

# AI Replugged

Das Gute-Äffchen-Böse-Äffchen-Spiel mit *Orange 3*

von Lennard Kerber

**Themen zum maschinellen Lernen und zu Big Data eignen sich zu einer gemeinsamen Behandlung in einer Einheit: Bei der Big-Data-Analyse spielen KI-Algorithmen eine zentrale Rolle. In der im Folgenden vorgestellten Unterrichtsreihe waren beide Teile unterrichtlich gleichgewichtig; in diesem Beitrag liegt der Schwerpunkt auf der Einführung des überwachten Lernens mit Entscheidungsbäumen. Dafür wird die bekannte AI-Unplugged-Aktivität *Gute-Äffchen-Böse-Äffchen* unter Nutzung des Werkzeugs *Orange 3* wieder in die digitale Form überführt, mit dem Ziel die Abläufe, Möglichkeiten und Grenzen von Datenanalysen an diesem Beispiel zu verdeutlichen. Es entstehen dabei drei verschiedene Entscheidungsbäume zur Lösung des Problems: eine manuell erzeugte unplugged-Variante und, vom *Orange*-Werkzeug erzeugt, ein beliebig verzweigender Entscheidungsbaum und ein binärer Entscheidungsbaum. Auf diesen exemplarischen Grundlagen werden typische Aussagen über maschinelles Lernen plausibilisiert.**

*Modelle zur Prognose (Entscheidungsbaum)* einen Vorschlag von Annabel Lindner und Stefan Seegerer aus der Reihe *AI Unplugged* (vgl. Lindner/Seegerer, 2020a und 2020b), dem Gute-Äffchen-Böse-Äffchen-Spiel, und einen Vorschlag von Andreas Grillenberger zum Thema *Big Data* (siehe den vorhergehenden Beitrag *Datenflussorientierte Big-Data-Analyse*, Seite 61 ff.) kombiniert.

Der Unterricht fand im Rahmen des Enrichment-Unterrichts Mathematik statt. Deshalb stand die Technik der Vorhersage mittels überwachtem Lernen mit Entscheidungsbäumen im Fokus. Ein Überblick über die Reihe findet sich in Bild 1. Im Rahmen der Reihe wurden folgende Schwerpunkte vorgesehen:

Die Schülerinnen und Schüler ...

- ▷ beschreiben überwachtes Lernen als ein Verfahren mit Trainings- und Testdaten am Beispiel von Entscheidungsbäumen.
- ▷ wenden automatisch generierte Entscheidungsbäume für die Analyse von gegebenen Problemstellungen an.
- ▷ identifizieren mögliche Schwachstellen des eingesetzten Verfahrens.

## Und es ist *fast* alles nur geklaut

Um Unterricht schnell und zügig vorzubereiten, stützt man sich gerne auf bekannte und erprobte Vorschläge. Deshalb habe ich bei der Reihe *Mathematische*

