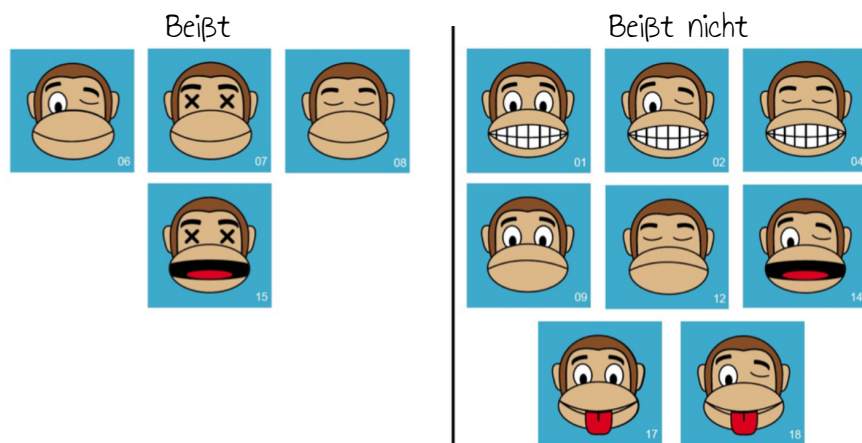


Aufgabenblatt: Klassifikation mit Entscheidungsbäumen

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

Trainingsdaten

Von den Äffchen, die bereits im Zoo leben, können wir lernen. Sie sind unsere **Trainingsdaten**. Mit Hilfe dieser Trainingsdaten können wir Regeln entwickeln, um beißende und nicht beißende Äffchen unterscheiden zu können. Regeln können bspw. von folgender Form sein: Wenn der Affe die Zunge zeigt, dann beißt er. Entwickle basierend auf den folgenden Äffchen selbst solche Regeln!



Regeln

Individuell

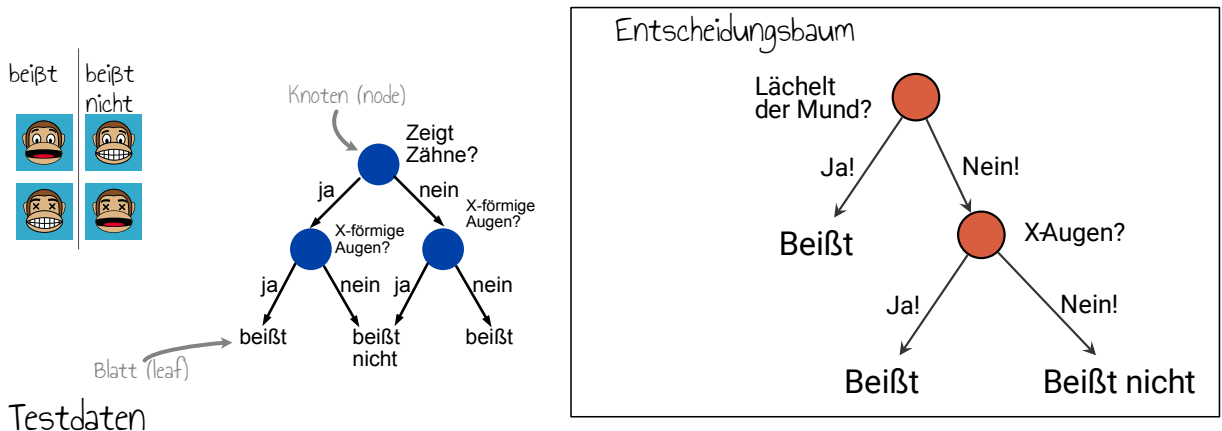
Z.B. „Hat ein Äffchen einen lächelnden Mund, so beißt es.“

Lächelt er nicht und hat X-Augen, so beißt es.

In allen anderen Fällen beißt es nicht.

Entscheidungsbaum

Eine besondere Art solche Regeln darzustellen sind **Entscheidungsbäume**. Ein Entscheidungsbaum stellt Regeln hierarchisch dar und wird von oben nach unten gelesen. Vollziehe das folgende Beispiel nach und notiere deine Regeln von Seite 1 ebenfalls als Entscheidungsbaum im Feld rechts.



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 Äffchen beißen (B) oder nicht beißen (NB) und trage deine Vorhersage in die freien Felder ein!
2. Vergleiche deine Lösung mit der Auflösung, die du von deiner Lehrkraft bekommst, und markiere, ob deine Vorhersage richtig oder falsch war!

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 | 2 | 1 |
| Tatsächlich: Äffchen beißt nicht | 1 | 2 |

2. Berechne das folgende Verhältnis und gib das Ergebnis als Prozentsatz an!