



FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Seminar Künstliche Intelligenz

Entscheidungsbäume – Eine Einführung

Autor

Michael Dörner

Erlangen, 16. Januar 2014

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

“ *Everything should be made as simple as possible, but not simpler.* ”
– *Albert Einstein*

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums
 - ▶ anhand der Trainingsmenge

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums
 - ▶ anhand der Trainingsmenge
 - ▶ Top-Down

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums
 - ▶ anhand der Trainingsmenge
 - ▶ Top-Down
- ▶ Anwendung eines Entscheidungsbaums

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums
 - ▶ anhand der Trainingsmenge
 - ▶ Top-Down
- ▶ Anwendung eines Entscheidungsbaums
 - ▶ Durchlaufen des Entscheidungsbaum von der Wurzel zu einem der Blätter (somit eindeutiger Pfad)

Definition

Was ist ein Entscheidungsbaum?

Definition

- ▶ Ein Entscheidungsbaum ist ein Baum mit folgenden Eigenschaften:
 - ▶ ein innerer Knoten repräsentiert ein **Attribut**
 - ▶ eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens
 - ▶ ein Blatt repräsentiert eine der **Klassen**
- ▶ Konstruktion eines Entscheidungsbaums
 - ▶ anhand der Trainingsmenge
 - ▶ Top-Down
- ▶ Anwendung eines Entscheidungsbaums
 - ▶ Durchlaufen des Entscheidungsbaum von der Wurzel zu einem der Blätter (somit eindeutiger Pfad)
 - ▶ Zuordnung des Objekts zur Klasse des erreichten Blattes

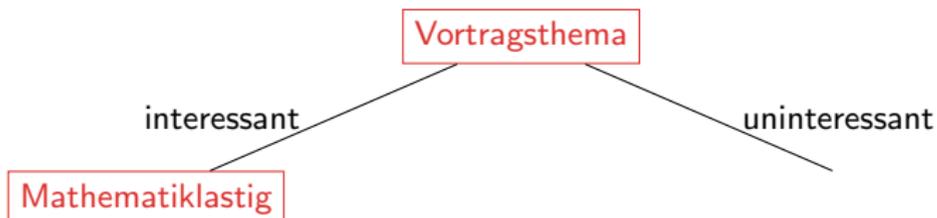
Beispiel (1)

Folge ich dem Vortrag aufmerksam?

Vortragsthema

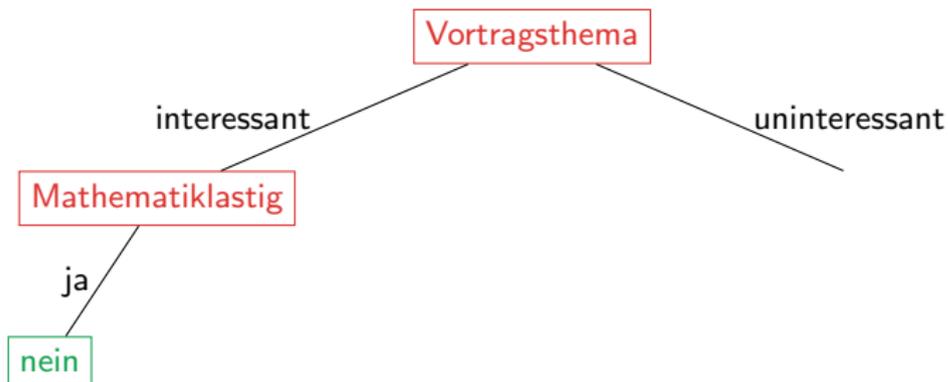
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



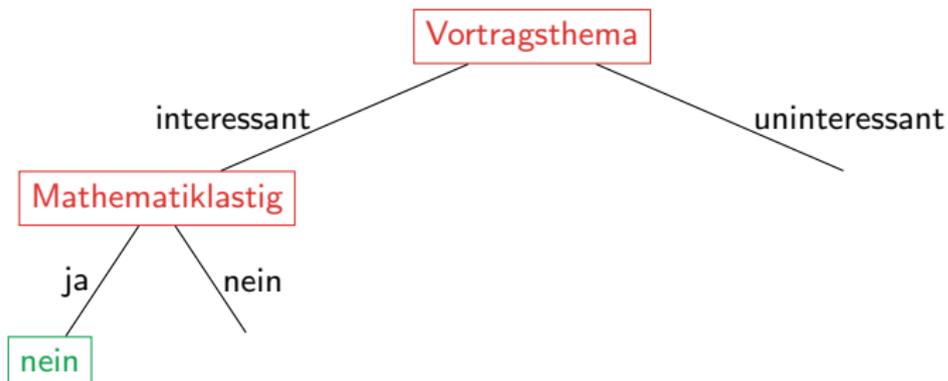
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



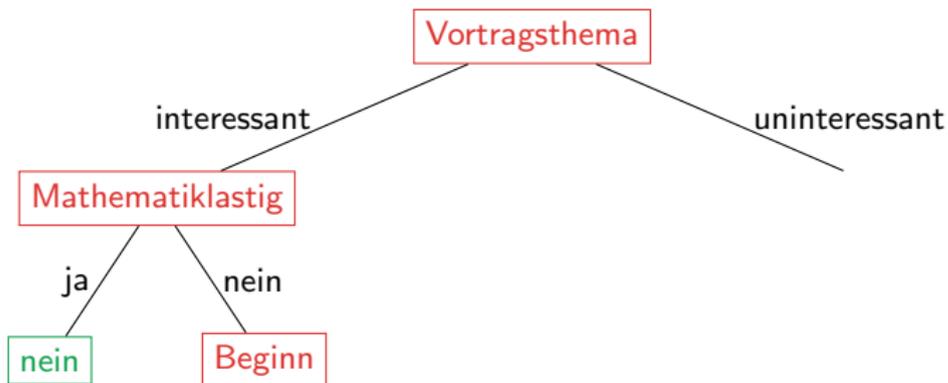
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



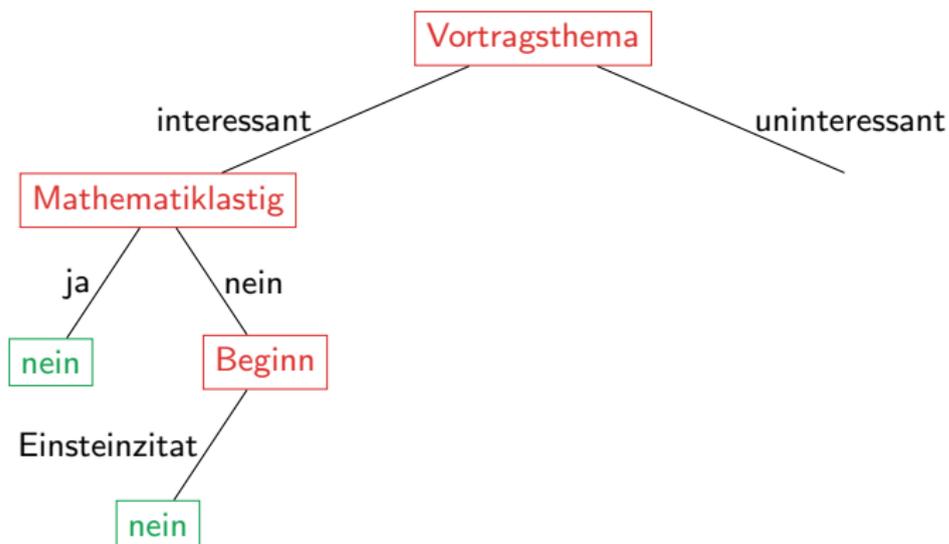
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



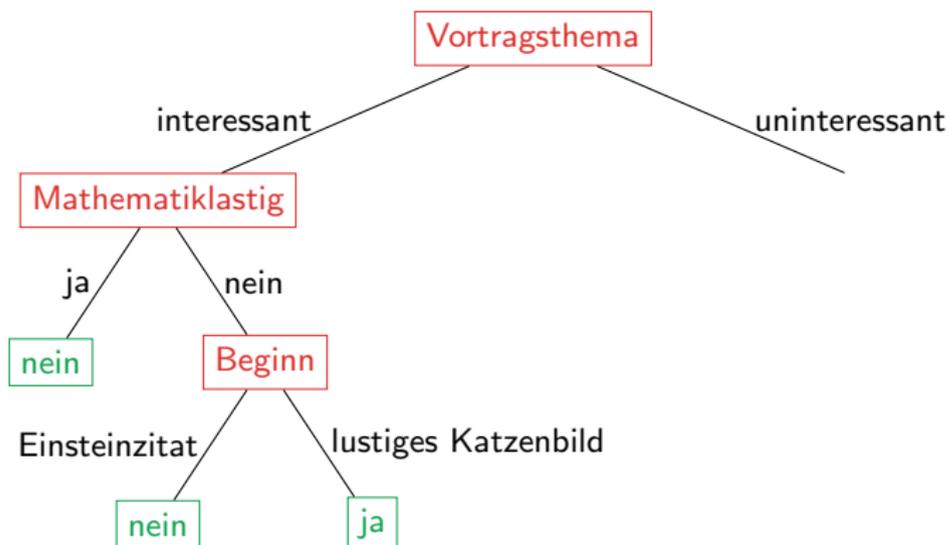
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



