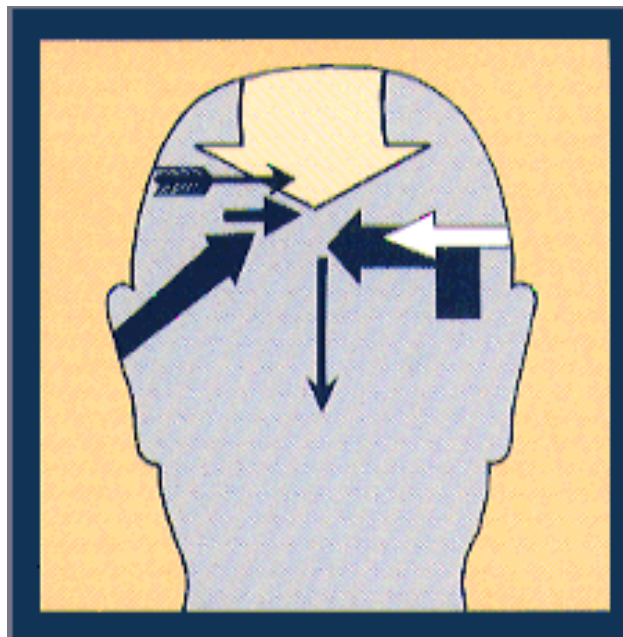


Künstliche Intelligenz und Maschinelles Lernen

KI erforscht,
wie man Computer Dinge machen lassen könnte,
die Menschen im Moment noch besser erledigen.

Elaine Rich



Künstliche Intelligenz (KI)

Artificial Intelligence (AI)

Ziele der KI

Ingenieurwissenschaftlich: Methoden zur maschinellen Lösung von Problemen (Techniken der Wissensrepräsentation, Inferenz, Suchalgorithmen, Klassifikation, Lernen, ...)

Wissenschaftliche Erkenntnis: Objektivierung und Modellierung von kognitiven Prozessen (Wahrnehmung, Denken, Schlussfolgern, Sprachverstehen, Wissenserwerb, ...)

Formalwissenschaftlich: (Weiter-)Entwicklung von Formalismen zur Beschreibung und Bewertung von Problemen und Algorithmen (Logik-Kalküle, Graphtheorie, Lernbarkeitstheorie, ...)

↪ **KI ist notwendigerweise interdisziplinär!**

++ Mathematik ++

++ Informatik ++

++ Neurobiologie ++

++ Kognitive Psychologie ++

++ Kognitive/formale Linguistik ++

++++ Philosophie der Erkenntnis ++++

Einschub: Algorithmus und Programm

Algorithmus: Eine detaillierte Folge von Handlungen, die ausgeführt werden, um eine bestimmte Aufgabe zu erfüllen.

Benannt nach dem Iranischen Mathematiker: *Al-Khawarizmi* (9. Jhd.)



Programm: Auf einem konkreten Computer realisierter Algorithmus.

Themengebiete der KI

Problemlösen und Planen: Suchstrategien und Techniken der Problemreduktion
(Teilzielbildung)

Anwendung: Formelmanipulation, Konfigurieren, Ablaufplanung, Spiele

Inferenz: logische Deduktion, Theorembeweis, Alltagsschliessen (unscharfes/statistisches Schliessen)

Anwendung: Frage-/Antwort-Systeme, Expertensysteme

Wissensrepräsentation: Semantische Netze, Schemata, Graphen

Grundlegend für alle KI-Anwendungen, die auf Wissensbasen aufsetzen

Maschinelles Lernen: Begriffslernen, Entdeckungslernen, Regellernen

Anwendung: Objekterkennung, Klassifikation, *data mining*, Prozess-Steuerung

Automatisches Beweisen/Programmieren: Generierung oder Überprüfung von
Beweisen/Programmen

Anwendung: Verifikation von Programmen

Schöner Bildband:

R. Kurzweil: Das Zeitalter der Künstlichen Intelligenz. Hanser.

↔ Maschinelles Lernen

Lernen ist Vielfältig

- Lernen, nicht auf eine heiße Herdplatte zu fassen.
↔ Reiz-Reaktions-Verknüpfung
- Ein-mal-Eins Lernen.
↔ Auswendig-Lernen (Speicher- und Abruf-Effizienz)
- Begriffe Lernen (Hund oder Katze?).
↔ Klassifikationslernen
- Schachspielen Lernen.
- Lernen, mathematische Beweise durchzuführen, zu programmieren.
- Handlungsstrategien im Alltag lernen.
↔ Regel-Lernen und Strategie-Lernen

↔ **Je flexibler und anpassungsfähiger ein Organismus, desto weniger verfügt er über ein fest vorgegebenes Verhaltensinventar und desto mehr muss (lebenslang) gelernt werden!**

Was ist Lernen?

