

Daten und Wissen

Warum Daten so wertvoll sind

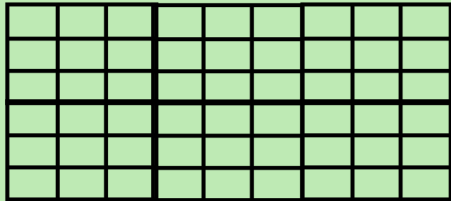
Wie lernt man, Eisverkäufe vorherzusagen?

Beispiel: Eisverkäufe vorhersagen

- Wie sind Sie dabei vorgegangen?
- 3 Gruppen:



Expert*innen gefragt



Daten analysiert

100 x	150 x
20 x	70 x
150 x	0 x
300 x	200 ✓
180 x	350 ✓

Versuch und Irrtum

Gruppe 1: Expert*innen gefragt

- Expert*innen können ihr Wissen z.B. in Form von Regeln oder Anweisungen weitergeben
- oft auch in Form von Schaubildern etc.
- Expertenwissen ist "extrahiert" aus bisherigen Erfahrungen und Kenntnissen der Expert*innen
- in irgendeiner Form abstrahiert
- dieses abstrakte Wissen nennt man manchmal **Modell**

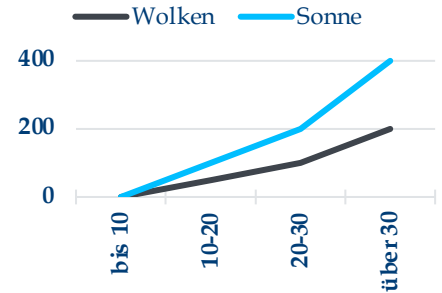
Expertentipp

bis 10 Grad: nichts bestellen
 10 bis 20 Grad: 50 Portionen
 20 bis 30 Grad: 100 Portionen
 über 30 Grad: 200 Portionen

bei Sonne Bestellung verdoppe

Modelle

Expertentipp



**Ein Modell ist eine abstrahierte
(also vereinfachte) Beschreibung
der Realität.**

Gruppe 2: Daten analysiert

- Strategie 1:
 - Regelmäßigkeiten in den Daten entdecken
 - ergibt Wissen ähnlich Expertenwissen
 - abstrahiert
 - im Gegensatz zu Expertenwissen nur aus einer Datenquelle gewonnen
- Strategie 2:
 - suche einen Tag mit ähnlichem Wetter in den Daten, übernehme Menge
 - keine Abstraktion

Bestellregeln

- Winter: 10 Portionen
Frühling und Herbst:
- bei schlechtem Wetter 50 Portionen
 - sonst 100 Portionen
- Sommer:
- bei sehr heißem Wetter 200 Portionen
 - sonst 150 Portionen

Bestellregeln

- bis 10 Grad: nichts bestellen
10 bis 20 Grad: 50 Portionen
20 bis 30 Grad: 100 Portionen
über 30 Grad: 200 Portionen
- bei Sonne Bestellung verdoppeln

Bestellregel

Bestellmenge = Temperatur mal 10

Bestellregel

Bestelle so viel Eis, wie an einem sehr ähnlichen Tag in der Vergangenheit verkauft wurde

Modelle

Gruppe 2: Daten analysiert

- Strategie 1:
 - Regelmäßigkeiten in den Daten entdecken
 - ergibt Wissen ähnlich Expertenwissen
 - abstrahiert
 - im Gegensatz zu Expertenwissen nur aus einer Datenquelle gewonnen

Maschinelles Lernen

- Strategie 2:
 - suche einen Tag mit ähnlichem Wetter in den Daten, übernehme Menge
 - keine Abstraktion

Instanzbasiertes Lernen (instance-based learning, "lazy" learning)

Gruppe 3: Versuch und Irrtum

- anfangs wie Ratespiel
- aber mit jedem Versuch lernt man dazu
- jeder Versuch erzeugt neue Erfahrungen
(= Daten)

Bestellregel

Bestelle 100 Portionen

Bestellregel

Wenn das Wetter schlecht ist,
bestelle 100 Portionen
sonst bestelle 300 Portionen

Bestellregel

Wenn das Wetter schlecht ist,
bestelle 100 Portionen
sonst bestelle 300 Portionen

Im Winter bestelle 200 Portionen
weniger

Gruppe 3: Versuch und Irrtum

- anfangs wie Ratespiel
- aber mit jedem Versuch lernt man dazu
- jeder Versuch erzeugt neue Daten

Verstärkendes Lernen
(reinforcement learning)

Maschinelles Lernen

Daten enthalten Wissen.