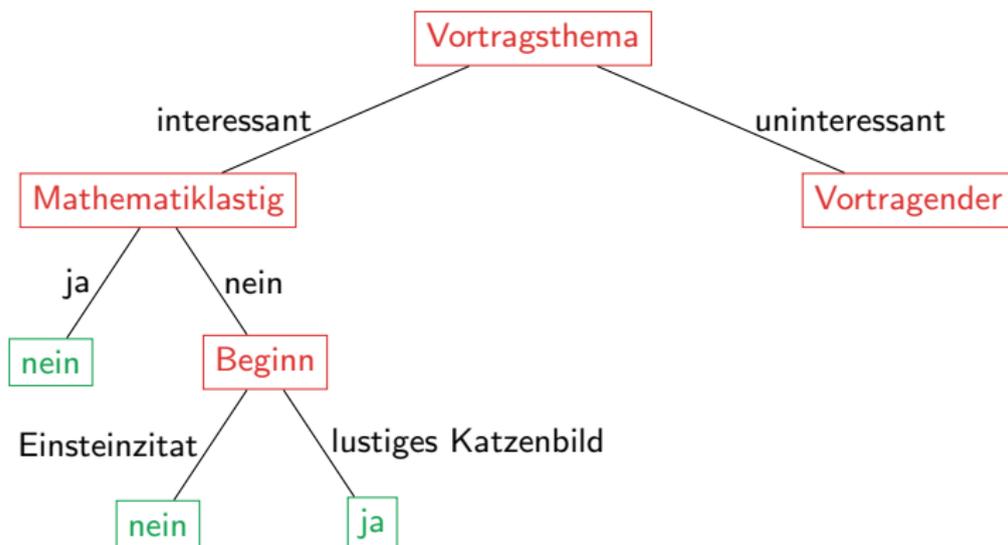
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



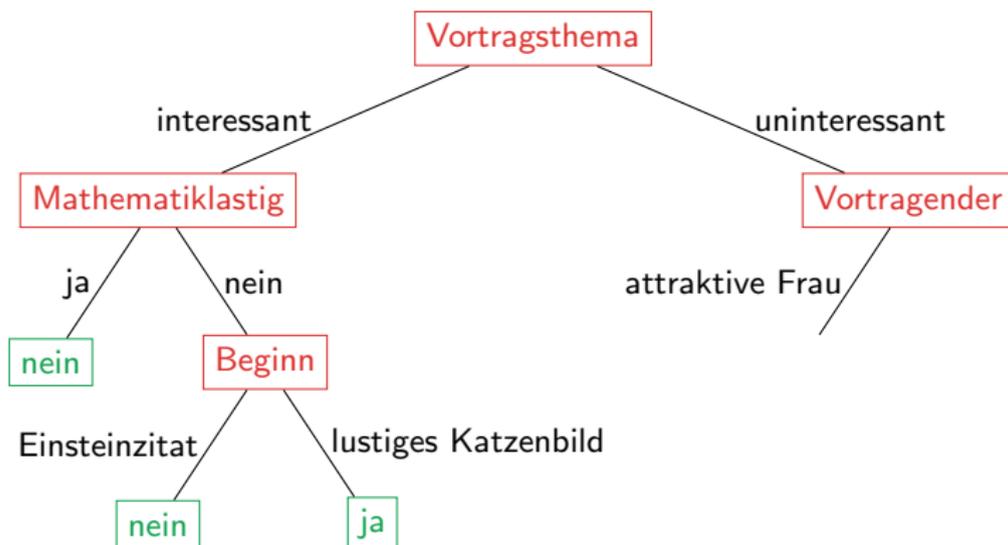
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



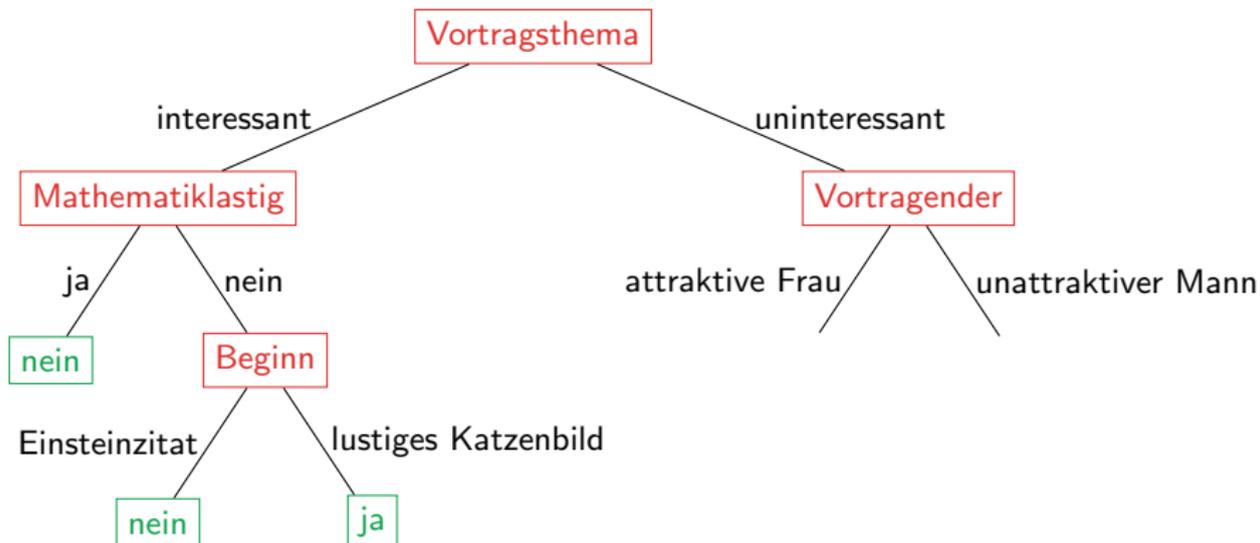
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



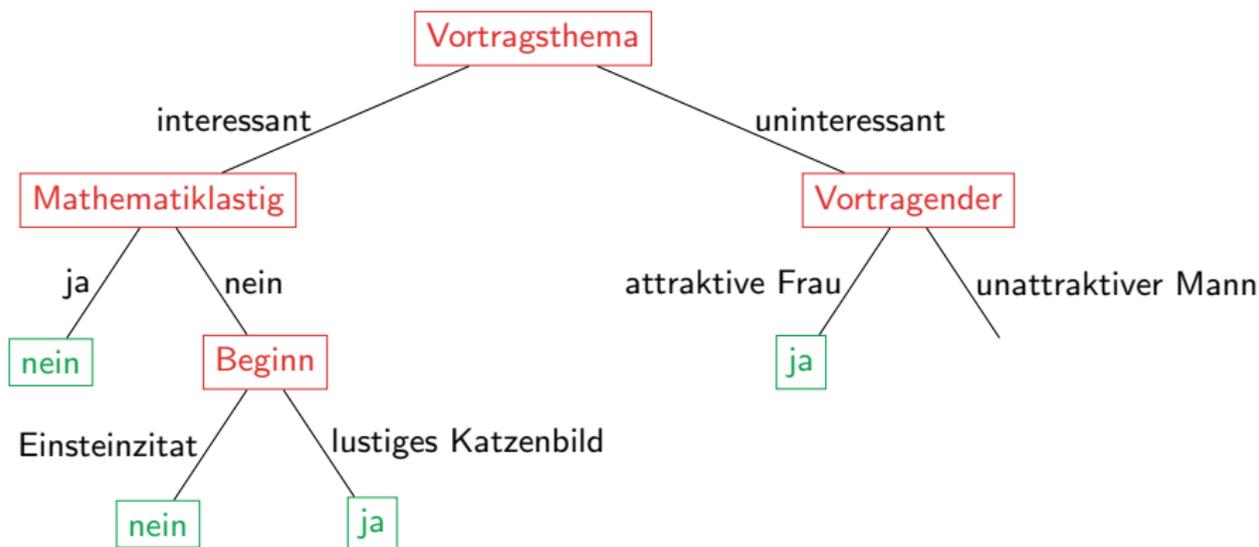
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



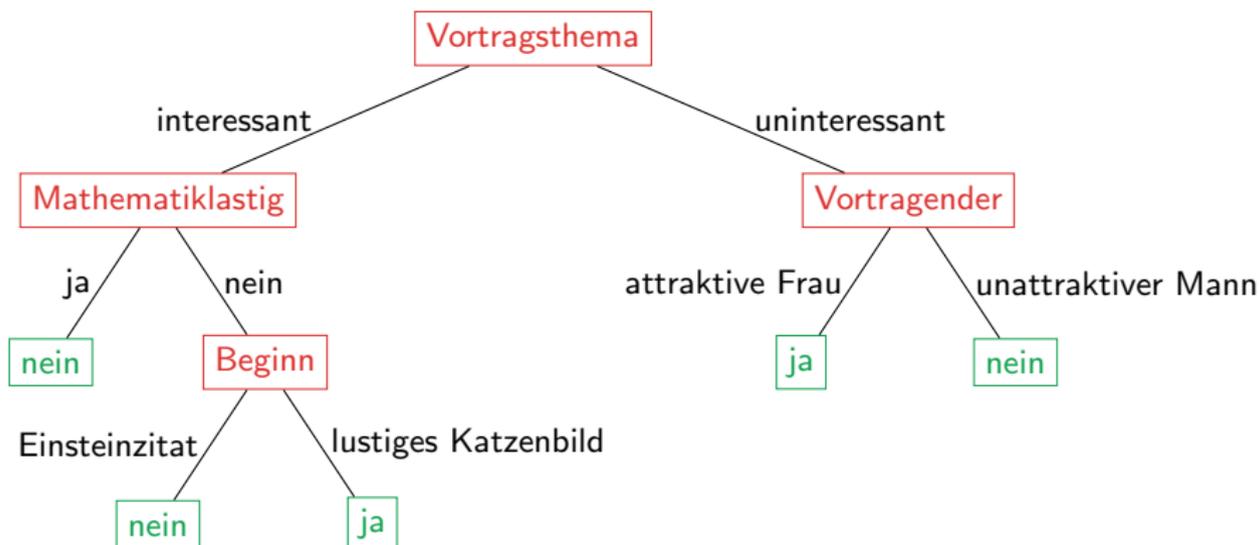
Beispiel (1)

Folge ich dem Vortrag aufmerksam?



