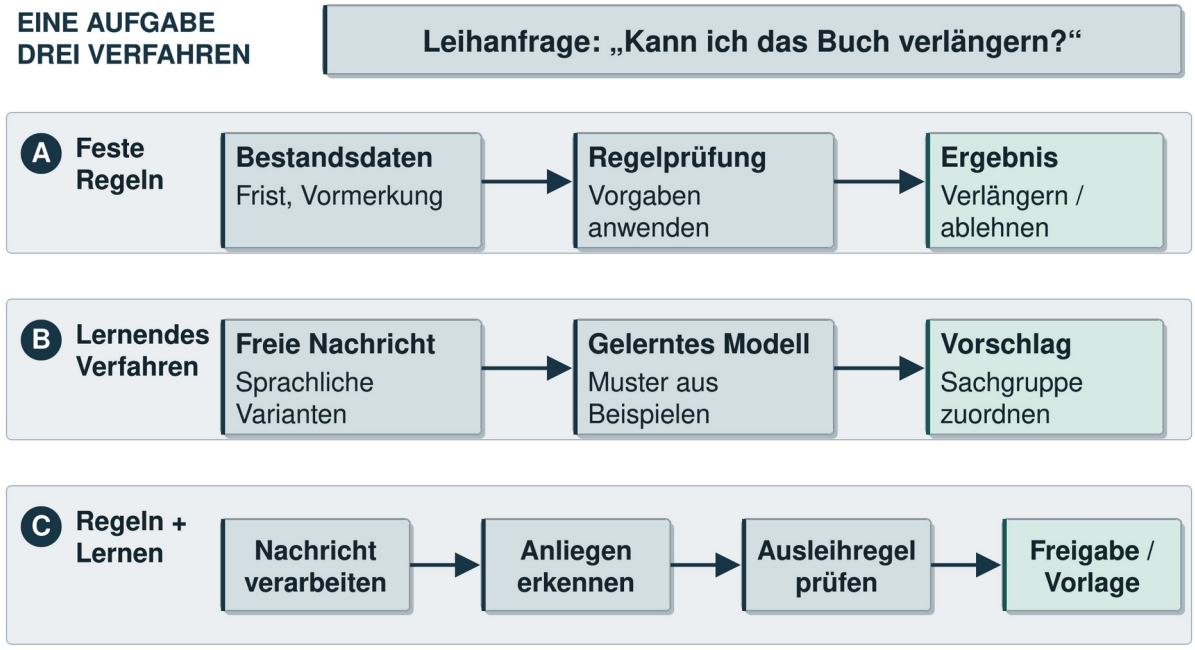
Symbolische Verfahren arbeiten mit ausdrücklich dargestellten Zeichen, Begriffen und Beziehungen. Eine Regel kann beispielsweise festlegen, dass eine Leihfrist nur verlängert werden darf, wenn keine Vormerkung vorliegt. Liegen die erforderlichen Tatsachen eindeutig vor, kann das System die Regel nachvollziehbar anwenden. Der Vorteil besteht darin, dass eine fachlich festgelegte Bedingung unmittelbar erkennbar ist. Der Nachteil liegt häufig davor: Die benötigten Tatsachen müssen korrekt und ausreichend vollständig in eine verarbeitbare Form gebracht werden.
Für „Ich bräuchte das Buch noch bis nach dem Seminar, aber jemand wartet wohl schon darauf“ muss ein Regelsystem Buch, tatsächliche Vormerkung und gemeintes Datum ermitteln. Die Entscheidungsregel kann einfach sein, ihre Voraussetzungen sind es nicht notwendig.
Symbolische Verfahren können mit umfangreichen Regelwerken, logischen Ausdrücken und vielen Handlungsfolgen arbeiten. Eine nachvollziehbare Einzelregel garantiert daher noch keine rechnerisch leicht beherrschbare Suche nach einer Lösung.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

1.3.2. Maschinelles Lernen

Beim maschinellen Lernen wird eine veränderbare Abbildung anhand von Beispielen oder anderer Erfahrung angepasst. Eine gebräuchliche Grundidee verbindet drei Angaben: eine Aufgabe, eine Form von Erfahrung und ein Maß dafür, ob die Leistung besser geworden ist. Ohne diese Angaben bleibt der Satz „Das System lernt“ weitgehend unbestimmt.[7]
Die Bibliothek könnte frühere, fachlich eingeordnete Nachrichten als Erfahrung nutzen und den Anteil richtig zugeordneter neuer Nachrichten als Leistungsmaß wählen. Damit steht die Modellart noch nicht fest: Ein einfaches statistisches Verfahren oder ein neuronales Netz kommen je nach Datenmenge, Fehlerrisiken und Aufwand infrage.
Das Lernergebnis muss keine sprachlich formulierte Regel sein. Ein Mitarbeiter begründet die Zuordnung mit „Es geht um eine Verlängerung“; ein Modell berechnet einen Wert aus Wortkombinationen. Dieselbe Ausgabe kann auf unterschiedlichen Verfahren beruhen.
Eine neue Bibliotheksdienstleistung lässt sich möglicherweise sofort durch eine Regel ergänzen, während ein gelerntes Modell neue Beispiele oder eine andere Einbindung benötigt. Tausende sprachlicher Varianten wiederum sind schwer von Hand zu erfassen. Die Aufgabe entscheidet über den geeigneten Ansatz.

Abb. 5: Drei Wege zur Bearbeitung einer Leihanfrage



Im kombinierten Ablauf: Unsichere Anliegen werden einem Mitarbeiter vorgelegt.

Quelle: Michaelis, eigene Darstellung (didaktisches Beispiel).

Keiner der drei Wege ist für jede Aufgabe überlegen. Die Kombination ist im Beispiel sinnvoll, weil die sprachliche Vielfalt und die verbindliche Ausleihregel unterschiedliche Anforderungen stellen. Eine unsichere Texterkennung darf nicht dazu führen, dass eine eindeutige Benutzungsregel frei erfunden wird. Das lernende Verfahren erschließt die Eingabe; die geregelte Prüfung begrenzt die anschließende Handlung.

[7] Mitchell, Tom M., Machine Learning, New York 1997, Kap. 1, insbesondere S. 2.

1.4. KI, maschinelles Lernen und Deep Learning einordnen

KI, maschinelles Lernen und Deep Learning benennen unterschiedlich weite Verfahrensgruppen. Generierung, Modalität und Handlungsbefugnis beschreiben dagegen andere Eigenschaften, die im Kantinenfall zusammenwirken.

1.4.1. Verwandte Begriffe mit unterschiedlichen Reichweiten

Maschinelles Lernen ist ein wichtiger Teilbereich der KI. Deep Learning bezeichnet wiederum Verfahren des maschinellen Lernens, bei denen mehrere Verarbeitungsschichten eine gelernte Darstellung aufbauen. Die Schichten verändern die Eingangsdaten schrittweise. Bei Bildern können dabei zunächst einfachere und anschließend komplexere Muster für die Aufgabe nutzbar werden. Die Fachliteratur beschreibt dieses Lernen mehrstufiger Darstellungen als zentrales Merkmal des Deep Learning.[8]

Nicht jede KI verwendet Deep Learning. Die Beziehungen lauten: KI als weiteres Forschungsgebiet, darin maschinelles Lernen und darin Deep Learning. Symbolische Verfahren liegen außerhalb dieser letzten Teilmenge. Grenzfälle des fachlichen Sprachgebrauchs bleiben vom jeweiligen Einordnungszweck abhängig.

Ein neuronales Netz ist kein maßstabsgetreues Gehirn. Der Name verweist auf biologische Anregungen; seine technischen Einheiten führen mathematische Operationen aus. Fühlen, Müdigkeit oder Motivation folgen daraus nicht. Die Funktionsweise erklärt Kapitel 3.

1.4.2. Generativ, multimodal und agentisch

Generative KI erzeugt Inhalte wie Text, Bilder oder Ton. Klassifikation weist dagegen einer Eingabe eine Kategorie zu, etwa „Recherchefrage“. Ein generatives Modell kann auch diese Zuordnung übernehmen und anschließend eine Antwort formulieren. Die Erzeugung allein belegt weder Wahrheit noch Zulässigkeit oder Nutzen.

Multimodal bedeutet, dass mehrere Datenarten verarbeitet oder verbunden werden. Eine Bibliotheksanwendung kann etwa ein fotografiertes Inhaltsverzeichnis und eine schriftliche Frage gemeinsam auswerten. Kleine Schrift, Verdeckungen und Widersprüche bleiben mögliche Fehlerquellen; mehr Eingabemöglichkeiten garantieren keine fehlerfreie Erkennung.