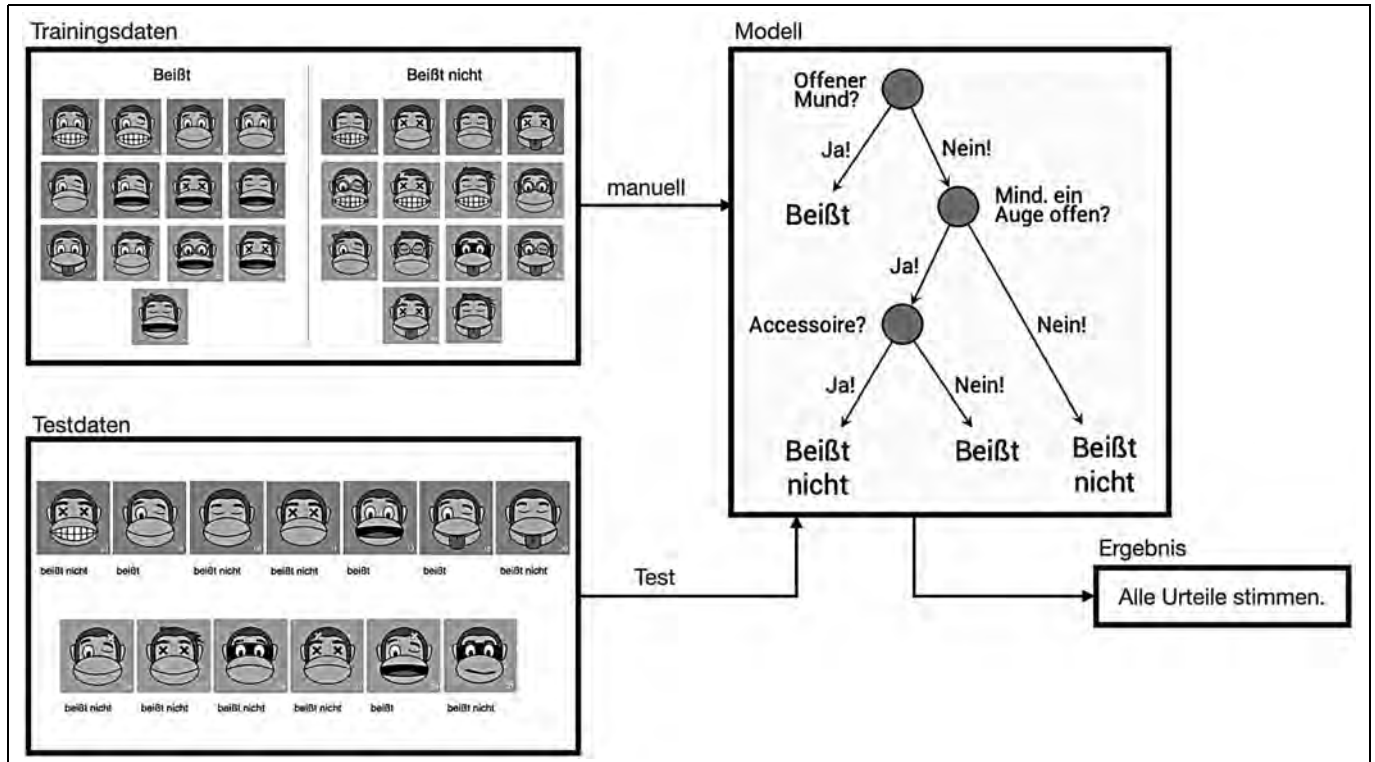
Der vorliegende Beitrag konzentriert sich auf den Erkenntnisprozess im Bezug auf den letztgenannten Schwerpunkt.

Beim Gute-Äffchen-Böse-Äffchen-Spiel werden zwei Kategorien von Äffchen ikonisch dargestellt, die *beißen* und die *nicht beißenden*. Ziel ist, basierend auf einem gelernten Modell aus einem Trainingsdatensatz von Äffchenbildern, weitere Äffchen zu klassifizieren. Entsprechend sollen die Schülerinnen und Schüler selbst Entscheidungsbäume erstellen, die diese Trainingsdaten korrekt klassifizieren. Zur Beurteilung ihres Entscheidungsbaums werden Testdaten genutzt. Die Trainings- und Testdaten sowie eine mögliche Lösung und ihre Güte werden in Bild 2 (nächste Seite) gezeigt.