Beispiel (1)

Folge ich dem Vortrag aufmerksam?



Beispiel (2)

Wie sieht der Entscheidungsbaum als Tabelle aus?

Vortragsthema	Mathematiklastig	Vortragender	Beginn	Aufmerksam?
interessant	nein	Frau	Katzenbild	ja
uninteressant	ja	Frau	Definition	ja
uninteressant	ja	Mann	Definition	nein
uninteressant	nein	Mann	Definition	nein
interessant	nein	Mann	Katzenbild	ja
interessant	ja	Mann	Definition	nein

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Wie konstruiere ich Entscheidungsbäume generisch?

Algorithmus 1 : Konstruktion von Entscheidungsbäumen

1. Wähle ein (bestes) Attribut

Wie konstruiere ich Entscheidungsbäume generisch?

Algorithmus 2 : Konstruktion von Entscheidungsbäumen

1. Wähle ein (bestes) Attribut
2. Erstelle Knoten für dieses Attribut

Wie konstruiere ich Entscheidungsbäume generisch?

Algorithmus 3 : Konstruktion von Entscheidungsbäumen

1. Wähle ein (bestes) Attribut
 2. Erstelle Knoten für dieses Attribut
 3. Füge für jeden Attributwert eine Verzweigung zu dem neuen Knoten
-

Wie konstruiere ich Entscheidungsbäume generisch?

Algorithmus 4 : Konstruktion von Entscheidungsbäumen

1. Wähle ein (bestes) Attribut
 2. Erstelle Knoten für dieses Attribut
 3. Füge für jeden Attributwert eine Verzweigung zu dem neuen Knoten
 4. Partitioniere die Trainingsdaten entsprechend der Attributwerte
-

Wie konstruiere ich Entscheidungsbäume generisch?

Algorithmus 5 : Konstruktion von Entscheidungsbäumen

1. Wähle ein (bestes) Attribut
 2. Erstelle Knoten für dieses Attribut
 3. Füge für jeden Attributwert eine Verzweigung zu dem neuen Knoten
 4. Partitioniere die Trainingsdaten entsprechend der Attributwerte
 5. Wiederhole 1. – 4., bis alle Daten im neuen Knoten der gleichen Klasse angehören
-

Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

Beispielkonstruktion eines Entscheidungsbaums (1)

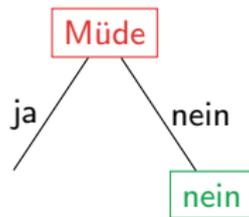
Möchte ich jetzt schlafen?

Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja

Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

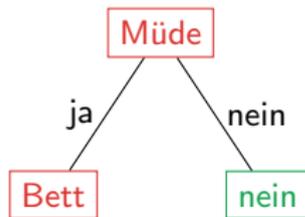
Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja



Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

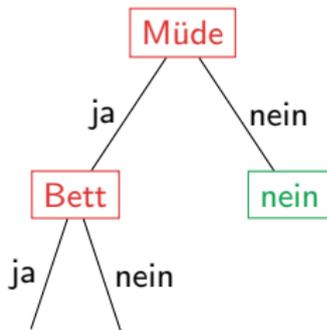
Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja



Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

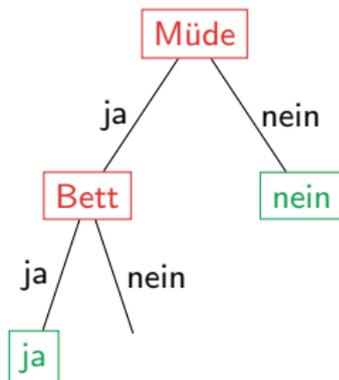
Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja



Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

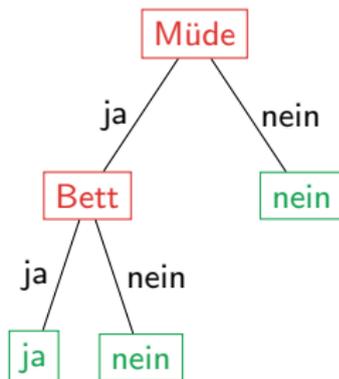
Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja



Beispielkonstruktion eines Entscheidungsbaums (1)

Möchte ich jetzt schlafen?

Müde	Bett	Schlafen?
nein	nein	nein
nein	ja	nein
ja	nein	nein
ja	ja	ja



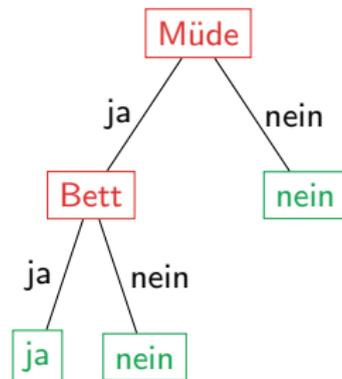
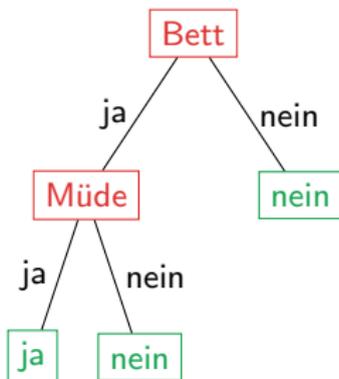
Beispielkonstruktion eines Entscheidungsbaums (2)

Gibt es einen weiteren/besseren Entscheidungsbaum?

Beispielkonstruktion eines Entscheidungsbaums (2)

Gibt es einen weiteren/besseren Entscheidungsbaum?

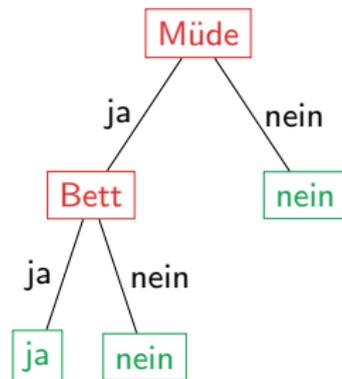
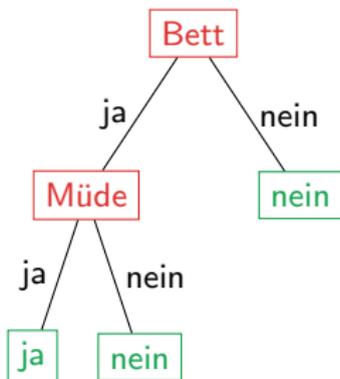
Ja, offensichtlich, denn



Beispielkonstruktion eines Entscheidungsbaums (2)

Gibt es einen weiteren/besseren Entscheidungsbaum?

Ja, offensichtlich, denn



Welches ist das beste Attribut?

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Warum haben Atomkraftwerke Kühltürme?





Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Definition

Den mittleren Informationsgehalt $H(P)$ einer Wahrscheinlichkeitsverteilung P über einer endlichen Menge S bezeichnet man als Entropie von P :

$$H(S) = - \sum_{i=1} P(C_i) \log_2 P(C_i)$$

mit $P(C_i)$ als Auftrittswahrscheinlichkeit der Klasse C_i in S .

Informationsgehalt & Informationsgewinn

Definition

Den mittleren Informationsgehalt $H(P)$ einer Wahrscheinlichkeitsverteilung P über einer endlichen Menge S bezeichnet man als Entropie von P :

$$H(S) = - \sum_{i=1} P(C_i) \log_2 P(C_i)$$

mit $P(C_i)$ als Auftrittswahrscheinlichkeit der Klasse C_i in S .

Definition

Um den Informationsgewinn (*Information Gain*) von Attribut A zu quantifizieren, bilden wir die Differenz der ursprünglichen Information und der Restinformation:

$$G(S, A) = H(S) - \sum_{i \in \text{Werte}(A)} \frac{|S_i|}{|S|} H(S_i)$$

mit

- ▶ $\text{Werte}(A)$ als alle Ausprägungen von A und
- ▶ S_i als Teilmenge von S , wobei A den Wert i annimmt.

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

- ▶ 'aaaa':

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

▶ 'aaaa':

▶ Wahrscheinlichkeitsverteilung:

$$P('a') = \frac{4}{4}$$

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

▶ 'aaaa':

▶ Wahrscheinlichkeitsverteilung:

$$P('a') = \frac{4}{4}$$

▶ Entropie:

$$H('aaaa') = - \left(\frac{4}{4} \log_2 \frac{4}{4} \right) = 0$$

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

▶ 'aaaa':

- ▶ Wahrscheinlichkeitsverteilung:

$$P('a') = \frac{4}{4}$$

- ▶ Entropie:

$$H('aaaa') = - \left(\frac{4}{4} \log_2 \frac{4}{4} \right) = 0$$

▶ 'aabc':

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

▶ 'aaaa':

- ▶ Wahrscheinlichkeitsverteilung:

$$P('a') = \frac{4}{4}$$

- ▶ Entropie:

$$H('aaaa') = - \left(\frac{4}{4} \log_2 \frac{4}{4} \right) = 0$$

▶ 'aabc':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = \frac{2}{4}, \quad P('b') = P('c') = \frac{1}{4}$$

Beispiel (1)

Welchen Informationsgehalt (Entropie) haben folgende Nachrichten?