Agentisch wird häufig für Systeme verwendet, die zur Verfolgung einer Aufgabe mehrere Schritte ausführen, Werkzeuge aufrufen oder Zwischenergebnisse berücksichtigen. Der Begriff wird unterschiedlich verwendet und sollte konkretisiert werden. Kann die Anwendung nur eine Suche vorschlagen? Darf sie selbst suchen? Darf sie eine E-Mail versenden oder Daten verändern? Für den Nutzer ist diese Befugnisfrage meist aussagekräftiger als die Etikettierung als „Agent“.

[8] LeCun/Bengio/Hinton, Nature 521, 2015, S. 436–444, hier S. 436–438.

Tab. 4: Begriffe nach der Frage ordnen, die sie beantworten

Begriff	Beantwortete Frage	Beispiel	Daraus folgt nicht
Maschinelles Lernen	Wie entsteht die Abbildung?	Anpassung anhand alter Anfragen	Ständiges Lernen während jeder Nutzung
Deep Learning	Welche Verfahrensfamilie wird genutzt?	Mehrschichtiges neuronales Netz	Gleichheit mit einem menschlichen Gehirn
Generativ	Welche Aufgabe steht im Vordergrund?	Entwurf einer Antwort	Sachliche Richtigkeit des Entwurfs
Multimodal	Welche Datenarten werden verbunden?	Textfrage und Buchfoto	Fehlerfreie Wahrnehmung
Agentisch	Wie werden Schritte und Werkzeuge verknüpft?	Suche, Prüfung und Entwurf	Unbegrenzte oder eigenmächtige Befugnis

Quelle: Michaelis, eigene Darstellung auf Grundlage der im Abschnitt genannten Quellen.



Fall 3: Eine Kantine plant den nächsten Mittag

Eine fiktive Betriebskantine möchte Lebensmittelabfälle verringern und zugleich genügend Essen anbieten. Sie bietet zwei Hauptgerichte und eine vegetarische Alternative an. Frühere Bestellungen, Beschäftigtenzahl und Speiseplan dienen als Ausgangsdaten. Nachfrage beschreiben und schätzen, Zubereitung planen sowie Fehlerfolgen bewerten sind dabei getrennte Aufgaben.

Die Kasse erfasst verkaufte, der Küchenleiter benötigt nachgefragte Portionen. Ist ein Gericht früh ausverkauft, wechseln Gäste zur Alternative oder kaufen nichts. Ausschließlich aus Kassenumsätzen könnte das Modell dann eine durch die Produktion begrenzte Menge als Nachfrage lernen. Verkauf und Nachfrage sind deshalb getrennte Zielgrößen.

Die Kantine erfasst zunächst Ausverkaufszeiten und beobachtet an ausgewählten Tagen abgewiesene Wünsche. Dafür können zusammengefasste Angaben je Gericht und Zeitabschnitt genügen. Namen, Zahlungsdetails und vollständige Besuchshistorien wären nicht automatisch aussagekräftiger. Die Datenauswahl muss die nicht bediente Nachfrage erfassen und sich am Planungszweck orientieren.

Als Vergleichslösung verwendet die Kantine den Durchschnitt der letzten vier vergleichbaren Wochen je Wochentag. Für Dienstagsverkäufe von 90, 100, 110 und 100 Portionen ergibt sich der Lehrwert 100. Ausverkauf oder Sonderveranstaltungen können die Vergleichbarkeit beeinträchtigen und werden dokumentiert. Gegen diese einfache Lösung wird der zusätzliche Nutzen einer KI-Prognose geprüft.

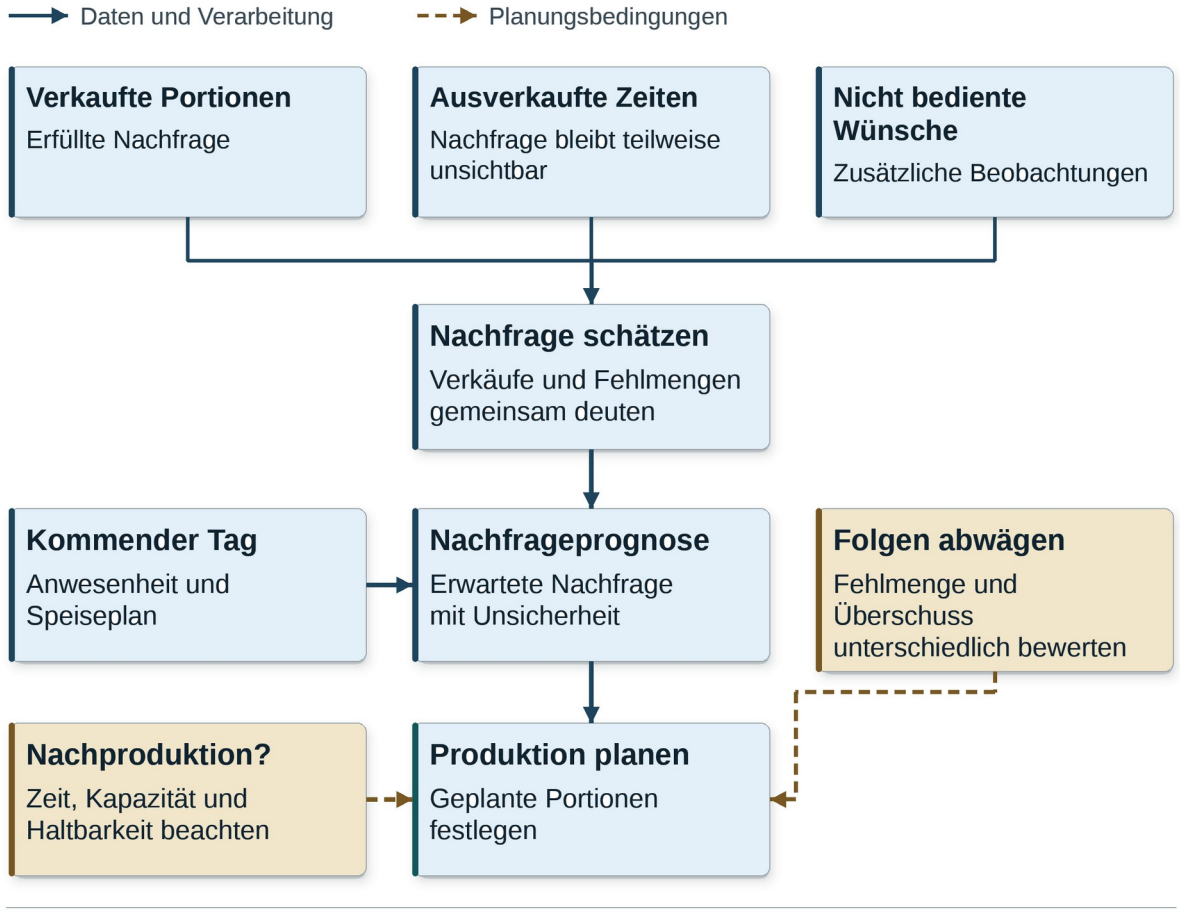
Als Prognosemerkmale kommen erwartete Anwesenheit, Wochentag, Ferienzeiten und Gericht infrage. Ihre Relevanz ist im Fall eine Annahme, kein Forschungsbefund. Hohe Verkäufe eines Gerichts können mit einer Veranstaltung zusammenhängen; ähnliche Gerichte können einander Nachfrage entziehen. Einzelprognosen müssen deshalb zum gesamten Speiseplan passen.

Eine Prognose von 100 Portionen lässt mehrere Produktionsentscheidungen zu: genau 100, eine Reserve von 110 oder zunächst 80 mit Nachproduktion. Die Wahl hängt von Zubereitungszeit, kurzfristiger Nachproduzierbarkeit und sicherer, qualitätsgerechter

Weiterverwendung nicht verkaufter Portionen ab. Planung reicht damit über den erwarteten Zahlenwert hinaus.
Im Rechenbeispiel ist Nachproduktion ausgeschlossen. Eine übrige Portion verursacht drei Euro Aufwand; eine fehlende wird intern mit acht Euro Nachteil bewertet. Letzteres ist eine Planungsannahme, kein objektiver Preis für Unzufriedenheit. Bei 105 nachgefragten Portionen entstehen aus einer Produktion von 100 fünf Fehlmengen, bewertet mit 40 Euro; bei Produktion von 110 kosten fünf übrige Portionen 15 Euro. Andere Nachfrage oder Gewichtung kann das Urteil verändern.
Die Rechnung bestimmt keine allgemein optimale Menge. Sie zeigt unterschiedliche Folgen gleich großer Über- und Unterschätzungen. Ein mittlerer absoluter Prognosefehler belegt deshalb noch keine geeignete Produktionsstrategie. Auch die Grenzen der Kostenannahmen bleiben wichtig: Zufriedenheit, Ernährungsvorlieben und Arbeitsbelastung lassen sich nicht ohne Weiteres auf eine Geldzahl reduzieren.
Die Oberfläche sollte Nachfrageprognose, Unsicherheit, Nachproduktionsmöglichkeit und empfohlene Menge unter den festgelegten Annahmen gemeinsam zeigen. Sonst bleibt etwa bei „110 Portionen“ offen, ob die Zahl die erwartete Nachfrage, einen Sicherheitszuschlag oder eine Kostenabwägung bezeichnet. Eine eindeutige Beschriftung unterstützt die fachliche Entscheidung.
Im Kantinensystem können mehrere Funktionen zusammenwirken: Klassifikation ordnet Kategorien zu, Prognose schätzt eine zukünftige Größe, Optimierung sucht unter Zielen und Bedingungen eine bevorzugte Handlung, und Generierung erläutert den Plan. Jede Ausgabe muss auf ihre eigene Aufgabe bezogen bleiben. Gemeinsam unterstützen die Funktionen Entscheidungen unter Unsicherheit.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

Abb. 6: Kantinenplanung: Verkauf, Nachfrage, Prognose und Produktionsentscheidung



Nach dem Einsatz: Verkäufe, Ausverkäufe und Reste getrennt prüfen. Eine gute Prognose ist noch keine begründete Produktionsentscheidung.