Lernen ist ein Prozess, der zu einer Performanzsteigerung des Systems führt.

Herb Simon

System: Mensch oder Computer

Performanzsteigerung: Verhaltensänderung im Hinblick auf die Optimierung einer Bewertungsfunktion

Bewertungsfunktion: Anzahl von Fehlern, Zeitbedarf beim Lösen einer Aufgabe

Lernen ist *induktiv*: aus Beispielen (Erfahrung) wird allgemeineres Wissen erworben.

↔ Durch Lernen erworbenes Wissen ist nicht gesichert!

Maschinelles Lernen als Teilgebiet der KI

Technologisches Interesse: Lernfähige Systeme (Belehren statt Programmieren)

Erkenntnis-Interesse: Verstehen des menschlichen Lernens durch die Entwicklung von Computermodellen.

Analytisches Interesse: Gegeben welche Informationen ist welche Art von Wissen mit welchem Aufwand automatisch lernbar?

- “Scheitern” der Expertensystem-Forschung:
Wissen und Regeln zur Lösung von Problemen müssen alle von menschlichen Experten erfragt und computergerecht aufbereitet werden.
- Alternative: Automatischer Regelerwerb aus Beispielen!

Klassifikationslernen

Ohne Begriffe würden wir in einer Informationsflut ersticken:

- Zusammenfassung vieler Objekte zu einem Begriff (gemeinsame Merkmale).
- Statt *Bello, der immer knurrt, Waldi, Struppi, Bello von Herrn Huber, Wotan, Zorro, ...*

Klassifikation eines Exemplars als Hund

(ist zwischen 20cm und 100cm hoch, hat Fell, kann bellen, ...)

Good Lord...tell them we're not cats.



www.cartoonsoup.com
copyright Bill Brown 1999

Begriffslernen mit Entscheidungsbaum-Verfahren

- Gegeben: Durch Merkmale beschriebene Objekte (Beispiele)
- Gesucht: Begriff (Entscheidungsbaum)

Bewegt sich: ja

Bewegt sich: nein

Ist haarig: ja

Ist haarig: ja

Ist klein: ja

Ist klein: ja

↪ **Hund**

↪ **Teddy-Bär**

Abstrakte Darstellung der Objekte (**Merkmalsvektoren**)

| | | | | | |
|---|----|-------|-------|-------|------|
| $\hat{x}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ | Nr | x_1 | x_2 | x_3 | K |
| | 1 | 1 | 1 | 1 | k |
| | 2 | 0 | 1 | 1 | $-k$ |

Abstrakte Darstellung eines Begriffs (**Entscheidungsbaum**)

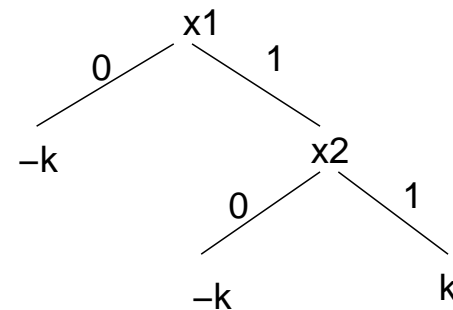
als Term: $x_1(x_2(k - k), k)$

Als bedingte Regeln:

WENN $x_1 = 0$ UND $x_2 = 0$ DANN k

WENN $x_1 = 0$ UND $x_2 = 1$ DANN $-k$

WENN $x_1 = 1$ DANN k



Anwendungen

- Medizinische Diagnostik: Experten beurteilen Daten (Röntgenbilder, Ultraschall, Laborwerte, ...) und klassifizieren nach “Krankheit liegt vor: ja/nein”
Entscheidungsbaum: generalisierte Regeln zur Diagnose von Krankheiten
 - Objekterkennung: Merkmale aus Pixel-Bildern generieren und vor-klassifizieren nach “Objekt vorhanden: ja/nein”
 - Spracherkennung
 - ...
- + Aus einer Menge vor-klassifizierter Beispieldaten werden allgemeinere Regeln induziert!
- + Ein Entscheidungsbaum ist ein Programm zur Klassifikation von Daten!
- + Vorteil für Expertensystem-Forschung: Experten müssen nicht explizit benennen, auf welche Kriterien sie ihr Urteil stützen. Stattdessen: Ermittlung der relevanten Merkmale aus den Urteilen!
-

Warum ist der gezeigte Entscheidungsbaum eine Generalisierung über die Daten?