▶ 'aaaa':

- ▶ Wahrscheinlichkeitsverteilung:

$$P('a') = \frac{4}{4}$$

- ▶ Entropie:

$$H('aaaa') = - \left(\frac{4}{4} \log_2 \frac{4}{4} \right) = 0$$

▶ 'aabc':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = \frac{2}{4}, \quad P('b') = P('c') = \frac{1}{4}$$

- ▶ Entropie:

$$H('aabc') = - \left(\frac{2}{4} \log_2 \frac{2}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right) = 1.5$$

Beispiele (2)

▶ 'abcd':

Beispiele (2)

- ▶ 'abcd':
 - ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = P('b') = P('c') = P('d') = \frac{1}{4}$$

Beispiele (2)

▶ 'abcd':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = P('b') = P('c') = P('d') = \frac{1}{4}$$

- ▶ Entropie:

$$H('abcd') = - \left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right) = 2$$

Beispiele (2)

▶ 'abcd':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = P('b') = P('c') = P('d') = \frac{1}{4}$$

- ▶ Entropie:

$$H('abcd') = - \left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right) = 2$$

▶ 'aaaabcdefg':

Beispiele (2)

▶ 'abcd':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = P('b') = P('c') = P('d') = \frac{1}{4}$$

- ▶ Entropie:

$$H('abcd') = - \left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right) = 2$$

▶ 'aaaabcfg':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = \frac{4}{10}, \quad P('b') = P('c') = P('d') = P('e') = P('f') = P('g') = \frac{1}{10}$$

Beispiele (2)

▶ 'abcd':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = P('b') = P('c') = P('d') = \frac{1}{4}$$

- ▶ Entropie:

$$H('abcd') = - \left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right) = 2$$

▶ 'aaaabcfg':

- ▶ Wahrscheinlichkeitsverteilungen:

$$P('a') = \frac{4}{10}, \quad P('b') = P('c') = P('d') = P('e') = P('f') = P('g') = \frac{1}{10}$$

- ▶ Entropie:

$$H('aaaabcfg') = - \left(\frac{4}{10} \log_2 \frac{4}{10} + 6 \cdot \left(\frac{1}{10} \log_2 \frac{1}{10} \right) \right) = 2,52193$$

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Pseudocode

Algorithmus 6 : ID3 Algorithmus

1. Wähle ~~ein (bestes) Attribut~~ das Attribut mit größtem Informationgewinn
 2. Erstelle Knoten für dieses Attribut
 3. Füge für jeden Attributwert eine Verzweigung zu dem neuen Knoten
 4. Partitioniere die Trainingsdaten entsprechend der Attributwerte
 5. Wiederhole 1. – 4., bis alle Daten im neuen Knoten der gleichen Klasse angehören
-

Beispiel

Informationsgehalt von Krankheit

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Beispiel

Informationsgehalt von Krankheit

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

$$H(S) = - \sum_{i=1} P(C_i) \log_2 P(C_i) = -5/10 \log_2 5/10 - 5/10 \log_2 5/10 = 1$$

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber
hoch
mittel
niedrig
ohne
Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl
hoch	3
mittel	2
niedrig	3
ohne	2

 Σ

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
hoch	3	$\frac{2}{3}$
mittel	2	$\frac{2}{2}$
niedrig	3	$\frac{1}{3}$
ohne	2	$\frac{0}{2}$

 Σ

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
hoch	3	$\frac{2}{3}$
mittel	2	$\frac{2}{2}$
niedrig	3	$\frac{1}{3}$
ohne	2	$\frac{0}{2}$

 Σ

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC
hoch	3	$\frac{2}{3}$	$\frac{1}{3}$
mittel	2	$\frac{2}{2}$	$\frac{0}{2}$
niedrig	3	$\frac{1}{3}$	$\frac{2}{3}$
ohne	2	$\frac{0}{2}$	$\frac{2}{2}$

 Σ

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel
hoch	3	$2/3$	$1/3$	$3/10 \cdot (-2/3 \log_2 2/3 - 1/3 \log_2 1/3)$
mittel	2	$2/2$	$0/2$	$2/10 \cdot (-2/2 \log_2 2/2 - 0/2 \log_2 0/2)$
niedrig	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$
ohne	2	$0/2$	$2/2$	$2/10 \cdot (-0/2 \log_2 0/2 - 2/2 \log_2 2/2)$
				Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel	Ergebnis
hoch	3	$2/3$	$1/3$	$3/10 \cdot (-2/3 \log_2 2/3 - 1/3 \log_2 1/3)$	$= 0,2755$
mittel	2	$2/2$	$0/2$	$2/10 \cdot (-2/2 \log_2 2/2 - 0/2 \log_2 0/2)$	$= 0$
niedrig	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$	$= 0,2755$
ohne	2	$0/2$	$2/2$	$2/10 \cdot (-0/2 \log_2 0/2 - 2/2 \log_2 2/2)$	$= 0$
				Σ	

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel	Ergebnis
hoch	3	$2/3$	$1/3$	$3/10 \cdot (-2/3 \log_2 2/3 - 1/3 \log_2 1/3)$	$= 0,2755$
mittel	2	$2/2$	$0/2$	$2/10 \cdot (-2/2 \log_2 2/2 - 0/2 \log_2 0/2)$	$= 0$
niedrig	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$	$= 0,2755$
ohne	2	$0/2$	$2/2$	$2/10 \cdot (-0/2 \log_2 0/2 - 2/2 \log_2 2/2)$	$= 0$
Σ					0,5510

Informationsgewinn für Fieber

$$G(S, \text{Fieber}) = H(S) - \sum_{i \in \{\text{hoch, mittel, niedrig, ohne}\}} \frac{|S_i|}{|S|} H(S_i) =$$

Beispiel

Informationsgewinn für Fieber

$$\begin{aligned}
 G(S, \text{Fieber}) &= H(S) - \sum_{i \in \{\text{hoch, mittel, niedrig, ohne}\}} \frac{|S_i|}{|S|} H(S_i) = \\
 &= 1 - 0,5510 = 0,4490
 \end{aligned}$$

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten
stark
mittel
leicht
Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl
stark	3
mittel	2
leicht	5
	Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE
stark	3	$\frac{2}{3}$
mittel	2	$\frac{1}{2}$
leicht	5	$\frac{2}{5}$
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE
stark	3	$\frac{2}{3}$
mittel	2	$\frac{1}{2}$
leicht	5	$\frac{2}{5}$
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE	TBC
stark	3	$\frac{2}{3}$	$\frac{1}{3}$
mittel	2	$\frac{1}{2}$	$\frac{1}{2}$
leicht	5	$\frac{2}{5}$	$\frac{4}{5}$
			Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE	TBC	Formel
stark	3	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{3}{10} \cdot (-\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3})$
mittel	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{10} \cdot (-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2})$
leicht	5	$\frac{2}{5}$	$\frac{4}{5}$	$\frac{3}{10} \cdot (-\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5})$
				Σ <input type="text"/>

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE	TBC	Formel	Ergebnis
stark	3	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{3}{10} \cdot (-\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3})$	$= 0,2755$
mittel	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{10} \cdot (-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2})$	$= 0,2$
leicht	5	$\frac{2}{5}$	$\frac{4}{5}$	$\frac{3}{10} \cdot (-\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5})$	$= 0,4855$
				Σ	

Beispiel

Informationsgehalt für Husten

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Husten	Anzahl	LE	TBC	Formel	Ergebnis
stark	3	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{3}{10} \cdot (-\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3})$	$= 0,2755$
mittel	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{10} \cdot (-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2})$	$= 0,2$
leicht	5	$\frac{2}{5}$	$\frac{4}{5}$	$\frac{3}{10} \cdot (-\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5})$	$= 0,4855$
				Σ	$0,9610$

Informationsgewinn für Husten

$$G(S, \text{Husten}) = H(S) - \sum_{i \in \{\text{stark, mittel, leicht}\}} \frac{|S_i|}{|S|} H(S_i) =$$

Beispiel

Informationsgewinn für Husten

$$\begin{aligned}
 G(S, \text{Husten}) &= H(S) - \sum_{i \in \{\text{stark, mittel, leicht}\}} \frac{|S_i|}{|S|} H(S_i) = \\
 &= 1 - 0,9610 = 0,0390
 \end{aligned}$$

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber
blubbernd
fiepend
normal
Σ <input type="text"/>

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl
blubbernd	4
fiepend	3
normal	3

 Σ

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
blubbernd	4	$\frac{4}{4}$
fiepend	3	$\frac{1}{3}$
normal	3	$\frac{0}{3}$
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
blubbernd	4	$\frac{4}{4}$
fiepend	3	$\frac{1}{3}$
normal	3	$\frac{0}{3}$
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC
blubbernd	4	$\frac{4}{4}$	$\frac{0}{4}$
fiepend	3	$\frac{1}{3}$	$\frac{2}{3}$
normal	3	$\frac{0}{3}$	$\frac{3}{3}$
Σ <input type="text"/>			

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel
blubbernd	4	$4/4$	$0/4$	$4/10 \cdot (-4/4 \log_2 4/4 - 0/4 \log_2 0/4)$
fiepend	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$
normal	3	$0/3$	$3/3$	$3/10 \cdot (-0/3 \log_2 0/3 - 3/3 \log_2 3/3)$
				Σ <input style="width: 100px; height: 20px;" type="text"/>