Quelle: Michaelis, eigene Darstellung (didaktischer Fall; keine empirischen Forschungsdaten).

1.5. Was während des Lernens geschieht und was bei der Nutzung

Eine neue Antwort beweist noch kein neues Lernen. Gerade bei juristischen Assistenten muss erkennbar sein, ob eine Information aus Modellparametern, aus dem aktuellen Gespräch oder aus einer zusätzlich durchsuchten Datenbank stammt. Jandek hebt diese Unterscheidung für die Quellenkontrolle im Studium hervor. Sie ist auch für die spätere Berichtigung wichtig: Ein falscher Datenbankeintrag, eine gespeicherte Gesprächsnotiz und eine im Modell verankerte Information verlangen unterschiedliche Eingriffe. Der Fahrradladenfall macht diese Wege sichtbar.[9]

[9] Jandek, JuS 2026, 919 (920 f.), Abschn. I.2.a–b, zur Unterscheidung generativer Modelle und quellenangebundenen Arbeitens; Sesing-Wagenpfeil, KIR 2026, 209 (209 f.), Abschn. II.1.a–b, zur getrennten Prüfung von Trainingsdaten und trainiertem Modell. Technische Begriffe: Art. 3 Nr. 29–33 KI-VO.

1.5.1. Training und Inferenz

> **Begriff: Training und Inferenz**
> Training ist der Vorgang, in dem ein Modell anhand eines festgelegten Verfahrens angepasst wird. Inferenz bezeichnet hier die Anwendung eines Modells auf eine Eingabe, um eine Ausgabe zu berechnen. Die Unterscheidung ist wichtig, weil ein System auch dann neue Antworten erzeugen kann, wenn seine Modellparameter während des Gesprächs unverändert bleiben. Eine wechselnde Ausgabe beweist kein erneutes Training.

Das Bibliotheksmodell kann am Montag eine neue Nachricht mit unveränderten Parametern einordnen. Am Dienstag kann es zusätzlich mitgeteilte Öffnungszeiten einer Sonderausstellung im aktuellen Kontext berücksichtigen. Erst eine gesonderte Parameteranpassung verändert das trainierte Modell; neue Eingaben allein tun dies nicht.
Unterscheiden Sie Trainingsmaterial, aktuellen Kontext und spätere Speicherung oder Weiterverwendung. Letzteres hängt von der konkreten Anwendung und ihren Bedingungen ab, nicht allein von der Modellart. Kapitel 7 zeigt, wie diese Fragen vor dem Einsatz geklärt werden.

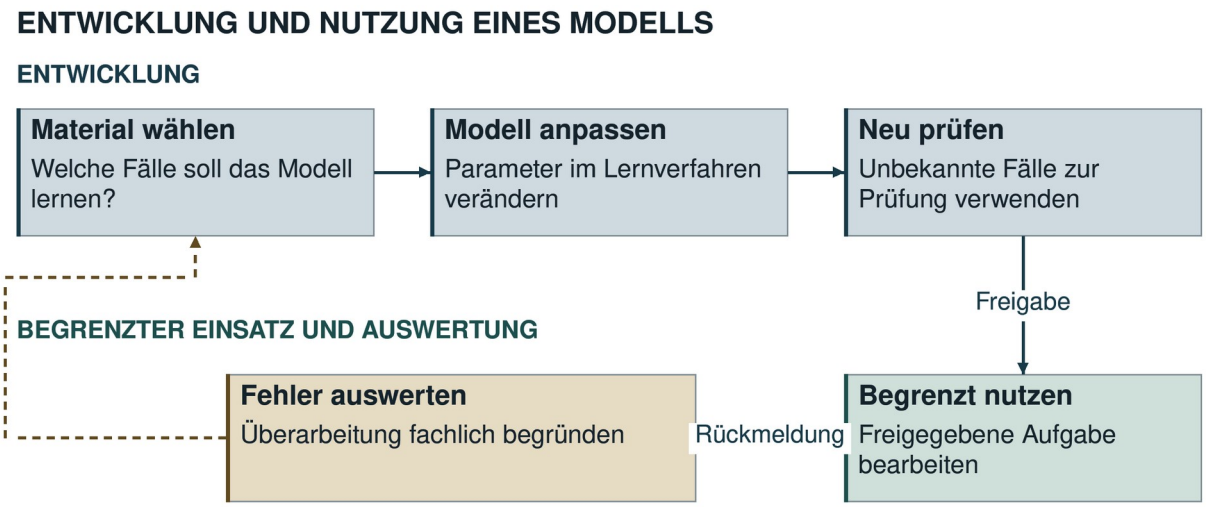
1.5.2. Verschiedene Formen von Erfahrung

Beim überwachten Lernen enthalten Beispiele Eingaben und Zielangaben, etwa Nachrichtentext und Sachgruppe „Ausleihe“. „Überwacht“ bezeichnet diese Zielinformationen, keinen während jeder Rechnung anwesenden Menschen. Uneinheitliche Einordnungen vergleichbarer Nachrichten sind ein Problem der Kennzeichnung, das zusätzliche Rechenleistung allein nicht löst.
Unüberwachtes Lernen sucht Strukturen ohne vorgegebene Zielklassen. Die Bibliothek könnte ähnliche Nachrichten gruppieren. Eine Gruppe ist noch keine fachliche Kategorie: Sie könnte Sprache oder Textlänge statt Anliegen abbilden. Erst die fachliche Auswertung verbindet die Gruppierung mit dem Organisationszweck.
Beim selbstüberwachten Lernen werden Lernaufgaben aus den Daten selbst gewonnen. Ein System kann beispielsweise fehlende oder nachfolgende Teile eines Textes vorhersagen. Dadurch stehen Lernsignale zur Verfügung, ohne dass für jedes Beispiel ein Mitarbeiter eine fachliche Sachgruppe zuweisen muss. Der Ausdruck darf nicht mit einer vollständigen Unabhängigkeit von menschlichen Entscheidungen verwechselt werden: Datenauswahl, Lernziel und Auswertung werden weiterhin gestaltet.
Beim bestärkenden Lernen wird das Verhalten anhand eines Rückmeldungssignals für Handlungen angepasst. Entscheidend ist dabei, welche Rückmeldung das System erhält und welche langfristigen Folgen in der Aufgabe berücksichtigt werden. Ein hoher Messwert kann am eigentlichen Zweck vorbeigehen, wenn die Belohnung falsch gewählt ist. Die Grundstruktur aus Zuständen, Handlungen und Rückmeldungen wird in der einschlägigen Fachliteratur ausführlich entwickelt.[10]

[10] Sutton, Richard S./Barto, Andrew G., Reinforcement Learning: An Introduction, 2. Aufl., Cambridge, Massachusetts 2018, Kap. 1 und 3.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

Abb. 7: Vom Lernmaterial zur geprüften Nutzung und zurück



Durchgezogen: Arbeitsfolge. Gestrichelt: bewusst ausgelöste Überarbeitung; kein automatisches Lernen im Betrieb.

Quelle: Michaelis, eigene Darstellung (didaktisches Beispiel).

Der Rückweg in der Grafik ist eine organisatorische Entscheidung. Nicht jeder beobachtete Fehler rechtfertigt sofort ein neues Training. Vielleicht ist eine Öffnungszeit falsch hinterlegt, eine Eingabemaske missverständlich oder eine Sachgruppe schlecht definiert. Wer bei jedem Problem das Modell austauscht, übersieht solche einfacheren Ursachen.