Ein einfacher Lern-Algorithmus

- Starte mit dem Entscheidungsbaum (*) (Klasse unbekannt).
- Solange sich noch Änderungen ergeben:
 - Wenn das neue Objekt korrekt klassifiziert wird, ändere nichts.
 - Wenn das neue Objekt als (*) klassifiziert wird, trage die Klasse ein.
 - Wenn das neue Objekt falsch klassifiziert wird, nimm das nächste Merkmal in den Baum hinzu.
Trage für die entsprechende Ausprägung die aktuelle Klasse ein und für alle anderen Klassen (*).

Funktioniert auch für

- Merkmale mit mehr als zwei Ausprägungen und
- mehr als zwei Klassen.

Beispiel

Begriff

*

1 1 0 Hund

Hund

(Schaeferhund, der rennt)

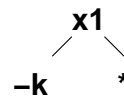
1 1 1 Hund

Hund

(Dackel, der laeuft und bellt)

Ueber-Generalisierung

0 0 0 kein Hund

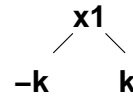


(Haengebauchschwein, das schlaeft)

Differenzierung

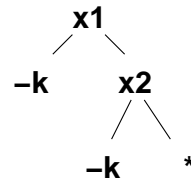
x1: hat Fell nein(0), ja(1)

1 1 0 Hund



(Collie, der spielt)

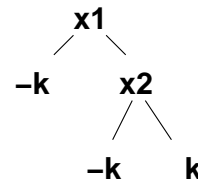
1 0 0 kein Hund



Teddy

x2: bewegt sich nein(0), ja(1)

1 1 0 Hund



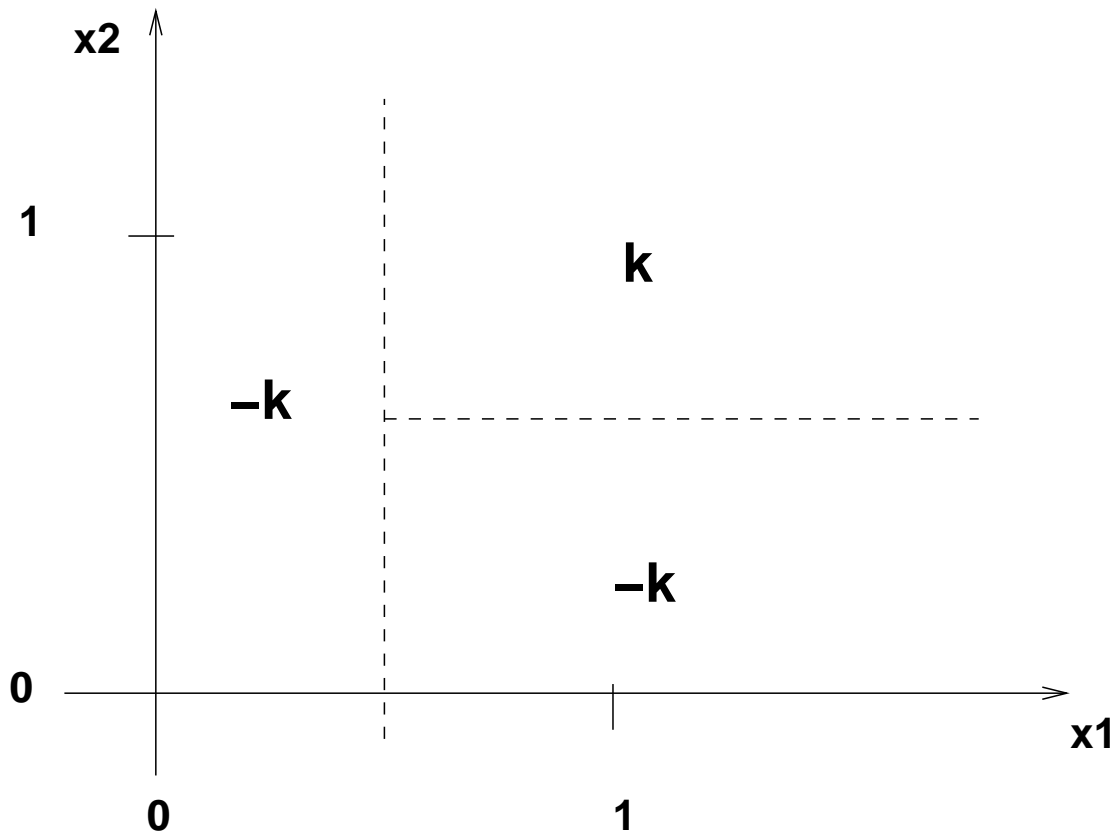
(Bernhardiner, der laeuft)

