

January	February	March	April
M T W T F S S	M T W T F S S	M T W T F S S	M T W T F S S
1 30 31 1 2 3 4 5	5 27 28 29 30 31 1 2	9 24 25 26 27 28 1 2	14 31 1 2 3 4 5 6
2 6 7 8 9 10 11 12	6 3 4 5 6 7 8 9	10 3 4 5 6 7 8 9	15 7 8 9 10 11 12 13
3 13 14 15 16 17 18 19	7 10 11 12 13 14 15 16	11 10 11 12 13 14 15 16	16 14 15 16 17 18 19 20
4 20 21 22 23 24 25 26	8 17 18 19 20 21 22 23	12 17 18 19 20 21 22 23	17 21 22 23 24 25 26 27
5 27 28 29 30 31 1 2	9 24 25 26 27 28 1 2	13 24 25 26 27 28 29 30	18 28 29 30 1 2 3 4
6 3 4 5 6 7 8 9	10 3 4 5 6 7 8 9	14 31 1 2 3 4 5 6	19 5 6 7 8 9 10 11
May	June	July	August
M T W T F S S	M T W T F S S	M T W T F S S	M T W T F S S
18 28 29 30 1 2 3 4	25 26 27 28 29 30 1 2	27 30 1 2 3 4 5 6	31 28 29 30 31 1 2 3
19 5 6 7 8 9 10 11	23 2 3 4 5 6 7 8	28 7 8 9 10 11 12 13	32 4 5 6 7 8 9 10
20 12 13 14 15 16 17 18	24 9 10 11 12 13 14 15	29 14 15 16 17 18 19 20	33 11 12 13 14 15 16 17
21 19 20 21 22 23 24 25	25 16 17 18 19 20 21 22	30 21 22 23 24 25 26 27	34 18 19 20 21 22 23 24
22 26 27 28 29 30 31	26 23 24 25 26 27 28 29	31 28 29 30 31 1 2 3	35 25 26 27 28 29 30 31
23 2 3 4 5 6 7 8	27 30 1 2 3 4 5 6	32 4 5 6 7 8 9 10	36 1 2 3 4 5 6 7

Handwritten notes on the calendar:

- Übungen* (Exercises) written over June 25-30.
- Theorie* (Theory) written over May 29-30.
- April 24 is circled in red.
- May 6 is circled in blue.
- June 9-15, 19-20, and July 3, 25 are circled in purple.

KW 19 [0311][0312][0318]

KW 20 [0319][0325][0331]

KW 21 [0401][0408][0415]

KW 22 [0416][0422][0423]

KW 23 [0429][0430][0506]

KW 24 [0507][0513][0514][0520][0521]

KW 25 PROBEKLAUSUR.

KW 26 Q&A

Übungen

Übungen

Schlachtplan
Prüfung.

ENTSCHEIDUNGSBAUME (CART).

KONZEPT: Verunreinigung der Information.
Die Verunreinigung misst die Homogenität einer Datenprobe. Wenn die Daten in der Probe homogen sind, gehören die Stichproben zur gleichen Klasse und die Verunreinigung ist NULL (\emptyset).

Wir messen die Verunreinigung mit dem Gini-Index

$$\text{Gini Index} = 1 - \sum_{i=1}^n p_i^2$$

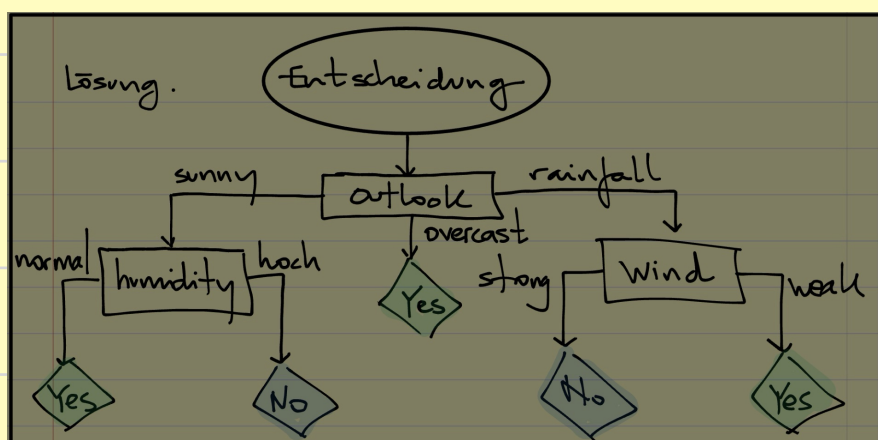
$p_i \in [0,1] \equiv \text{W. dafür, dass die Probe zur Klasse gehört.}$

Gini-Index ist ein Maß für die Verunreinigung einer Stichprobe und hat einen Wert zw. 0 und 1.

- Gini-Index = 0 bedeutet, dass die Stichprobe vollkommen homogen ist, und Alle Elemente zur Klasse gehören.
- Gini-Index = 1 bedeutet, dass die Stichprobe maximal verunreinigt ist bzw. maximale Ungleichheit zw. den Elementen.

Beispiel:

Wandern



Lösung

Day	outlook	temperature	humidity	wind	Decision
1	sunny	hot	high	weak	No
2	sunny	hot	high	strong	No
3	overcast	hot	high	weak	Yes
4	rainfall	mild	high	weak	Yes
5	rainfall	cool	normal	weak	Yes
6	rainfall	cool	normal	strong	No
7	overcast	cool	normal	wtrong	Yes
8	sunny	mild	high	weak	No
9	sunny	cool	normal	weak	Yes
10	rainfall	mild	normal	weak	Yes
11	sunny	mild	normal	strong	Yes
12	overcast	mild	high	strong	Yes
13	overcast	hot	normal	weak	Yes
14	rainfall	mild	high	strong	No

Day ♂. Harshalt

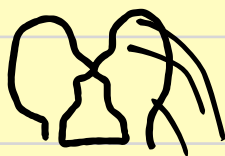
Ja

Nein



da

Nein



of +

Selten

Gar Nicht

• • •



Ja

Nein