1. Begriffsbildung: Baum / Entscheidungsbaum (Knoten, Blatt, Wurzel, Verzweigungsgrad, binärer Baum)	} 90 min
2. AI Unplugged: Gute-Äffchen-Böse-Äffchen-Spiel (Spiel 1 und 2)	} 90 min
3. AI Replugged: Gute-Äffchen-Böse-Äffchen-Spiel (Spiel 2) mit <i>Orange 3</i>	
4. Big Data mit <i>Orange</i> und vier Szenarien inkl. Präsentation	2x90 min
5. Reflexion:	45 min

### **Bild 1: Reihenplanung der Reihe *Mathematische Modelle zur Prognose (Entscheidungsbaum)* für 10. Klassen (Enrichment-Unterricht Mathematik).**

Der Enrichment-Unterricht findet im Land Berlin im Rahmen der Begabtenförderung statt. Hier sollen abseits des Rahmenlehrplans Projekte zur Vertiefung durchgeführt werden.



The screenshot shows a monkey face with a bandage on its head. Below it is a table with the following data:

beißt	Nummer	Auge links	Auge rechts	Mund	Accessoir
nein	26	offen	zu	Grinsen	Pflaster rechts

Arrows point from the table to labels: 'ja/nein' (under 'beißt'), 'offen/zu/x' (under 'Auge links' and 'Auge rechts'), 'Grinsen/Zunge/Zahn/grimmig/geöffnet/verschmitzt' (under 'Mund'), and 'Pflaster rechts/Pflaster links/keins/Brille/Haarschleife/Augenbinde' (under 'Accessoir').

**Bild 2 (oben):**  
AI Unplugged – Gute-Affchen-Böse-Affchen mit einer möglichen Lösung.

**Bild 3 (links):**  
Attributwerte des (Problem-)Affen 26.

Um diese Erfahrung mit in die digitale Form zu nehmen, wurden die Trainings- und Testdaten in Tabellenform (CSV) zur Verfügung gestellt und mit einem Werkzeug verarbeitet. Daran anschließend wurden für vier unterschiedliche Projekte große Datenmengen Entscheidungsbäume mit dem für Nicht-Informatiker bzw. Nicht-Data-Scientists entworfenen Analysewerkzeug *Orange 3* erstellt und mit Testdaten bewertet. Abschließend wurden die Projektergebnisse und der Modellierungsprozess reflektiert.

tensätze werden von der Lehrkraft gestellt, da dieser Modellierungsschritt aufgrund der Fokussierung auf die Big-Data- und Machine-Learning-Aspekte auch in der folgenden Projektphase nicht durchgeführt wird. Ein Datensatz für ein Gesicht besteht hier aus den Attributwerten für folgende Attribute: *beißt*, *Nummer*, *Auge links*, *Auge rechts*, *Mund* und *Accessoire* (siehe Bild 3).

Gehen wir zunächst den Weg der Schülerinnen und Schüler weiter, bevor wir den Schlamassel auflösen, den diese Modellierung mit sich bringt.

Zur Verarbeitung mit *Orange* müssen – nach dem Import der Daten ins Werkzeug – den Attributen Rollen und Datentypen zugewiesen werden.

## Den Stecker zurück in die Dose

Um das Äffchen-Spiel digital nachzubauen, müssen die Affengesichter als Datensätze vorliegen. Diese Da-

▷ Rollen (beginnen mit einem Großbuchstaben):

- Rolle *Classification*: Das Attribut *beißt* ist Ziel der Untersuchung, nach dem klassifiziert werden soll.
- Rolle *Meta*: Das Attribut *Nummer* ist ein zusätzliches Merkmal. Es dient nur dazu, dass man am

	A	B	C	D	E	F
1	mS#Nummer	D#Auge links	D#Auge rechts	D#Mund	D#Accessoir	cD#beißt
2		1 offen	offen	Zahn	keins	ja
3		2 offen	zu	Zahn	keins	ja

Ende der Untersuchung herausfinden kann, welcher Affe richtig / nicht richtig klassifiziert wurde. Es wird nicht in das Modell einbezogen.