Fall 4: Ein Fahrradladen lernt nicht bei jedem Gespräch neu

Ein Fahrradkunde pendelt täglich zwölf Kilometer, transportiert gelegentlich ein Kind und hat keinen Aufzug zum Keller. Der Verkaufsassistent berücksichtigt dies bei seinen Vorschlägen für leichte Räder mit Transportlösung. Die passendere Empfehlung beweist keine Parameteränderung: Die Angaben können allein im aktuellen Gesprächskontext stehen.

Speichert eine berechtigte Kundendatenbank den Wunsch nach einem leichten Rad, kann die Anwendung ihn beim nächsten Gespräch abrufen und in den Kontext aufnehmen. Diese scheinbare Erinnerung ist dann ein Datenbankabruf. Angaben lassen sich gezielt berichtigen oder löschen; falsche Notizen oder eine vertauschte Kundennummer würden die Beratung jedoch irreführen.

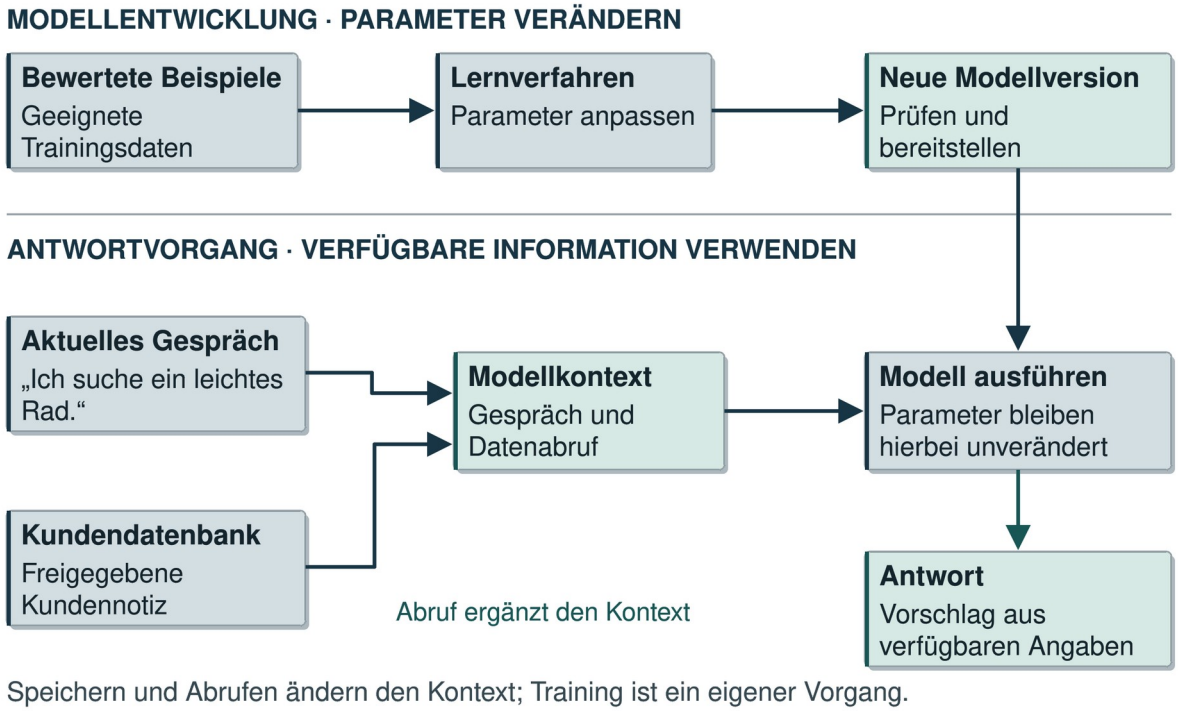
Training wäre ein gesonderter Vorgang: Fachlich bewertete Beratungssituationen könnten zur Modellanpassung dienen. Dafür braucht es ein Lernziel, geeignete Beispiele und eine anschließende Prüfung. Ob eine Kundenäußerung tatsächlich und zulässigerweise dafür verwendet wird, lässt sich nicht aus der Antwort ableiten. Bei personenbezogenen Kundendaten sind insbesondere Zweckbindung und Rechtsgrundlage nach Art. 5 Abs. 1 Buchst. b und Art. 6 Abs. 1 DSGVO sowie je nach Fall Art. 6 Abs. 4 und Art. 13 Abs. 3 DSGVO zu prüfen. Kontextnutzung, Speicherung und Parameteränderung sind getrennt zu dokumentieren.

Bei einer zu schweren Empfehlung kommen mehrere Fehlerorte infrage: übersehene Gesprächsangabe, falsches Kataloggewicht oder eine alte Kundennotiz, die die aktuelle Mitteilung verdrängt. Ein korrigierter Eintrag oder eine klare Rangfolge widersprüchlicher Angaben kann genügen. Die Diagnose geht dem möglichen Neutraining voraus.

Die Fehlerprüfung fragt daher: Welche Information lag bei dieser Antwort tatsächlich

vor? Ein Protokoll von Produktdaten, Gesprächsangaben und freigegebenen Kundennotizen kann dies zeigen. Der Kunde benötigt eine verständliche Begründung anhand seiner Bedürfnisse und der Produktmerkmale sowie einen Hinweis auf offene Punkte.
Die OECD erläutert die mögliche Anpassungsfähigkeit nach der Bereitstellung als ein Merkmal unterschiedlicher KI-Systeme. Sie setzt nicht voraus, dass jedes System bei jeder Nutzung seine Parameter verändert.[11]

Abb. 8: Aktueller Kontext, gespeicherte Daten und Training sind drei verschiedene Wege



Quelle: Michaelis, eigene Darstellung (didaktischer Fall; keine empirischen Forschungsdaten).

1.6. Aus Daten wird nicht automatisch Wissen

Datenauswahl bestimmt, welche Fälle ein Modell überhaupt kennenlernt. Selbst ein zuverlässig erkanntes Muster erklärt noch nicht seine Ursache; der Parkplatzfall zeigt die Folgen für Handlungsempfehlungen.

1.6.1. Daten entstehen durch Auswahl

Eine Sammlung von Nachrichten ist kein unverfälschtes Abbild aller Anliegen. Sie enthält nur, was über diesen Kommunikationsweg eingegangen und gespeichert worden ist. Studenten, die telefonisch nachfragen oder gar nicht wissen, dass es eine Beratung gibt, fehlen möglicherweise. Auch die Festlegung einer Zielangabe enthält Entscheidungen: Gilt eine Nachricht mit zwei Anliegen als ein Fall oder als zwei Fälle? Wird sie nach dem zuerst genannten oder nach dem wichtigsten Anliegen eingeordnet?

[11] OECD, Explanatory memorandum on the updated OECD definition of an AI system, OECD Artificial Intelligence Papers Nr. 8, 2024, S. 6, Abschn. „Role of humans, autonomy and adaptiveness“, sowie S. 8, Abschn. „Building AI systems and models“, DOI: 10.1787/623da898-en.

Datenauswahl legt fest, was ein Modell lernen kann. Welche Personengruppe fehlt, ist aus dem vorhandenen Bestand allein häufig nicht erkennbar. Datenblätter dokumentieren deshalb Entstehung, Zweck, Grenzen und geeignete Verwendung eines Datensatzes.[12]
Sammelt die Bibliothek nur leicht zuordenbare Nachrichten, fehlen unvollständige und mehrdeutige Anfragen im Training. Gute Ergebnisse auf diesem ausgewählten Bestand belegen keine Eignung für den gesamten Eingang. Die Lücke liegt in der Auswahl und lässt sich nicht allein durch ein größeres Modell schließen.

1.6.2. Zusammenhang, Erklärung und Ursache

Ein Modell kann einen statistischen Zusammenhang für eine Vorhersage nutzen, ohne dessen Ursache zu kennen. Wenn in einem bestimmten Zeitraum viele Anfragen zugleich mit einem hohen Ausleihbedarf auftreten, kann dies eine nützliche Beobachtung sein. Daraus folgt jedoch nicht, dass die Anfragen den Bedarf verursachen. Beide können durch den Beginn einer Lehrveranstaltung ausgelöst werden. Eine Handlungsempfehlung verlangt deshalb oft mehr als eine zutreffende Vorhersage.
Studenten mit häufigen Beratungsterminen könnten besser bewertete Arbeiten abgeben. Daraus folgt nicht, dass zusätzliche Pflichttermine die Bewertung verbessern: Motivation, Vorwissen oder Arbeitsart können mitwirken. Ein kausales Untersuchungsdesign muss solche Erklärungen berücksichtigen; Kapitel 5 vertieft die Bewertung.
Das gilt auch bei scheinbar plausiblen Erklärungen eines Modells. Ein generiertes „weil“ ist zunächst Bestandteil einer Ausgabe. Es ist nicht automatisch der nachgewiesene tatsächliche Grund für die Berechnung. Selbst spezielle Erklärungsverfahren benötigen eine genaue Einordnung dessen, was sie sichtbar machen. Die Arbeit von Ribeiro, Singh und Guestrin zeigt beispielsweise einen Ansatz, mit dem einzelne Klassifikationen lokal durch ein einfacheres Modell angenähert und dadurch untersucht werden können.[13]
Für den Anwender lautet die praktische Konsequenz: Fragen Sie, welche Behauptung eine Erklärung tatsächlich trägt. Beschreibt sie einen nachvollziehbaren fachlichen Grund? Zeigt sie einen statistischen Einfluss in einem Modell? Oder ist sie eine sprachlich passende Begründung, deren Entstehung nicht weiter untersucht wurde? Diese drei Möglichkeiten können ähnlich aussehen und dennoch einen sehr unterschiedlichen Beweiswert haben.