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel	Ergebnis
blubbernd	4	$4/4$	$0/4$	$4/10 \cdot (-4/4 \log_2 4/4 - 0/4 \log_2 0/4)$	$= 0$
fiepend	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$	$= 0,2755$
normal	3	$0/3$	$3/3$	$3/10 \cdot (-0/3 \log_2 0/3 - 3/3 \log_2 3/3)$	$= 0$
				Σ	

Beispiel

Informationsgehalt für Abhören

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel	Ergebnis
blubbernd	4	$4/4$	$0/4$	$4/10 \cdot (-4/4 \log_2 4/4 - 0/4 \log_2 0/4)$	$= 0$
fiepend	3	$1/3$	$2/3$	$3/10 \cdot (-1/3 \log_2 1/3 - 2/3 \log_2 2/3)$	$= 0,2755$
normal	3	$0/3$	$3/3$	$3/10 \cdot (-0/3 \log_2 0/3 - 3/3 \log_2 3/3)$	$= 0$
				Σ	$0,2755$

Informationsgewinn für Abhören

$$G(S, \text{Abhören}) = H(S) - \sum_{i \in \{\text{blubbernd}, \text{fiepend}, \text{normal}\}} \frac{|S_i|}{|S|} H(S_i) =$$

Beispiel

Informationsgewinn für Abhören

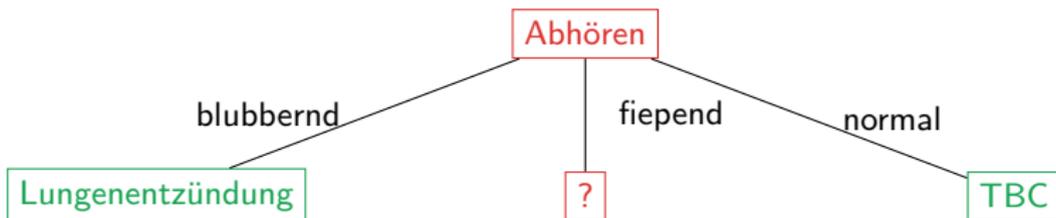
$$\begin{aligned}
 G(S, \text{Abhören}) &= H(S) - \sum_{i \in \{\text{blubbernd}, \text{fiepend}, \text{normal}\}} \frac{|S_i|}{|S|} H(S_i) = \\
 &= 1 - 0,2755 = 0,7245
 \end{aligned}$$

Wahl des ersten Knoten

Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
0,4490	0,0390	0,4390	0,3958	0,7245	1

Wahl des ersten Knoten

Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
0,4490	0,0390	0,4390	0,3958	0,7245	1



Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber
niedrig
hoch
ohne
Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl
niedrig	1
hoch	1
ohne	1
	Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
niedrig	1	1/1
hoch	1	0/1
ohne	1	0/1
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE
niedrig	1	1/1
hoch	1	0/1
ohne	1	0/1
		Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC
niedrig	1	1/1	0/1
hoch	1	0/1	1/1
ohne	1	0/1	1/1
Σ <input type="text"/>			

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel
niedrig	1	$1/1$	$0/1$	$1/3 \cdot (-1/1 \log_2 1/1 - 0/1 \log_2 0/1)$
hoch	1	$0/1$	$1/1$	$1/3 \cdot (-0/1 \log_2 0/1 - 1/1 \log_2 1/1)$
ohne	1	$0/1$	$1/1$	$1/3 \cdot (-0/1 \log_2 0/1 - 1/11 \log_2 1/1)$
				Σ <input type="text"/>

Beispiel

Informationsgehalt für Fieber

Patient	Fieber	Husten	Röntgen	BSG	Abhören	Krankheit
1	hoch	stark	Flocken	normal	blubbernd	Lungenentzündung
2	mittel	stark	Flocken	normal	blubbernd	Lungenentzündung
3	niedrig	leicht	Punkt	normal	fiepend	Lungenentzündung
4	hoch	mittel	Flocken	normal	blubbernd	Lungenentzündung
5	mittel	leicht	Flocken	normal	blubbernd	Lungenentzündung
6	ohne	leicht	Streifen	normal	normal	Tuberkulose
7	hoch	stark	Loch	schnell	fiepend	Tuberkulose
8	niedrig	leicht	Streifen	normal	normal	Tuberkulose
9	ohne	leicht	Punkt	schnell	fiepend	Tuberkulose
10	niedrig	mittel	Flocken	schnell	normal	Tuberkulose

Fieber	Anzahl	LE	TBC	Formel	Ergebnis
niedrig	1	$1/1$	$0/1$	$1/3 \cdot (-1/1 \log_2 1/1 - 0/1 \log_2 0/1)$	$= 0$
hoch	1	$0/1$	$1/1$	$1/3 \cdot (-0/1 \log_2 0/1 - 1/1 \log_2 1/1)$	$= 0$
ohne	1	$0/1$	$1/1$	$1/3 \cdot (-0/1 \log_2 0/1 - 1/11 \log_2 1/1)$	$= 0$
				Σ	0

Informationsgewinn für Fieber

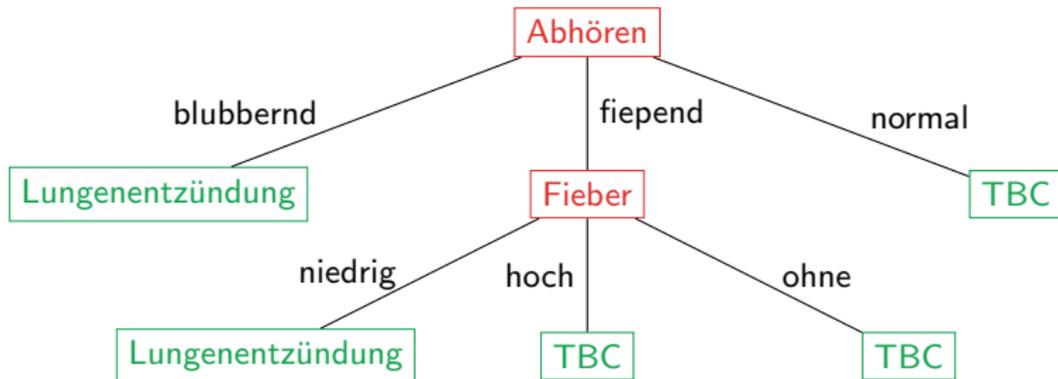
$$G(S, \text{Fieber}) = H(S) - \sum_{i \in \{\text{niedrig, hoch, normal}\}} \frac{|S_i|}{|S|} H(S_i) =$$

Informationsgewinn für Fieber

$$\begin{aligned}
 G(S, \text{Fieber}) &= H(S) - \sum_{i \in \{\text{niedrig, hoch, normal}\}} \frac{|S_i|}{|S|} H(S_i) = \\
 &= 0,9183 - 0 = 0,9183
 \end{aligned}$$

Beispiel

Resultierender Entscheidungsbaum



Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Allgemeines

- ▶ Daten aus meiner Bachelorarbeit

Allgemeines

- ▶ Daten aus meiner Bachelorarbeit
- ▶ Nicht vollständig

Allgemeines

- ▶ Daten aus meiner Bachelorarbeit
- ▶ Nicht vollständig
- ▶ Uniformisierte Daten (vier Klassen à 876 Samples)

Allgemeines

- ▶ Daten aus meiner Bachelorarbeit
- ▶ Nicht vollständig
- ▶ Uniformisierte Daten (vier Klassen à 876 Samples)
- ▶ GPS Geschwindigkeit und Beschleunigungslevel

Allgemeines

- ▶ Daten aus meiner Bachelorarbeit
- ▶ Nicht vollständig
- ▶ Uniformisierte Daten (vier Klassen à 876 Samples)
- ▶ GPS Geschwindigkeit und Beschleunigungslevel

Allgemeines

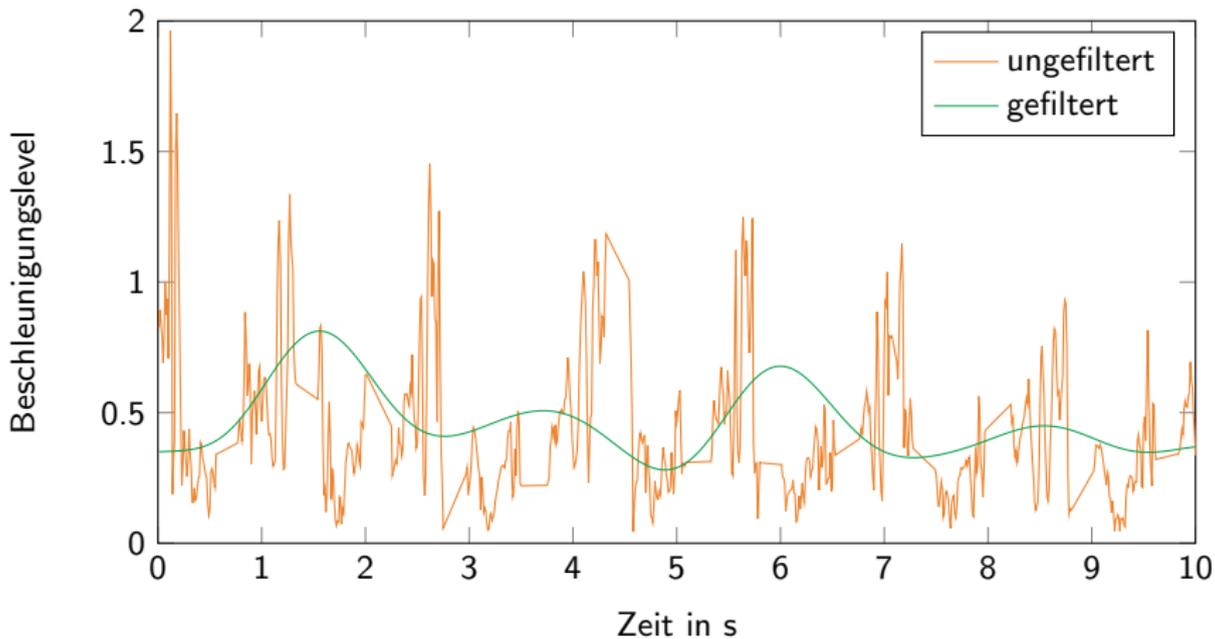
- ▶ Daten aus meiner Bachelorarbeit
- ▶ Nicht vollständig
- ▶ Uniformisierte Daten (vier Klassen à 876 Samples)
- ▶ GPS Geschwindigkeit und Beschleunigungslevel