- Alle anderen Attribute sind normale Eingabemerkmale, die untersucht werden, um das Zielmerkmal zu bestimmen.
- ▷ Datentypen (beginnen mit einem Kleinbuchstaben):
  - Datentyp *string*: Das Attribut *Nummer* kann auch als *string* aufgefasst werden, da mit diesem Meta-Attribut keine numerischen Operationen bzw. Vergleiche durchgeführt werden. Es dient nur zur Benennung der Affen.
  - Datentyp *discrete*: Alle weiteren Attribute in diesem Beispiel sind *discrete*, d.h. dass es eine (endliche) Auflistung an Merkmalsausprägungen gibt.

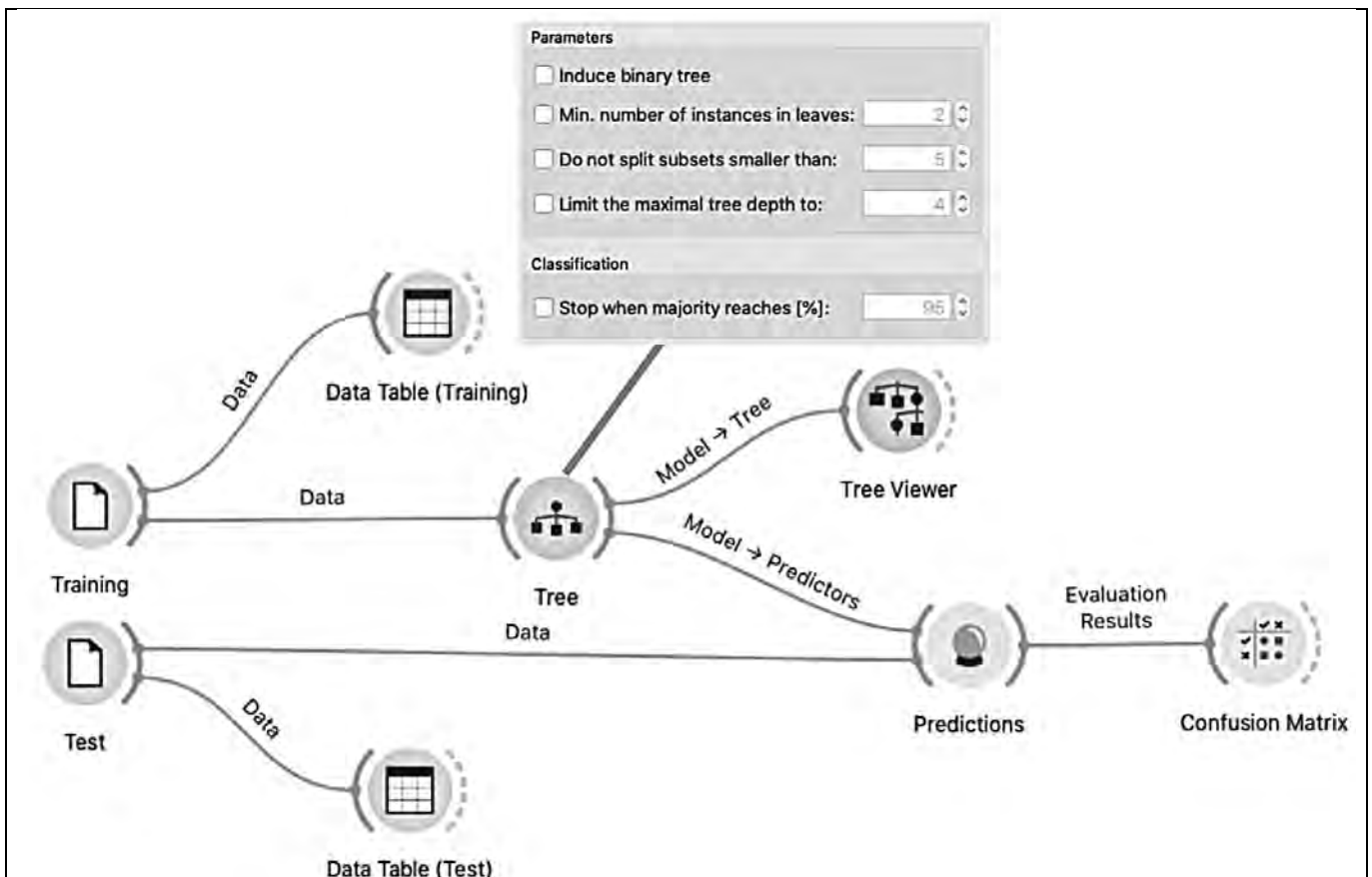
Diese Rollen und Datentypen können bereits in der CSV-Datei der Eingabe festgelegt werden (siehe Bild 4) und mussten von den Schülerinnen und Schülern dort ergänzt werden.