x3: bellt wird nicht mehr benoetigt

Geometrische Interpretation

Entscheidungsbaum definiert achsenparallele Trennebenen im Merkmalsraum

↔ Generalisierung (Trainingsbeispiele als Stützstellen)

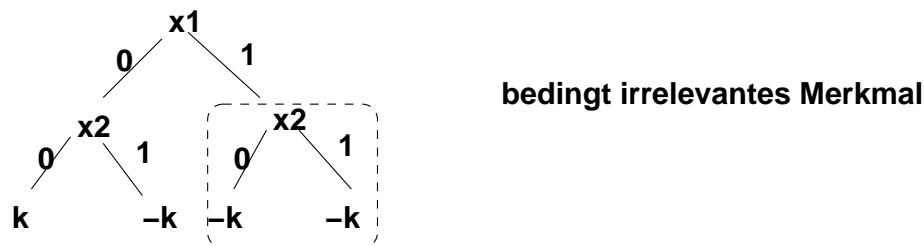
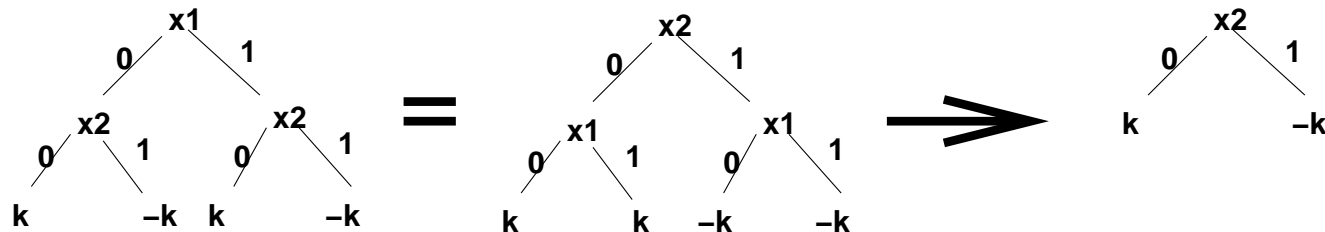
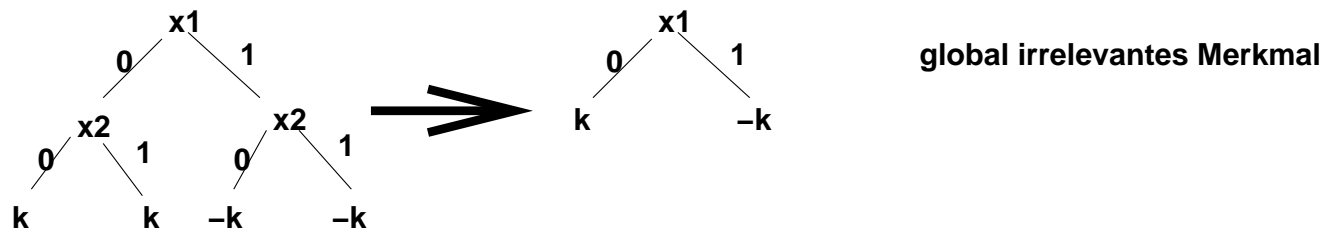


Allgemeinere Lernverfahren: Künstliche Neuronale Netze

↔ *beliebige Klassifizierungsfunktionen*

Effizienzüberlegungen

- Klassifizieren heisst: Im Entscheidungsbaum einen Weg von der Wurzel zu einem Blatt zu gehen.
- Ein Entscheidungsbaum kann *irrelevante Merkmale* beinhalten, die nicht zur Klassifikation beitragen!
- Elimination im Nachhinein (Formale Regeln zur Transformation von Entscheidungsbäumen) oder Vorab-Erkennung (Informationsgehalt).