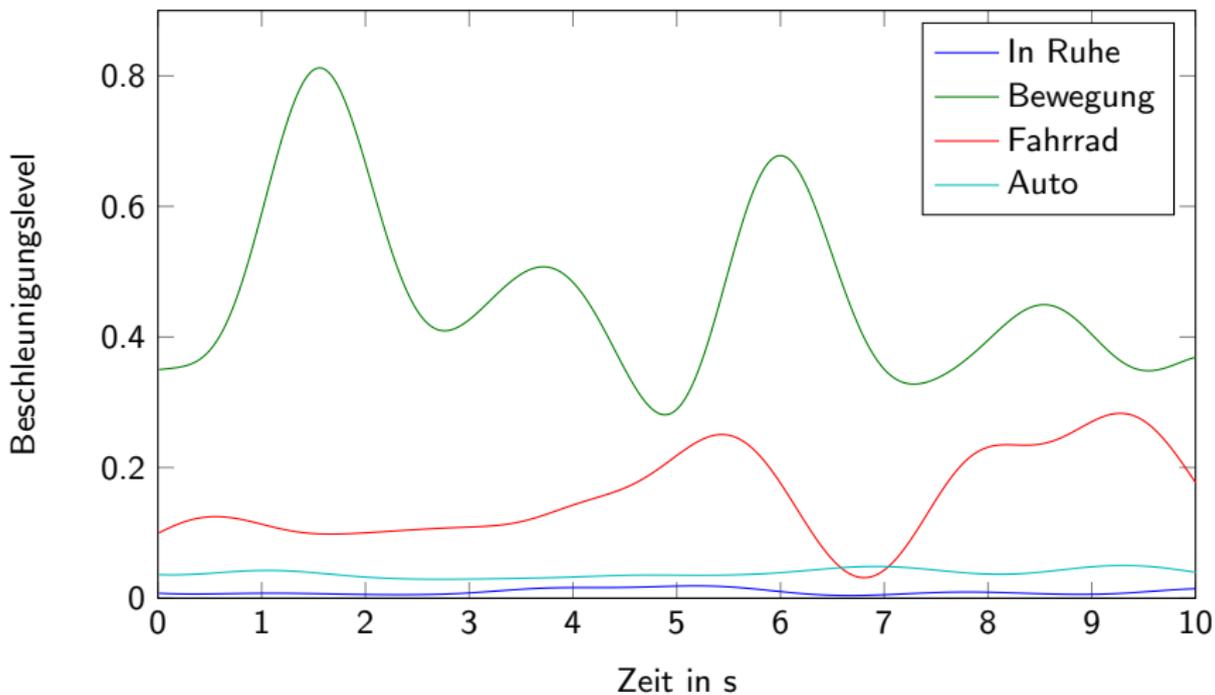
Berechnung

Betrag der Beschleunigung

$$m_i = \sqrt{x_i^2 + y_i^2 + z_i^2}, \quad i = 1, \dots, n \quad (1)$$

mit x_i, y_i, z_i als Beschleunigungswerte entlang der entsprechenden Achse für jeden Datenpunkt i der gesamtem Samples n .





Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

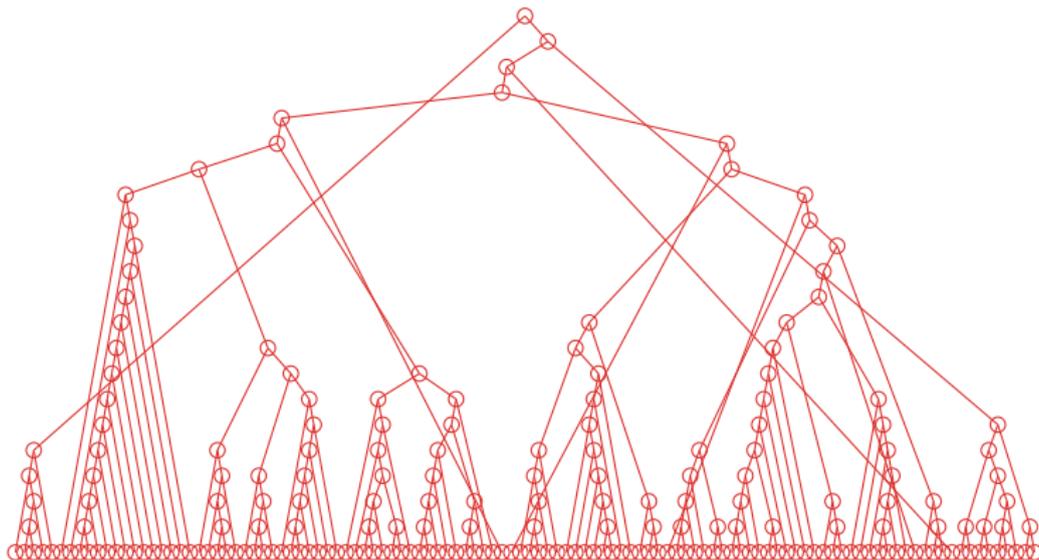
Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Entscheidungsbaum für Bewegungsmuster



Tiefe: 21

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung Klassifikation einer Eiche anhand ihrer Silhouette

Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung
Klassifikation einer Eiche anhand ihrer Silhouette





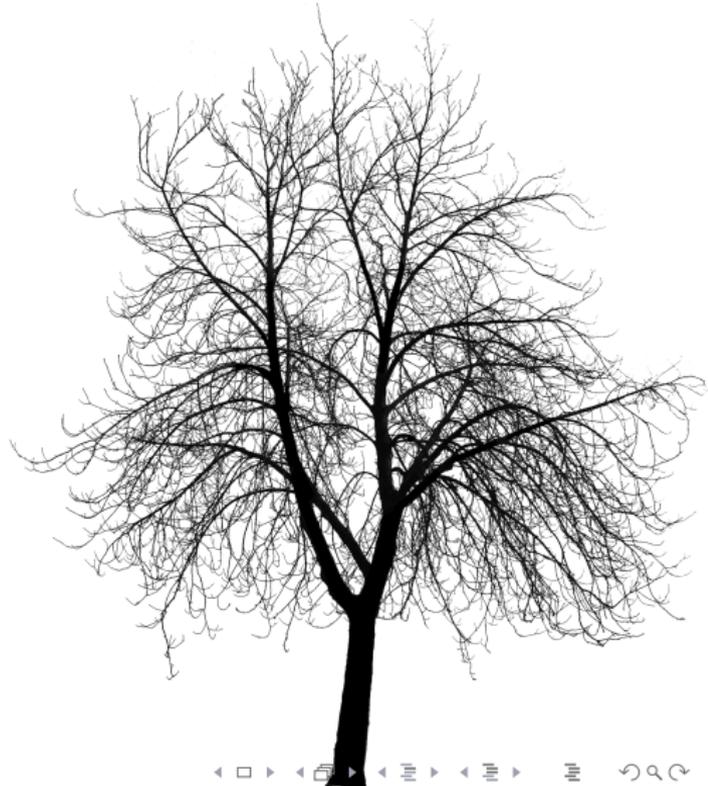
Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung
Klassifikation einer Eiche anhand ihrer Silhouette



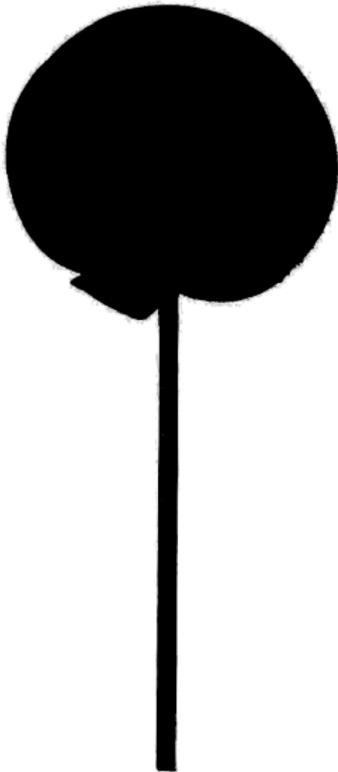
Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung
Klassifikation einer Eiche anhand ihrer Silhouette



Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung
Klassifikation einer Eiche anhand ihrer Silhouette



Under- & Overfitting (1)

Was ist Overfitting? – Eine anschauliche Erklärung
Klassifikation einer Eiche anhand ihrer Silhouette