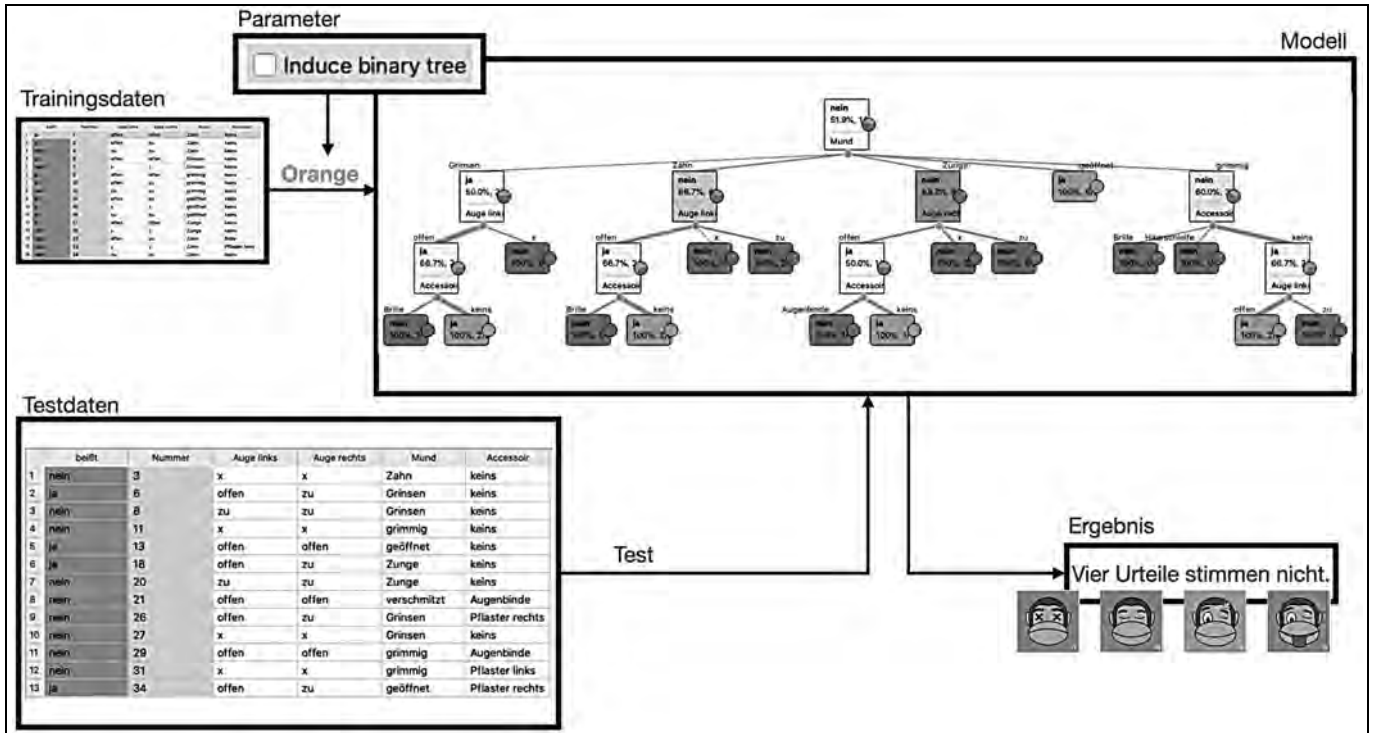
Um nun ein Modell unter Nutzung von *Orange 3* zu erstellen, muss die Analyse in diesem Werkzeug spezifiziert werden. Die Struktur dieser *Orange 3*-Spezifikation ist in Bild 5 dargestellt. Die zwei Dateien für die Trainings- und Testdaten (Training, Test) werden eingelesen. Zur