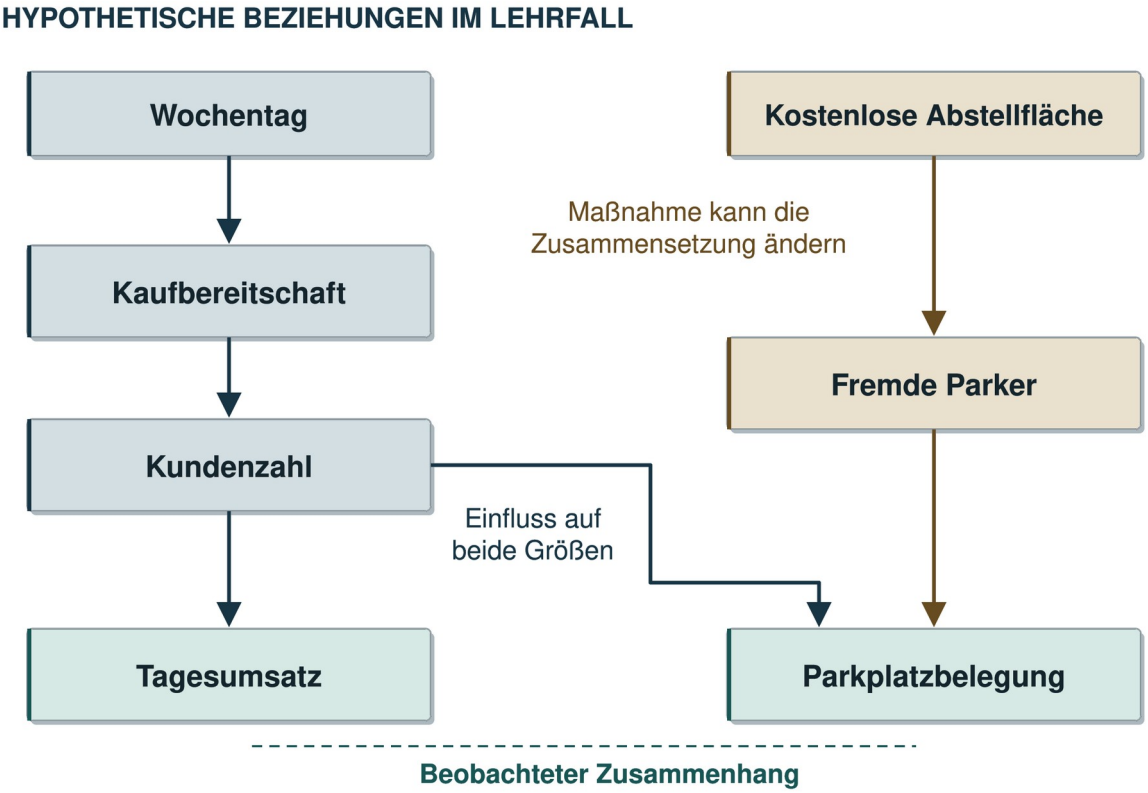
Fall 5: Ein voller Parkplatz erklärt noch keinen Umsatz

Ein Einzelhändler nutzt im Fall neben Wochentag und Öffnungsdauer die Parkplatzbelegung zur Umsatzprognose. Viele Fahrzeuge gehen gewöhnlich mit hohem Umsatz einher. Daraus folgert er, kostenloses Fremdparken werde mehr Autos und damit mehr Umsatz bringen. Er wechselt damit von einer Vorhersage zur unbelegten Wirkungsaussage über eine Maßnahme.
Verkaufsstarke Tage können mehr Autokunden anziehen; Veranstaltungen können zugleich Besucherzahl und Umsatz erhöhen. Beschäftigte benachbarter Betriebe könnten dagegen parken, ohne einzukaufen. Dieselbe Belegung lässt deshalb unterschiedliche Erklärungen zu, zwischen denen der Modellwert nicht entscheidet.
Eine Beziehungsskizze kann Wochentag, Kundenzahl, Parkplatzbelegung und Umsatz verbinden. Fremde Parker erhöhen möglicherweise nur die Belegung. Das beweist keine Ursache, zeigt aber den Informationsbedarf: Welcher Anteil der Parker kauft ein, und ändern Veranstaltungen die Beziehung?
Ein Sprachmodell könnte sowohl mehr Aufmerksamkeit durch zusätzliche Parker als auch die Verdrängung kaufender Kunden plausibel begründen. Beides sind zunächst Hypothesen. Welche Erklärung trägt, muss ein geeignetes Untersuchungsdesign mit

[12] Gebru u. a., Communications of the ACM 64(12), 2021, S. 86–92, insbesondere Abschnitt zum

betrieblichen Informationen klären; sprachliche Überzeugungskraft entscheidet darüber nicht.
Für die Personalplanung kann die Parkplatzbelegung nützlich sein, ohne dass sämtliche Ursachen feststehen. Greift der Händler jedoch durch kostenloses Fremdparken in das Merkmal ein, kann die bisherige Beziehung verloren gehen. Die Prognose für den bisherigen Betrieb begründet daher noch keine Empfehlung für den veränderten.

Abb. 9: Ein gemeinsames Merkmal ist noch kein Beweis für eine Ursache



Pfeile zeigen angenommene Einflüsse, keinen Nachweis. Mehr parkende Fahrzeuge müssen durch einen Eingriff nicht mehr Umsatz erzeugen.

Quelle: Michaelis, eigene Darstellung (didaktischer Fall; keine empirischen Forschungsdaten).

1.7. Von passenden Beispielen zu neuen Fällen

Ein Modell soll neue Fälle bewältigen und darf nicht nur Besonderheiten des Lernmaterials wiedererkennen. Gezielte Varianten und veränderte Betriebsbedingungen prüfen unterschiedliche Grenzen dieser Übertragung.

1.7.1. Generalisierung und Überanpassung

Generalisierung bezeichnet die Übertragung auf geeignete neue Eingaben. Das bloße Wiedererkennen von Trainingsbeispielen genügt dafür nicht. Ähnlich kann ein Student eine Musterlösung auswendig kennen und an einem abgewandelten Fall scheitern. Die Analogie erläutert das Prüfproblem, ohne menschliches und maschinelles Lernen gleichzusetzen.

Fragebogen.

[13] Ribeiro/Singh/Guestrin, Proceedings of KDD 2016, S. 1135–1144, Abschn. 3.

Begriff: Überanpassung und sachfremde Abkürzungen
Überanpassung bindet ein Modell zu stark an Trainingsbesonderheiten und beeinträchtigt die Leistung auf geeigneten neuen Fällen. Sachfremde Abkürzungen sind davon zu unterscheiden: Enthalten Training und Test dieselben Zufallsbeziehungen, können sie in beiden gut funktionieren. Beginnen etwa Recherchefragen meist mit „Sehr geehrtes Bibliotheksteam“, Ausleihfragen mit „Hallo“, könnte das Modell die Anrede übergewichten. Eine höfliche Ausleihfrage würde dann falsch zugeordnet. Die Schwäche zeigt sich erst unter passend veränderten Bedingungen.