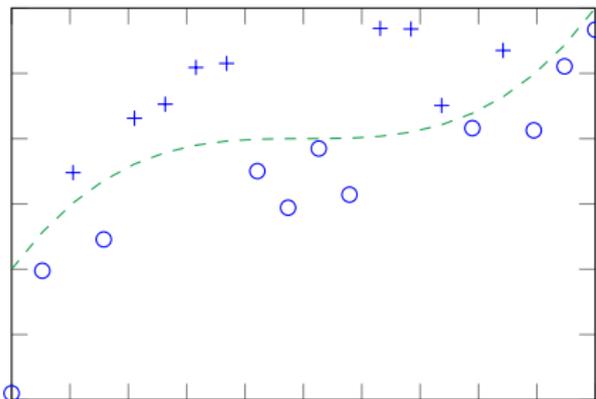


Under- & Overfitting (2)

Was ist Overfitting? – Eine mathematische Erklärung

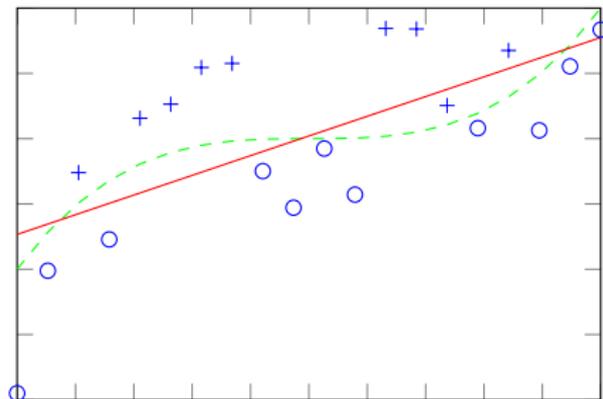
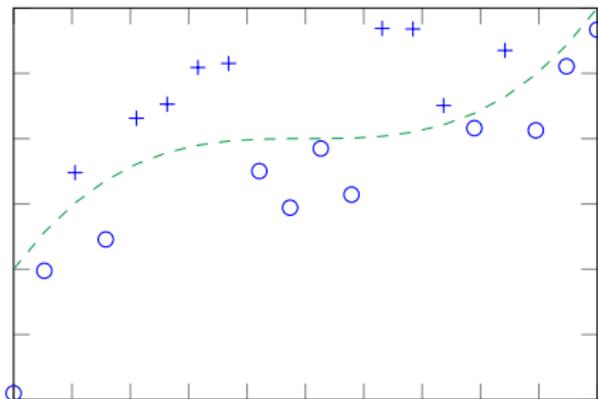
Under- & Overfitting (2)

Was ist Overfitting? – Eine mathematische Erklärung



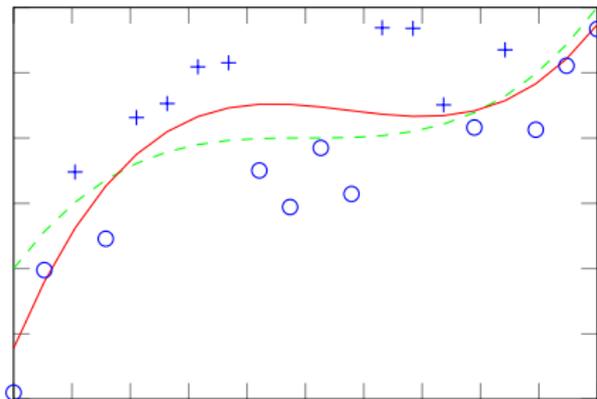
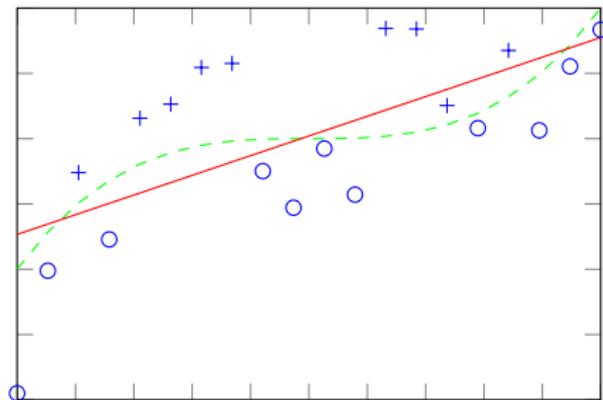
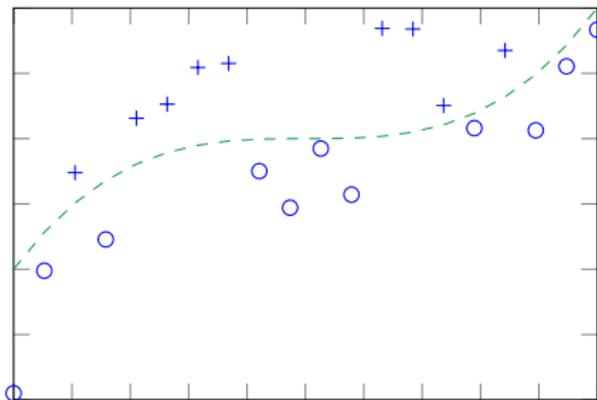
Under- & Overfitting (2)

Was ist Overfitting? – Eine mathematische Erklärung



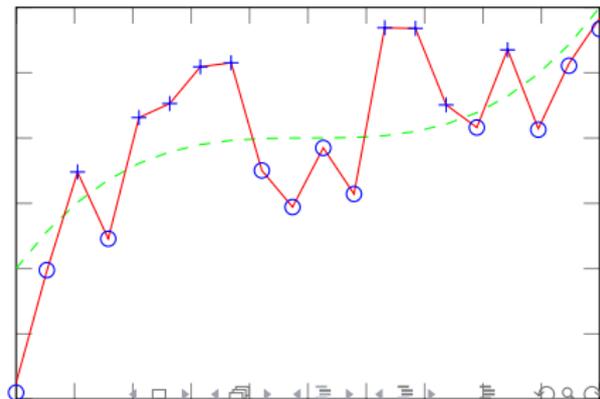
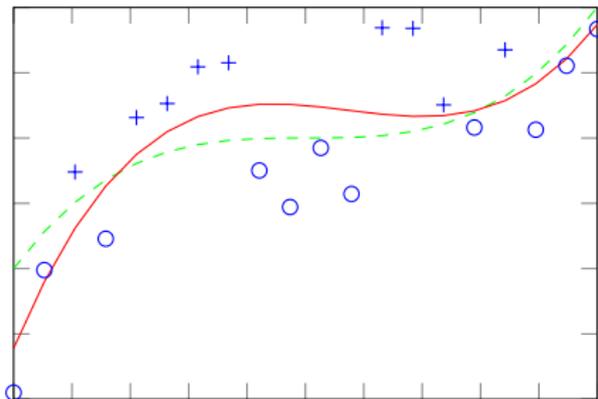
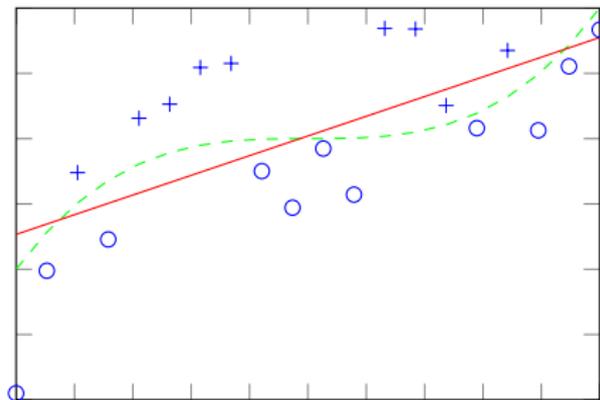
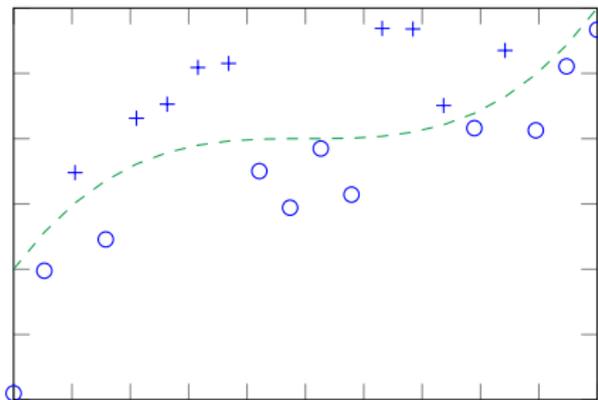
Under- & Overfitting (2)