**Bild 4: Anfang der Eingabedaten mit Angabe von Rolle und Datentyp je Attribut.**

Kontrolle können die Inhalte der Dateien als Tabellen (Data Table (Training), Data Table (Test)) ausgegeben werden. Die Trainingsdaten werden genutzt, um das Modell, einen Entscheidungsbaum, automatisch zu konstruieren (Tree). Dieses Konstruktionsprozess besitzt Parameter, die die Baumstruktur beeinflussen, wie beispielsweise die Entscheidung, ob dieser binär ist oder ob eine Maximaltiefe festgelegt wird. Zur Veranschaulichung des Modells kann dieses als Graph betrachtet werden (Tree Viewer). Unter Nutzung des Entscheidungsbaums wird die gesuchte Merkmalsausprägung für die Testdaten vorhergesagt (Predictions). Zur übersichtlichen Darstellung des Ergebnisses und als Basis für die Bewertung der Analysequalität wird dieses als Konfusionsmatrix angezeigt (Confusion Matrix).

**Bild 5: Orange 3 – Spezifikation des Affenspiels.**





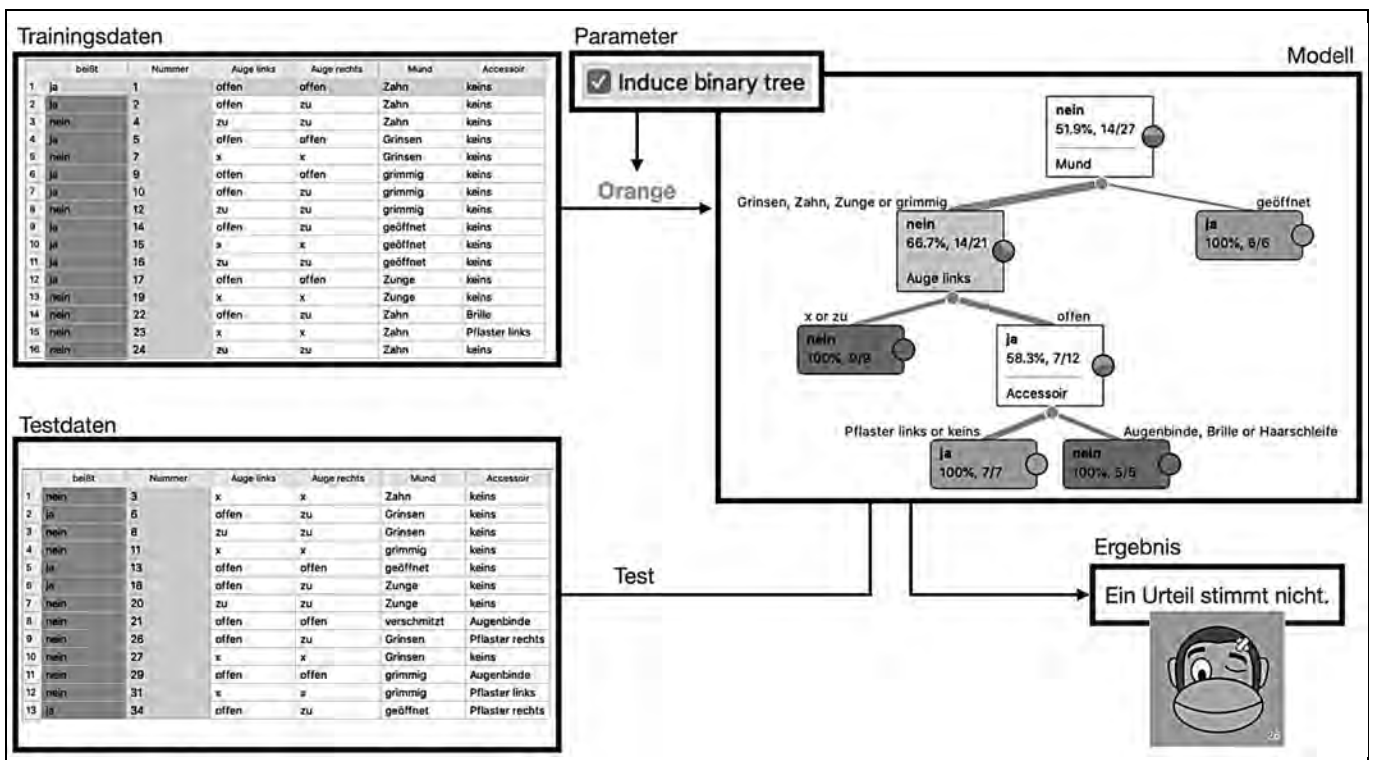
**Bild 6: AI Replugged – Gute-Äffchen-Böse-Äffchen mit einem allgemeinen Entscheidungsbaum.**