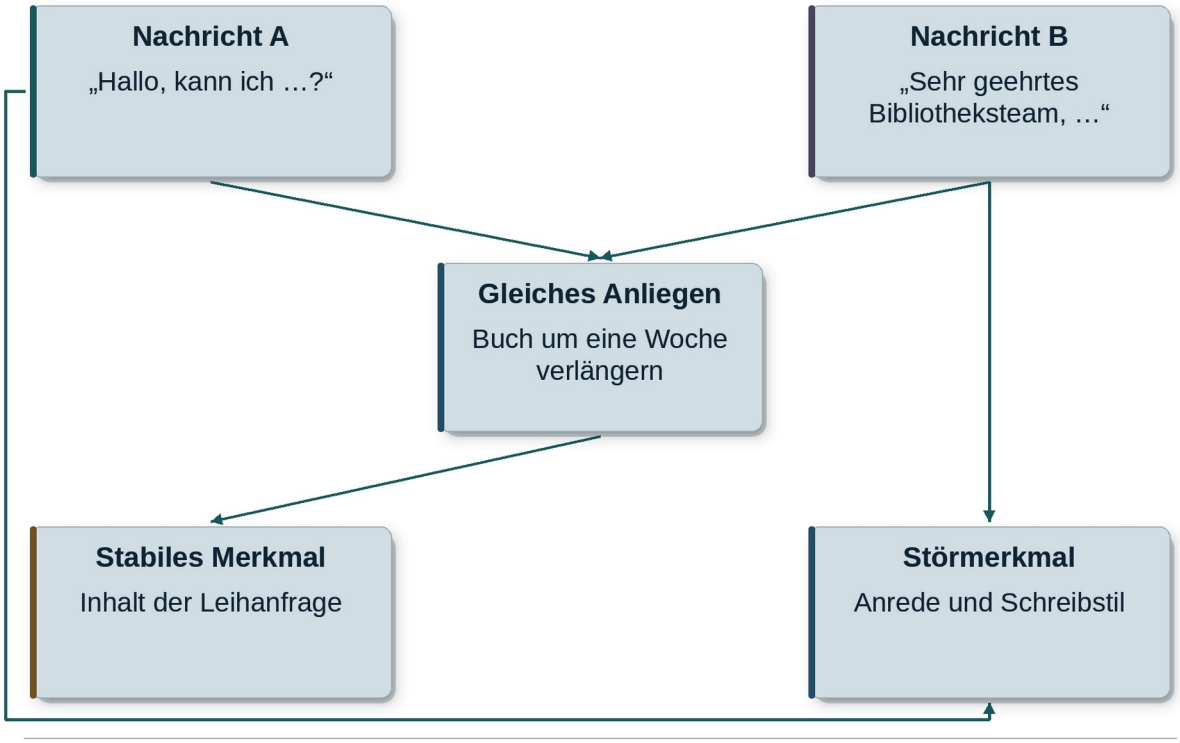
Die Forschung zu sogenanntem Shortcut Learning untersucht solche Abkürzungen: Modelle nutzen leicht verfügbare Merkmale, die in einer bekannten Datensituation erfolgreich sind, unter veränderten Bedingungen aber nicht die beabsichtigte Leistung tragen.[14] Für das Bibliotheksbeispiel folgt daraus ein konkreter Prüfplan. Man formuliert vergleichbare Anliegen mit verschiedenen Anreden, Längen und Schreibweisen. Bleibt die Sachgruppe bei sachlich gleichbleibender Frage stabil, spricht dies für eine brauchbare Robustheit in diesem Test. Eine allgemeine Fehlerfreiheit wäre damit noch nicht bewiesen.

1.7.2. Warum ein guter Test unbekannte Fälle benötigt

Identische oder nur leicht umformulierte Nachrichten in Training und Test können Wiedererkennung als Leistung an neuen Fällen erscheinen lassen. Auch weitergeleitete Kopien beeinträchtigen die Unabhängigkeit. Die Trennung muss daher der beabsichtigten Prüfaussage entsprechen.
Bei zeitlich veränderlichen Aufgaben kann es sinnvoll sein, eine spätere Zeitspanne als Test zu verwenden. Dann wird gefragt, ob ein früher vorbereitetes System mit später eintreffenden Anfragen zurechtkommt. Eine zufällige Aufteilung alter Daten beantwortet möglicherweise eine andere Frage. Welche Trennung geeignet ist, ergibt sich also aus dem geplanten Einsatz. Kapitel 5 vertieft die methodischen Einzelheiten einschließlich Datenleckage, Validierung und Testdaten. Wird das Modell nach jedem Test gezielt auf dessen Fehler abgestimmt, passt es sich zunehmend an den Test an. Entwicklungsentscheidungen und abschließende Prüfung benötigen deshalb getrennte Daten; die methodische Ausgestaltung folgt in Kapitel 5.

[14] Geirhos u. a., Nature Machine Intelligence 2, 2020, S. 665–673.

Abb. 10: Eine scheinbar gute Zuordnung kann auf einem falschen Merkmal beruhen



Quelle: Michaelis, eigene Darstellung (didaktisches Beispiel).

Das Beispiel eignet sich auch für einen ersten eigenen Versuch mit einer KI-Anwendung. Verändern Sie eine nebensächliche Eigenschaft und beobachten Sie die Antwort. Wichtig ist die anschließende Auswertung: Eine einzige stabile Antwort beweist noch keine robuste Funktion. Eine unerwartete Veränderung kann aber eine gezielte Untersuchung auslösen und zeigt, welche Unterschiede für den Test interessant sind.

Fall 6: Eine Qualitätskamera verwechselt das Teil mit seiner Umgebung

Ein Metallteilehersteller fotografierte einwandfreie Teile überwiegend auf hellen Tischen, beschädigte auf dunklen Matten. Im konstruierten Fall kann das Modell deshalb entweder Oberflächenschäden oder den leichter verfügbaren Hintergrund nutzen. Die fachlich zufällige Mattenfarbe ist nicht automatisch von der Berechnung ausgeschlossen; reale Messergebnisse werden hier nicht behauptet.

Eine zufällige Aufteilung dieser Sammlung kann den Zusammenhang zwischen Schaden und Hintergrund in Training und Test erhalten. Die Gesamtquote unterscheidet dann Schadenerkennung und Mattennutzung nicht. Zur gezielten Prüfung werden einwandfreie und beschädigte Teile jeweils auf beiden Unterlagen fotografiert.

Reflektiert die dunkle Matte weniger Licht, können Kratzer schlechter sichtbar werden. Ein Fehler wäre dann noch kein Beweis für die Mattenfarbe als Klassenmerkmal. Beleuchtung, Abstand und Auflösung sollten möglichst konstant bleiben, um diese alternative Erklärung einzugrenzen.

Die Aufnahmebedingungen bestimmen die Reichweite: Eine festgelegte Anlage erlaubt einen engeren Test als wechselnde Handykameras. Trotz gleicher Modellarchitektur müssen beide Einsatzbereiche gesondert geprüft werden. Ein Leistungsversprechen ohne Aufnahmebedingungen ist für den eigenen Betrieb deshalb wenig aussagekräftig.

Werden aussortierte Teile ungeprüft als beschädigt gespeichert, können eigene Fehlentscheidungen als Trainingsbeispiele zurückkehren. Eine unabhängige fachliche Kontrolle ausgewählter Teile unterbricht diese Selbstbestätigung. Abweichungen

zwischen Modell und Prüfer zeigen Untersuchungsbedarf am Modell, an der Prüfung oder an der Kategorie.
Ribeiro und seine Mitautoren entwickeln mit CheckList einen Ansatz zur gezielten Verhaltensprüfung von Sprachmodellen. Die Idee kontrollierter Varianten wird hier auf einen eigenen Bildfall übertragen; dies ist kein in CheckList untersuchtes Produktionsexperiment.[15]

Tab. 5: Gezielte Varianten für die Qualitätskamera

Oberfläche	Hintergrund	Fachlich erwartete Ausgabe
Einwandfrei	Heller Tisch	Einwandfrei
Einwandfrei	Dunkle Matte	Einwandfrei
Beschädigt	Heller Tisch	Beschädigt
Beschädigt	Dunkle Matte	Beschädigt

Quelle: Michaelis, eigene Darstellung (didaktischer Fall).

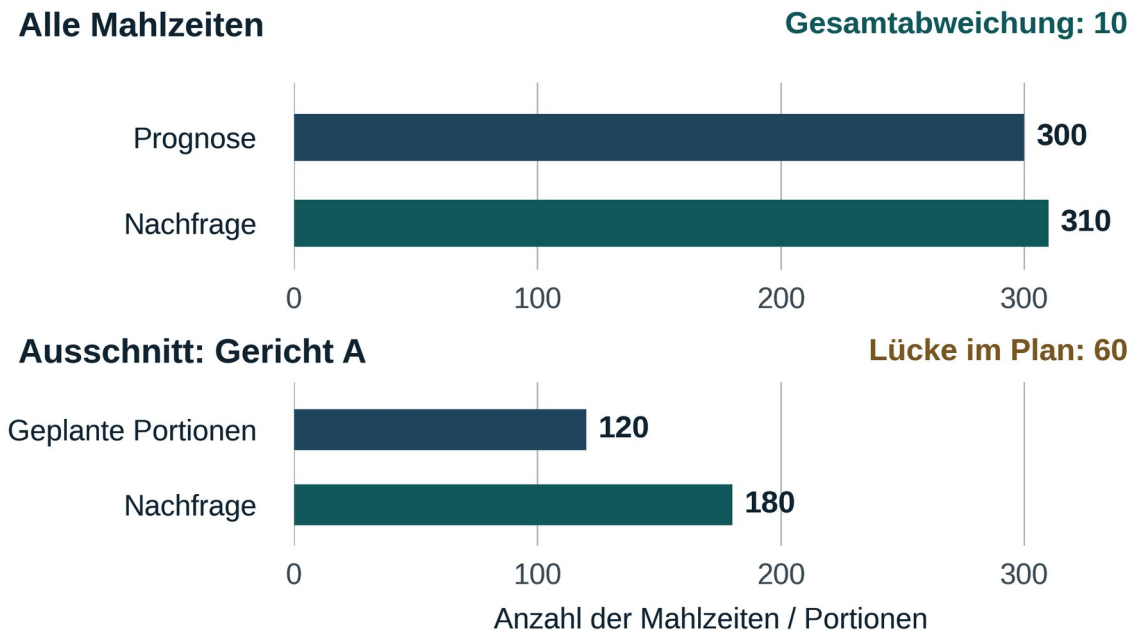
Fall 3a: Eine Messewoche verändert die Bedeutung der Daten

