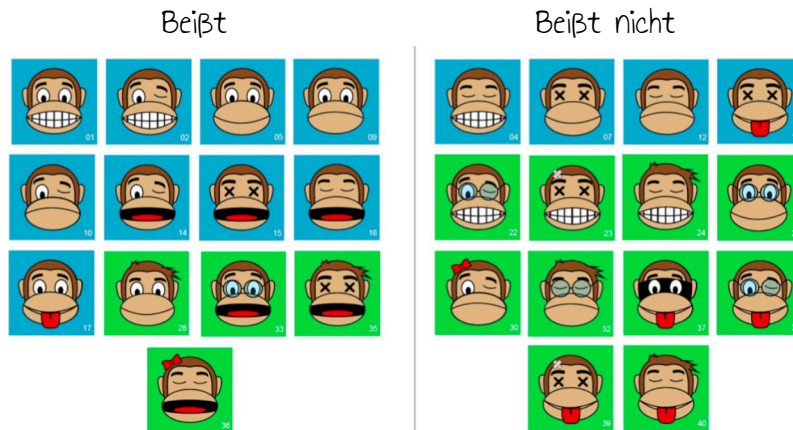


Aufgabenblatt: Klassifikation mit Entscheidungsbäumen

Als Tierpflegerin bzw. Tierpfleger sind wir für die Fütterung der Affen zuständig. Dabei müssen wir aufpassen: Manche Affen beißen. Von den Affen im Zoo wissen wir bereits, ob sie beißen. Allerdings werden bald neue Tiere zur Gruppe hinzukommen. Wir müssen uns daher überlegen, wie wir herausfinden können, welche neuen Affen beißen und welche nicht – am besten ohne ihren Zähnen zu nahe zu kommen.

Trainingsdaten

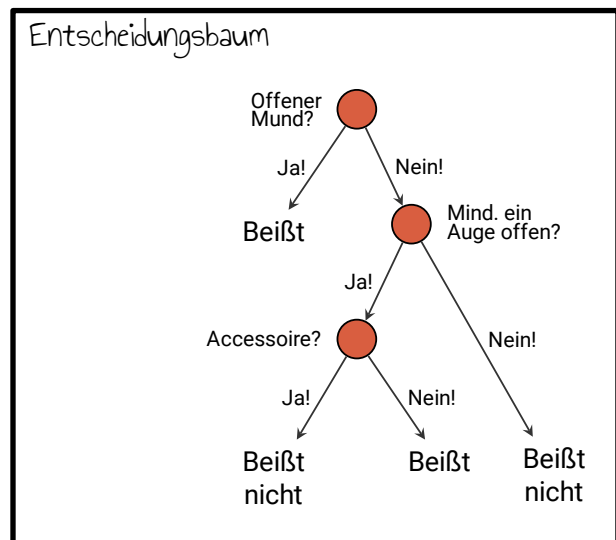
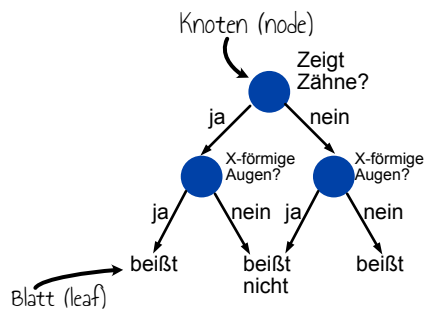
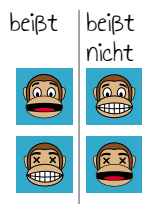
Von den Affen, die bereits im Zoo leben, können wir lernen. Sie sind unsere **Trainingsdaten**, mit deren Hilfe der **Entscheidungsbaum** erstellt wird.



Entscheidungsbaum

Ein Entscheidungsbaum stellt Regeln hierarchisch dar und wird von oben nach unten gelesen.

Vollziehe das Beispiel nach und erstelle einen Entscheidungsbaum für die gegebenen Affen.



Testdaten

Ein Entscheidungsbaum ist ein Modell, das nun auf weitere Daten angewendet werden kann. Unser Modell können wir nun mit weiteren Tieren testen, von denen wir bereits wissen, ob sie beißen oder nicht. Diese Tiere sind also unsere **Testdaten**.



1. Entscheide mithilfe deines Entscheidungsbaumes, ob die folgenden Affen beißen (B) oder nicht beißen (NB) und trage deine Vorhersage in die freien Felder ein!

2. Vergleiche deine Lösung mit der Auflösung und markiere, ob deine Vorhersage richtig oder falsch war!

						Deine Vorhersage: Beißt oder beißt nicht?
BN	B	BN	BN	B	B	
✓ X	✓ X	✓ X	✓ X	✓ X	✓ X	War deine Vorhersage richtig oder falsch?
						Deine Vorhersage: Beißt oder beißt nicht?
BN	BN	BN	BN	BN	B	
✓ X	✓ X	✓ X	✓ X	✓ X	✓ X	War deine Vorhersage richtig oder falsch?

Konfusionsmatrix

Jetzt hast du die Testdaten klassifiziert und ihre tatsächliche Zuordnung ist bekannt. Die Ergebnisse trägst du summarisch in die folgende Tabelle (die Konfusionsmatrix) ein.

1. Fülle die Felder der Konfusionsmatrix!

Beispiel: Das Feld links oben gibt an, für wie viele Äffchen du richtig die Kategorie beißt vorhergesagt hast (z.B. 2)!

	Vorhergesagt: Äffchen beißt	Vorhergesagt: Äffchen beißt nicht
Tatsächlich: Äffchen beißt	3	1
Tatsächlich: Äffchen beißt nicht	2	6

2. Berechne die folgenden Verhältnisse und gib das Ergebnis als Prozentsatz an!

$$\begin{array}{|c|c|c|} \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \end{array}
 \quad
 \frac{\blacksquare}{\blacksquare}
 = \frac{\# \text{korrekt vorhergesagt}}{\# \text{insgesamt vorhergesagt}}
 = \boxed{75\%}
 \quad
 \underline{\text{Genauigkeit (accuracy)}}$$

$$\begin{array}{|c|c|c|} \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \end{array}
 \quad
 \frac{\blacksquare}{\blacksquare}
 = \frac{\# \text{beißt richtig vorhergesagt}}{\# \text{beißt tatsächlich}}
 = \boxed{75\%}
 \quad
 \underline{\text{Trefferquote (recall)}}$$

$$\begin{array}{|c|c|c|} \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \blacksquare & \blacksquare & \blacksquare \\ \hline \end{array}
 \quad
 \frac{\blacksquare}{\blacksquare}
 = \frac{\# \text{beißt richtig vorhergesagt}}{\# \text{beißt vorhergesagt}}
 = \boxed{60\%}
 \quad
 \underline{\text{Präzision (precision)}}$$

Deine Ergebnisse

Hier benennen wir
gemeinsam diese Metrik.