**Es muss nur zugänglich gemacht
werden (zum Beispiel durch
Abstraktion).**

Von Daten zu Wissen: Expert*innen

- Expert*innen sammeln durch ihre Arbeit Erfahrung
 - sie werden laufend mit Daten (Fakten) konfrontiert und entdecken dabei Regelmäßigkeiten
 - zusätzlich gewinnen sie Wissen direkt in abstrahierter Form, z.B. durch Theorien und Ergebnisse anderer Expert*innen (die diese wiederum auf der Basis ihrer Daten und Erfahrungen aufgestellt haben)
- Ziel ist häufig das Erstellen von Modellen bzw. Theorien

Von Daten zu Wissen: direkt

- durch die Untersuchung der Daten auf Regelmäßigkeiten und Muster
 - durch Menschen
 - durch den Computer ⇒ **Maschinelles Lernen**
- Beim Maschinellen Lernen wird Wissen durch **Modelle** ausgedrückt
- Modelle basieren häufig auf weniger Daten (insbesondere auf weniger diversen Daten) als bei der Modellbildung durch Expert*innen

Maschinelles Lernen extrahiert Wissen aus Daten. Das extrahierte Wissen nennt man Modell*.

*Zur Abgrenzung von Modellen, die von Expert*innen erstellt wurden, spricht man manchmal explizit von KI-Modellen oder Machine Learning-Modellen.

Von Erfahrungen zu Wissen

- Lernen auf Basis von Versuch und Irrtum
- sehr ähnlich zum menschlichen Lernen
- fängt meist "bei Null" an
- oft indirekt: z.B. Schachspielen
 - nicht: dieser Zug war richtig/falsch
 - sondern: Lernen, welche Aktionen vielleicht erst Züge später zum Sieg geführt haben
- beim Reinforcement Learning:
 - technisch wie Lernen aus Daten (schließlich: Erfahrungen \approx Daten)
 - den lernenden Teil eines solchen KI-Systems bezeichnet man oft als **Agenten**

**Beim Reinforcement Learning lernt
ein System aus den Erfahrungen,
die es macht.**

**In diesem Fall bezeichnet man die
lernende Komponente des Systems
oft als Agenten.**

Dr. Antje Schweitzer

Universität Stuttgart
Institut für Maschinelle Sprachverarbeitung



Universität Stuttgart

Institut für Maschinelle Sprachverarbeitung
Institut für Software Engineering



Reutlingen | Tübingen | Zollernalb



Lizenzbestimmungen

“Daten und Wissen” von Antje Schweitzer, KI B³ / Uni Stuttgart

Das Werk - mit Ausnahme der folgenden Elemente:

- Logos der Verbundpartner und des Förderprogramms
- im Quellenverzeichnis aufgeführte Medien

ist lizenziert unter:

 [CC BY 4.0 \(https://creativecommons.org/licenses/by/4.0/deed.de\)](https://creativecommons.org/licenses/by/4.0/deed.de)

(Namensnennung 4.0 International)

Quellenverzeichnis

Seite 1, Titelfoto: [Markus Winkler \(https://unsplash.com/de/@markuswinkler\)](https://unsplash.com/de/@markuswinkler), ohne Titel, auf [Unsplash \(https://unsplash.com/de/fotos/IrRbSND5EUc\)](https://unsplash.com/de/fotos/IrRbSND5EUc), ist lizenziert unter [Unsplash-Lizenz \(https://unsplash.com/license\)](https://unsplash.com/license).

Bildausschnitt verändert

Seite 3, Clipart: Microsoft