Was ist Overfitting? – Eine mathematische Erklärung



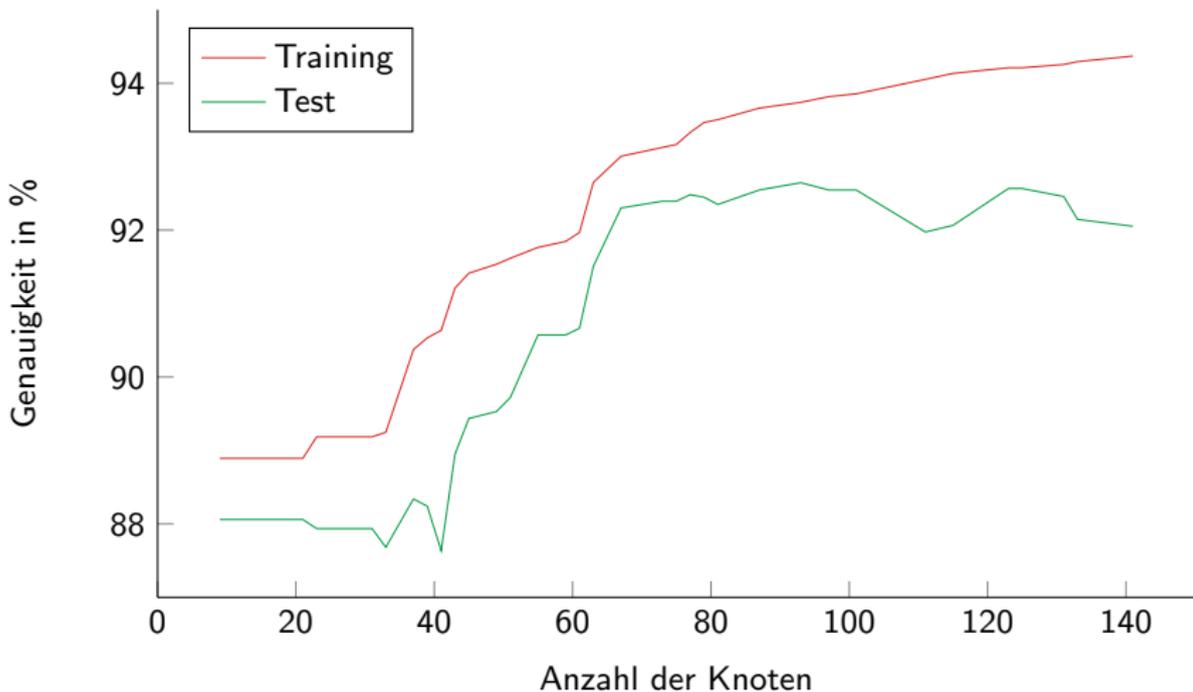
Under- & Overfitting (2)

Was ist Overfitting? – Eine mathematische Erklärung



Under- & Overfitting (3)

Was ist Overfitting? – Eine anwendungsorientierte Erklärung



Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

- ▶ Pruning mit Schwellwert

- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning

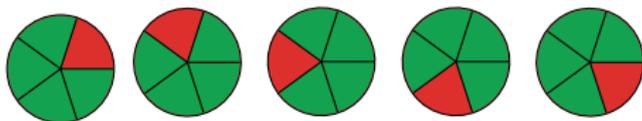
- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning
- ▶ Cross Validation

- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning
- ▶ Cross Validation

Optimierungen

- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning
- ▶ Cross Validation

Trainingsdaten und Testdaten:

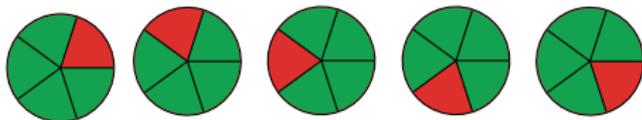


- ▶ Bagging (Bootstrap aggregating)

Optimierungen

- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning
- ▶ Cross Validation

Trainingsdaten und Testdaten:

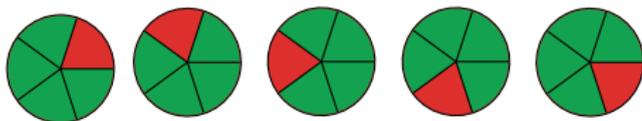


- ▶ Bagging (Bootstrap aggregating)
- ▶ Induktiver Bias

Optimierungen

- ▶ Pruning mit Schwellwert
- ▶ Reduced-Error-Pruning
- ▶ Cross Validation

Trainingsdaten und Testdaten:



- ▶ Bagging (Bootstrap aggregating)
- ▶ Induktiver Bias
- ▶ Anderes Kriterium für Bestenauswahl (Gini-Index)

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

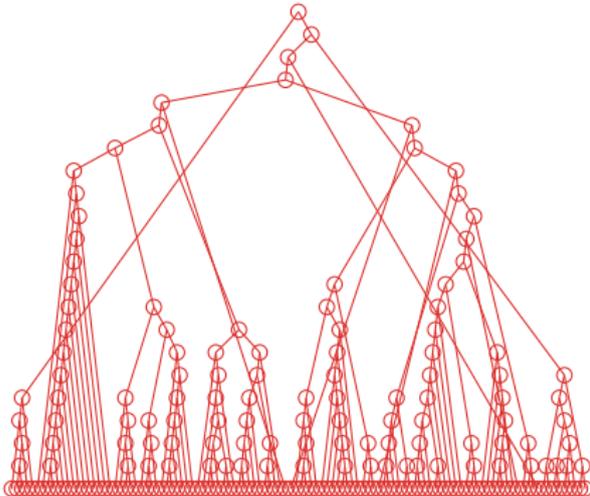
Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

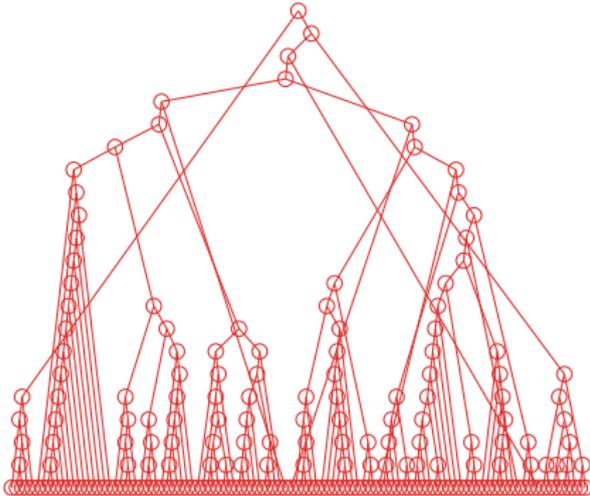
Zusammenfassung

Kompletter Entscheidungsbaum



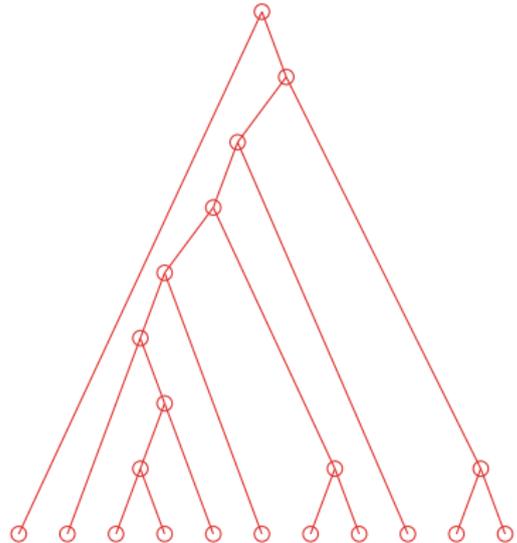
Tiefe: 21
Genauigkeit: 91,1 %

Kompletter Entscheidungsbaum



Tiefe: 21
Genauigkeit: 91,1 %

Bester Entscheidungsbaum



Tiefe: 8
Genauigkeit: 92,3 %

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

- ▶ Medizin

- ▶ Medizin
- ▶ Klassifikation

- ▶ Medizin
- ▶ Klassifikation
- ▶ Regression

- ▶ Medizin
- ▶ Klassifikation
- ▶ Regression
- ▶ Data-Mining

- ▶ Medizin
- ▶ Klassifikation
- ▶ Regression
- ▶ Data-Mining
- ▶ SPAM-Filter

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

▶ CHAID

- ▶ CHAID
- ▶ CART

- ▶ CHAID
- ▶ CART
- ▶ C4.5 und C5.0

- ▶ CHAID
- ▶ CART
- ▶ C4.5 und C5.0
- ▶ Entscheidungswälder

- ▶ CHAID
- ▶ CART
- ▶ C4.5 und C5.0
- ▶ Entscheidungswälder
- ▶ Kombination mit künstlichen neuronalen Netzen

Entscheidungsbäume

Definition

Beispielentscheidungsbaum

Konstruktion

Beispielkonstruktion

ID3 Algorithmus

Exkurs in die Physik

Informationsgehalt und -gewinn

Beispiel

Pseudocode

Beispiel

Anwendung – Bewegungsdaten

Einführung

Entscheidungsbaum

Under- & Overfitting

Optimierungen

Ergebnisse & Vergleiche

Ausblick & Zusammenfassung

Anwendungsgebiete

Ausblick

Zusammenfassung

- ▶ Was ist ein Entscheidungsbaum?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?
- ▶ Worin liegen die Stärken und Schwächen eines Entscheidungsbaumes?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?
- ▶ Worin liegen die Stärken und Schwächen eines Entscheidungsbaumes?
- ▶ Was ist Overfitting?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?
- ▶ Worin liegen die Stärken und Schwächen eines Entscheidungsbaumes?
- ▶ Was ist Overfitting?
- ▶ Wo gibt es Optimierungsmöglichkeiten?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?
- ▶ Worin liegen die Stärken und Schwächen eines Entscheidungsbaumes?
- ▶ Was ist Overfitting?
- ▶ Wo gibt es Optimierungsmöglichkeiten?
- ▶ Was sind die Einsatzgebiete von Entscheidungsbäumen?

- ▶ Was ist ein Entscheidungsbaum?
- ▶ Wie konstruiere ich einen Entscheidungsbaum?
- ▶ Warum haben Atomkraftwerke Kühltürme oder was ist Entropie?
- ▶ Was ist Informationsgehalt und -gewinn?
- ▶ Wie funktioniert der ID3-Algorithmus?
- ▶ Worin liegen die Stärken und Schwächen eines Entscheidungsbaumes?
- ▶ Was ist Overfitting?
- ▶ Wo gibt es Optimierungsmöglichkeiten?
- ▶ Was sind die Einsatzgebiete von Entscheidungsbäumen?
- ▶ Welche weiteren Algorithmen für Entscheidungsbäume gibt es neben des ID3-Algorithmus' noch?

Vielen Dank für die Aufmerksamkeit!

Fragen?