

Entscheidungsbäume

Kompaktes Lern- und Nachschlageheft · Informatik 11

1. Entscheidungsbäume

Ein **Entscheidungsbaum** ist ein Modell zur **Klassifikation**: Er ordnet neue Daten anhand ihrer Merkmale einer Klasse zu. Entscheidungsbäume können durch **überwachtes Lernen** aus Beispielen mit bekanntem Ergebnis erzeugt werden. Sie werden immer von unten nach oben durchlaufen.

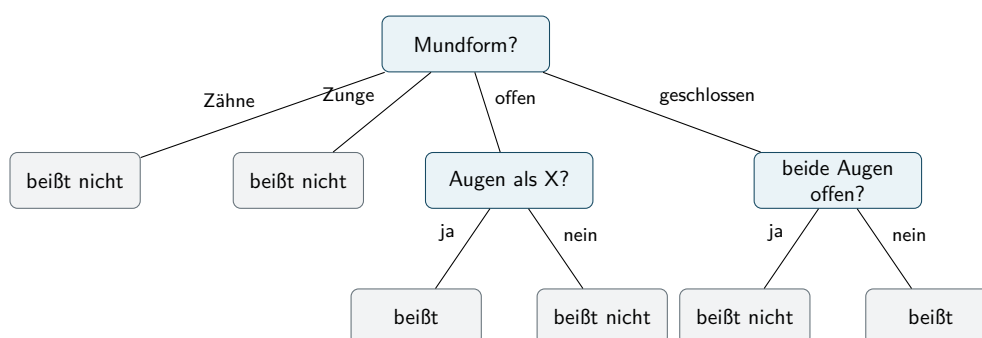
Definition: Überwachtes Lernen

Beim **überwachten Lernen** wird aus **gelabelten Daten** ein Modell erzeugt. Gelabelte Daten besitzen neben ihren Merkmalen ein bekanntes **Label**, also die richtige Klasse. Das Lernverfahren sucht Regeln, mit denen es anschließend neue Eingaben klassifizieren kann.



Abbildung: Stefan Seegerer, Tilman Michaeli und Sven Jatzlau, CC BY; übernommen aus dem Ausgangsmaterial.

2. Aufbau eines Entscheidungsbaums



- Ein **Entscheidungsknoten** prüft ein Attribut.
- Eine **Kante** steht für eine Merkmalsausprägung und führt zum nächsten Knoten.
- Ein **Blatt** enthält die vorhergesagte Klasse bzw. das Label.

Beispiel: eine Vorhersage

Ein Äffchen mit *Mundform* = *offen* und *Augenstellung* = *X* folgt im obigen Baum zweimal der passenden Kante. Es erreicht das Blatt *beißt*; dies ist die Vorhersage des Modells.