Fortsetzung zu Fall 3: Eine Kantine plant den nächsten Mittag

Während einer Fachveranstaltung ersetzt eine ähnlich große Gästegruppe viele abwesende Stammkunden der Kantine. Unveränderte Anwesenheitszahlen können nun andere Bestellzeiten und Gerichtswahlen bedeuten: Gäste kennen das Angebot nicht und kommen teilweise gemeinsam. Der Fall zeigt eine mögliche Bedeutungsverschiebung eines bekannten Merkmals.
Ob die veränderten Bedingungen die Prognose beeinträchtigen, muss geprüft werden. Vergleichbare Veranstaltungen könnten bereits im Training enthalten gewesen sein; Unterschiede könnten auch gering bleiben. Gute Werktagsergebnisse belegen dies aber nicht. Für Messewochen wird die Aussage deshalb zunächst begrenzt und die Planung gezielter beobachtet, ohne daraus pauschale Unzuverlässigkeit abzuleiten.
Zu klären sind neue Gäste, verkürzte Ausgabezeiten, Gutscheine und zusätzliche Reservierungen. Ein anderer Zahlungsweg kann die Kassenstatistik verändern, ohne die Nachfrage zu ändern; konzentrierte Besuchszeiten können trotz richtiger Gesamtmenge Engpässe auslösen. Mengenplanung und zeitliche Ausgabekapazität müssen daher getrennt geprüft werden.
Der Küchenleiter wertet frühere Veranstaltungstage gesondert aus. Wenige Fälle erlauben keine hohe statistische Sicherheit, können aber Fehlerarten sichtbar machen: Ein Gericht fehlt trotz passender Gesamtmenge, oder die Nachfrage konzentriert sich auf eine halbe Stunde. Zu prüfen sind deshalb Gesamtmenge, Aufteilung auf Gerichte und zeitliche Verteilung. Je nach Befund müssen Prognose oder Ausgabeorganisation angepasst werden.
Weitere Lehrzahlen: Prognostiziert sind 300 Mahlzeiten, nachgefragt werden 310. Die Gesamtabweichung beträgt nur zehn. Von den Wünschen entfallen jedoch 180 auf Gericht A, für das nur 120 Portionen geplant sind; Gericht B bleibt häufiger übrig. Der Verteilungsfehler verlangt eine Auswertung nach Gerichten, die der Tagesgesamtwert nicht ersetzt.

[15] Ribeiro/Wu/Guestrin/Singh, ACL 2020, S. 4902–4912, Abschnitt 2, DOI: 10.18653/v1/2020.acl-main.442.

Abb. 11: Eine kleine Gesamtabweichung kann eine große Lücke bei einem Gericht verdecken



Quelle: Michaelis, eigene Darstellung und Berechnung mit hypothetischen Lehrzahlen des Kantinenfalls: Gesamtprognose 300, Gesamtnachfrage 310; für Gericht A geplant 120, nachgefragt 180. Keine empirischen Forschungsdaten.

Die Reaktion richtet sich nach der Ursache: Eine einmalige Veranstaltung kann eine manuelle Anpassung erfordern, wiederkehrende Veranstaltungen zusätzliche Daten oder Merkmale. Ein neues Modell ist nur eine Möglichkeit. Eine Vorabfrage zur Speisenauswahl kann Unsicherheit stärker verringern als ein aufwendigeres Lernverfahren.
Bereitet der Küchenleiter nach einem Hinweis zusätzliche Portionen zu, beeinflusst sein Eingriff das Ergebnis. Ursprüngliche Prognose, Anpassung und verfügbare Menge werden deshalb getrennt dokumentiert. Sonst würden gute Versorgung oder Fehlmengen womöglich allein dem Modell zugeschrieben.

1.8. Was Leistungsangaben aussagen

Gesamtquote, Erkennungsleistung und Verlässlichkeit einer Sicherheitsangabe beantworten verschiedene Fragen. Seltene Wartungsausfälle zeigen, weshalb ihr Nutzen erst zusammen mit Fehlerfolgen und Bearbeitungskapazität sichtbar wird.

1.8.1. Eine Prozentzahl braucht einen Nenner

Ein Anbieter meldet „95 Prozent Genauigkeit“. Die Zahl klingt eindeutig, enthält aber ohne weitere Angaben wenig Entscheidungsinformation. Auf welchen Fällen wurde gemessen? Was galt als richtige Antwort? Waren schwierige Fälle enthalten? Wie oft kommen die einzelnen Sachgruppen vor? Wer die Zahl versteht, muss auch wissen, was in den Nenner aufgenommen wurde und welche Fehler damit verdeckt werden können.
Betrachten wir ein ausdrücklich hypothetisches Beispiel mit 1.000 Bibliotheksnachrichten. Davon betreffen 900 die Ausleihe und 100 andere Anliegen. Ein Verfahren, das jede Nachricht als Ausleihfrage einordnet, erreicht 900 richtige Zuordnungen und damit 90 Prozent. Es erkennt trotzdem kein einziges anderes Anliegen. Für eine Bibliothek, die gerade schwierige Recherchefragen zuverlässig weiterleiten möchte, wäre das Ergebnis unzureichend. Die hohe Gesamtquote würde den eigentlichen Zweck verfehlen.
Nun vergleichen wir zwei ebenfalls hypothetische Systeme. System A ordnet insgesamt 950 Nachrichten richtig zu, übersieht aber 40 der 100 anderen Anliegen. System B ordnet insgesamt

930 Nachrichten richtig zu, übersieht davon nur zehn andere Anliegen. Welches System günstiger ist, hängt von den Folgen der unterschiedlichen Fehler ab. Wenn eine übersehene Recherchefrage tagelang liegen bleibt, kann B trotz niedrigerer Gesamtquote vorzugswürdig sein. Wenn eine zusätzliche Weiterleitung kaum Aufwand verursacht, lässt sich eine vorsichtigere Zuordnung leichter rechtfertigen.

Tab. 6: Hypothetische Ergebnisse zweier Sortiersysteme

Merkmal	System A	System B	Bedeutung
Alle geprüften Nachrichten	1.000	1.000	Gleicher Testumfang
Insgesamt richtig	950	930	A hat die höhere Gesamtquote.
Andere Anliegen übersehen	40 von 100	10 von 100	B erkennt diese Anliegen zuverlässiger.
Fehler in der Ausleihgruppe	10 von 900	60 von 900	B verursacht mehr zusätzliche Prüfungen.
Auswahlfrage	Gesamtquote bevorzugen?	Kritische Fehler begrenzen?	Der Arbeitszweck entscheidet.

Quelle: Michaelis, eigene Darstellung auf Grundlage der im Abschnitt genannten Quellen; Lehrzahlen sind didaktische Beispiele.

Die Zahlen sind Rechenbeispiele, keine Ergebnisse eines realen Produkttests. Gerade dadurch lässt sich der Gedankengang prüfen: Bei A addieren sich 40 übersehene andere Anliegen und zehn Fehler in der Ausleihgruppe zu 50 Fehlern. Bei B ergeben zehn und 60 zusammen 70 Fehler. Beide Gesamtquoten sind korrekt, aber die Fehler sind verschieden verteilt.

1.8.2. Sicherheitssprache und kalibrierte Schätzungen

Ein System kann seine Antwort sprachlich sehr bestimmt formulieren. Daraus folgt keine gemessene Wahrscheinlichkeit ihrer Richtigkeit. Auch ein ausgegebener Zahlenwert von 0,95 ist zunächst ein Modellwert, dessen Bedeutung geprüft werden muss. **Kalibrierung** bezeichnet vereinfacht die Übereinstimmung zwischen solchen Wahrscheinlichkeitsangaben und der tatsächlich beobachteten Häufigkeit richtiger Ergebnisse in passenden Vergleichsfällen. Die Forschung zeigt, dass gute Klassifikationsleistung und gute Kalibrierung unterschiedliche Eigenschaften sind.[16]

Wenn bei vielen vergleichbaren Antworten mit einer angegebenen Sicherheit von 80 Prozent ungefähr acht von zehn richtig sind, ist diese Gruppe hinsichtlich dieses Maßes gut kalibriert. Daraus folgt nicht, dass jede einzelne Antwort zu 80 Prozent wahr wäre. Eine bestimmte Auskunft ist entweder in den relevanten Punkten zutreffend oder nicht; die Wahrscheinlichkeitsaussage betrifft Unsicherheit vor der Prüfung. Diese Unterscheidung ist besonders wichtig, wenn ein Nutzer eine einzelne folgenreiche Entscheidung treffen muss.