Informationsgehalt von Merkmalen

Entropie: Ausmaß an Unsicherheit

Anzahl notwendiger Entscheidungsfragen (ja/nein Fragen) zur Reduktion der Unsicherheit bis zur vollständigen Gewissheit

Einheit: *bit*

$$H(k) = - \sum_{i=1}^N p(k_i) \cdot \log_2(p_{k_i})$$

Beispiel: Münzwurf \Rightarrow 1 Entscheidungsfrage

Ereignis k : Was liegt oben

Ereignisalternativen: Kopf (k_1) oder Zahl (k_2)

$$p_{k_1} = p_{k_2} = 0.5 \quad H(k) = -[0.5 \cdot -1 + 0.5 \cdot -1] = 1 \text{ bit}$$

Bei nicht gleich-wahrscheinlichen Ereignissen reduziert sich die Unsicherheit:

| Nr | x_1 | x_2 | K |
|----|-------|-------|---|
| 1 | 0 | 0 | A |
| 2 | 0 | 1 | B |
| 3 | 1 | 0 | B |
| 4 | 1 | 1 | B |

Wahrscheinlichkeit für Klasse A: 0.25

Wahrscheinlichkeit für Klasse B: 0.75

$$H(k) = -[0.25 \cdot \log_2 0.25 + 0.75 \cdot \log_2 0.75] = 0.8 \text{ bit}$$

... Informationsgehalt

Wieviel trägt die Kenntnis der Ausprägung eines Merkmals x_j zur Verringerung der Unsicherheit bei?

(Je mehr, desto besser, also um so früher im Entscheidungsbaum verwenden.)

$$T(x_j) = H(k) - H(k|x_t)$$

- bedingte Entropie: Ausprägung des Ereignisses k gegeben Ausprägung des Merkmals x_t
- T kann minimal 0 sein (Merkmal trägt keine weitere Information)
- und maximal $H(k)$ (Merkmal trägt gesamte Information, bed. Entropie ist 0)

$$H(k|x_t) = - \sum_{x_t} [p(x_t) \cdot \sum_{i=1}^N p(k_i|x_t) \cdot \log_2 p(k_i|x_t)]$$

- innere Summe: H bei Vorliegen von Merkmal x_t
- äußere Summe: Mittelung über alle Ausprägungen von x_t gewichtet mit ihrer Auftretenswahrscheinlichkeit

... Informationsgehalt/Beispiel

| Nr | x_1 | x_2 | K |
|----|-------|-------|---|
| 1 | 0 | 0 | A |
| 2 | 0 | 1 | B |
| 3 | 1 | 0 | B |
| 4 | 1 | 1 | B |

$$p(x_1 = 0) = p(x_1 = 1) = 2/4 = 0.5$$

$$p(A|x_1 = 0) = 0.5 = p(B|x_1 = 0)$$

$$p(A|x_1 = 1) = 0, p(B|x_1 = 1) = 1$$

Ergebnis für $H(k|x_1) = 0.5 \Rightarrow T(x_1) = 0.8 = 0.5 = 0.3$

Zusammenfassung

- Erweiterungen des einfachen Entscheidungsbaum-Verfahrens:
 - überlappende Klassen (Wal-Fisch)
 - Risiko-Schwellen
- Herausforderung: Kontextabhängige Klassifikation (Flasche als Hammer)
↔ menschliches Lernen
- Beispiel für ein sogenanntes “symbolisches” Verfahren: interpretierbare Regeln
(vs. neuronale Netze)
- Aktueller Trend: “situierte Systeme”
statt “nur geistige” Prozesse: Interaktion von Denken und Handeln in der Umwelt
(*aktives Lernen*)

Roboter Fussball

Aktuelle Anwendung in der KI, um zu zeigen, was heute möglich ist.

“Mitte des 21-sten Jahrhunderts wird eine Mannschaft von autonomen humanoiden Roboter-Fussballspielern gegen den amtierenden Fussballweltmeister gewinnen (unter Beachtung der FIFA-Regeln).”

- Echtzeit-Sensorik, Vision: Wo ist der Ball? Mit- oder Gegenspieler? Wo ist das Tor?
- Motor Control, reaktives Verhalten: Zum Ball gehen, Passen, auf's Tor schießen, Gegner blocken
- Planen: z.B. Schritte, um den Ball zu bekommen
- Strategische Entscheidungen: Abseitsfalle, Spiel-
laufbau
- Multi-Agenten Interaktion: Abstimmung zwischen Spielern



CMU Sony Legged Robot Team Gallery

Ein Fussballspieler wird nicht als solcher geboren:

↪ auf allen Ebenen müssen **Fertigkeiten erworben** werden!



virtualWerder: 2

play_on 4168

Macbethians_2: 1