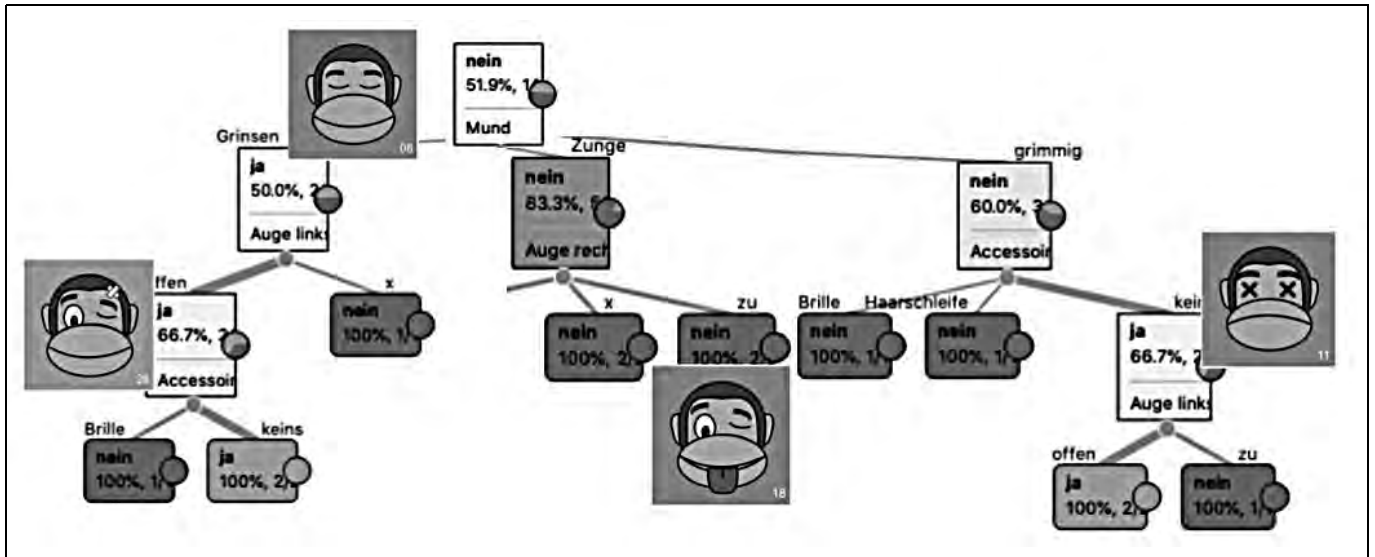
- ▷ der manuelle Entscheidungsbaum besser erscheint als die konstruierten und
- ▷ der eingeschränkte Baum, also der binäre Baum, besser erscheint als der nicht eingeschränkte, also der allgemeine Baum.