Für den Einstieg genügt eine praktische Regel: Übernehmen Sie keine Sicherheitsbehauptung ungeprüft, nur weil sie numerisch oder selbstbewusst erscheint. Fragen Sie nach dem zugrunde liegenden Verfahren und nach Erfahrungen mit vergleichbaren Fällen. Bei einer folgenarmen Formulierungsaufgabe kann eine unmittelbare Sichtprüfung ausreichen. Bei einer verbindlichen

[16] Guo/Pleiss/Sun/Weinberger, Proceedings of Machine Learning Research 70, 2017, S. 1321–1330, Abschn. 2–3.

Frist, einem medizinischen Hinweis oder einer rechtlichen Bewertung ist ein anderer Prüfmaßstab erforderlich.

Fall 7: Eine Wartungswarnung braucht mehr als eine hohe Trefferquote

Im Logistik-Fall treten in 20 von 1.000 Beobachtungszeiträumen Förderanlagenausfälle auf; 980 bleiben störungsfrei. Ein System, das niemals warnt, erreicht 98 Prozent richtige Gesamtentscheidungen und übersieht jeden Ausfall. Wie schon beim Bibliothekssortierer verdeckt die häufigere Klasse damit das eigentliche Prüfziel.

Das warnende System erkennt 16 der 20 Ausfälle und übersieht vier. Hinzu kommen 64 Fehlwarnungen: Von 80 Warnungen treffen 16 zu. Die Ausfallerkennung beträgt 80 Prozent, der Anteil zutreffender Warnungen 20 Prozent. Erst beide Bezugsgrößen zeigen Erkennung und Prüfaufwand.

Eine kurze Sichtkontrolle macht Fehlwarnungen eher vertretbar als ein mehrstündiger Anlagenstillstand. Ein übersehener Ausfall kann seinerseits den Versand lahmlegen. Die Fehlertypen werden daher nach ihren Folgen bewertet; die Warnschwelle verbindet Erkennungsleistung und Kontrollaufwand.

Eine niedrigere Warnschwelle kann mehr Ausfälle und zugleich mehr Fehlwarnungen erfassen. Der Betrieb vergleicht Einstellungen vorab an geeigneten Fällen und plant die Bearbeitungskapazität mit. Auch hohe Empfindlichkeit hilft wenig, wenn Meldungen nicht rechtzeitig geprüft werden; Kapitel 5 erläutert die Schwellenwahl.

Abb. 12: Seltene Ausfälle: Erkennung und Fehlwarnungen getrennt lesen

	Ausfall: 20 Zeiträume	Kein Ausfall: 980 Zeiträume
Warnung	16 richtig gewarnt	64 Fehlwarnungen
Keine Warnung	4 übersehen	916 richtig ohne Warnung

Eigene Lehrdaten: 1.000 Zeiträume; unter 80 Warnungen sind 16 richtige Warnungen (20 %).

Quelle: Michaelis, eigene Darstellung (didaktischer Fall; keine empirischen Forschungsdaten).

Wenn keine sinnvolle Aussage möglich ist

Fehlen Besucherzahl und vergleichbare Tage für eine Großveranstaltung, kann die Kantinenanwendung zwar eine Zahl erzeugen, aber keine hinreichend begründete Empfehlung. Für solche Informationslücken braucht sie einen besonderen Bearbeitungsweg. Das setzt keine zuverlässige Selbsterkennung sämtlicher Modellunsicherheit voraus.

Die Anwendung könnte die Abweichung benennen: Die Prognose beruht auf gewöhnlichen Werktagen, für morgen ist ein Veranstaltungsbetrieb unbekannter Größe angekündigt. Der Küchenleiter kann zusätzliche Angaben beschaffen oder vorsichtiger planen. Diese begründete Einschränkung hilft mehr als eine Prozentangabe ohne geklärten Bezug.

© Oliver Michaelis · Einführung in die KI, 2. Auflage

Auch Rückstellungen kosten Zeit und begrenzen den Nutzen. Prüfen Sie deshalb, welche Fälle übergeben werden, wann eine Zusatzangabe genügt hätte und welche Übergaben problematische Entscheidungen verhindern. Die menschliche Bearbeitung ist selbst eine zu bewertende Funktion.
Ein geeigneter Ablauf unterscheidet gewöhnliche Fälle, gezielte Rückfragen und fachliche Einzelentscheidungen. Übergänge bleiben möglich: Eine Teilnehmerzahl wird nachgereicht, ein bisher gewöhnlicher Tag verändert sich kurzfristig. Die Anwendung muss solche Änderungen berücksichtigen. Autonomie wird damit für beschriebene Bedingungen eingeräumt und bei relevanten Veränderungen neu beurteilt.

1.9. Menschliche Intelligenz als Vergleich und als Missverständnis

Menschenähnliche Sprache erleichtert die Bedienung, kann aber zu unbelegten Zuschreibungen führen. Beobachtbares Können lässt sich prüfen, ohne damit Bewusstsein oder menschliches Verstehen nachzuweisen.

1.9.1. Leistung und inneres Erleben unterscheiden

Wir verwenden menschliche Verben häufig als praktische Abkürzung. Ein Navigationssystem „weiß“ den Weg, eine Suchmaschine „findet“ eine Antwort und ein Sprachmodell „erklärt“ einen Begriff. Solche Formulierungen erleichtern Gespräche. Schwierigkeiten entstehen, wenn aus der Abkürzung unbemerkt eine Behauptung über innere Zustände wird. Dass eine Anwendung einen verständlichen Trostsatz erzeugt, beweist nicht, dass sie Mitgefühl erlebt. Dass sie ihre Ausgabe korrigiert, beweist nicht, dass sie sich im menschlichen Sinn schämt.
Bereits Turing[17] verschob die Diskussion von einer unklaren allgemeinen Frage nach dem Denken auf ein beschreibbares Verfahren des sprachlichen Vergleichs. Sein Beitrag ist für die Geschichte grundlegend, aber kein einfacher Nachweis, dass überzeugendes Gesprächsverhalten alle Fragen über Geist und Bewusstsein beantwortet.[18] Kapitel 2 erläutert den historischen Zusammenhang. Für die praktische Nutzung genügt zunächst, beobachtbare Leistung und weitergehende Zuschreibungen auseinanderzuhalten.
Searles Gedankenexperiment des chinesischen Zimmers richtet sich gegen die Gleichsetzung formaler Zeichenverarbeitung mit Verstehen. Seine Schlussfolgerungen sind Teil einer philosophischen Debatte und nicht als abschließendes technisches Prüfergebnis zu behandeln.[19] Ein Lehrbuch muss diese Frage nicht endgültig entscheiden, um brauchbare Regeln für den Alltag abzuleiten. Wir können die Richtigkeit einer Auskunft, ihre Belege und ihre Folgen prüfen, ohne zuvor eine vollständige Theorie des Bewusstseins zu entwickeln.

1.9.2. Verstehen als unterschiedlich anspruchsvolle Aussage

Wenn ein Student sagt, er habe eine Definition verstanden, erwarten wir häufig mehr als die Wiederholung des Satzes. Er sollte ein Beispiel erläutern, einen Grenzfall erkennen und den Begriff von einem ähnlichen unterscheiden können. Auch bei einer KI kann man mehrere solcher Aufgaben prüfen. Gute Ergebnisse zeigen dann bestimmte beobachtete Fähigkeiten unter bestimmten Bedingungen. Sie beantworten nicht automatisch, ob das System auf dieselbe Weise versteht wie ein Mensch.

---

17 Alan Mathison Turing (* 23.6.1912; † 7.6.1954) war ein britischer Mathematiker und Logiker. Seine Arbeiten zur Berechenbarkeit und sein Imitationsspiel gehören zu den Grundlagen der Informatik und der Diskussion über maschinelle Intelligenz. Zur Biografie: Copeland, Jack, Alan Turing 1912–1954, in: The Essential Turing, 2004, S. 1–4, https://doi.org/10.1093/oso/9780198250791.002.0003.

18 Turing, Mind 59(236), 1950, S. 433–460, insbesondere Abschn. 1–2 und 6.

19 Searle, Behavioral and Brain Sciences 3(3), 1980, S. 417–424, Gedankenexperiment und anschließende Erwiderungen im selben Heft.