Erzeugt man mit dieser Konfiguration einen allgemeinen Baum (siehe Bild 6), so werden vier Affen falsch klassifiziert, bei einem binären Baum (siehe Bild 7) nur einer.

Dieses Ergebnis ist für die Schülerinnen und Schüler überraschend, da (bei diesen Testdaten)

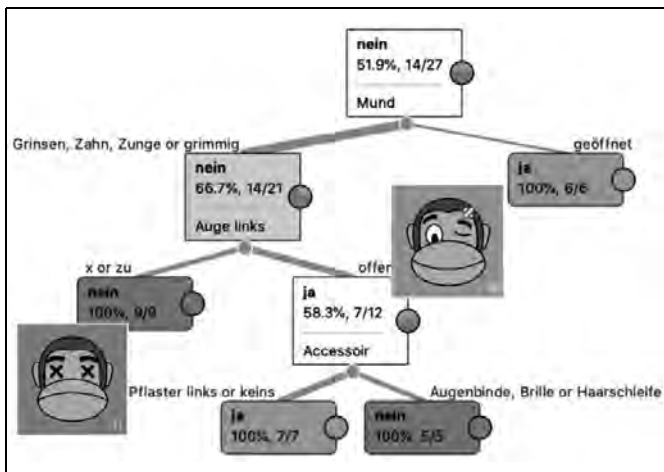
**Bild 7: AI Replugged – Gute-Äffchen-Böse-Äffchen mit einem binären Entscheidungsbaum.**





**Bild 8 (oben):** Ausschnitt aus dem allgemeinen Entscheidungsbaum mit den falsch beurteilten Testaffen.

**Bild 9 (unten):** Binärer Entscheidungsbaum mit einem richtig beurteilten Affen (11) und einem falsch beurteilten Affen (26).



Schaut man sich die drei Entscheidungsbäume genauer an, so haben diese individuelle Strukturmerkmale:

▷ **manueller Baum:**

Entscheidungen werden mit ja/nein-Fragen getroffen. Dies bedeutet, dass unabhängig von den Attributwerten jeder Testaffe bis zu einem Blattknoten vorrückt. Urteile werden also nur in den Blattknoten getroffen.

▷ **erzeugte Bäume:**

Alle Zweige werden mit einem oder mehreren Attributen beschriftet. Damit kann ein Affe, der für das gefragte Attribut einen Attributwert besitzt, der an keinem Ausgangszweig steht, im Baum hängen bleiben. Es müssen also auch Urteile an Zwischen- und Wurzelknoten möglich sein. In jedem Knoten ist deshalb ein Urteil notiert (ja oder nein). Wie dieses Ur-

teil zustande kommt, kann an dem Tortendiagramm des jeweiligen Knotens abgelesen werden. Dieses stellt die Verhältnisse der Trainingsdaten dar, die diesen Knoten repräsentieren. Es entscheidet die Mehrheit der Trainingsdaten, welches Urteil getroffen wird. Bei Gleichstand entscheidet der erste Datensatz der Trainingsdaten, in unserem Fall ein beißender Affe (Urteil: ja).

• **allgemeiner Baum:**

Jeder Ausgangszweig wird mit genau einem Attribut beschriftet. Dies bedeutet, dass jeder Knoten gewisse Ausprägungen eindeutig festlegt. Diese können am Pfad von der Wurzel zu diesem Knoten abgelesen werden. Für diese Ausprägungen muss es mindestens einen Affen in den Trainingsdaten gegeben haben, der diese besitzt.

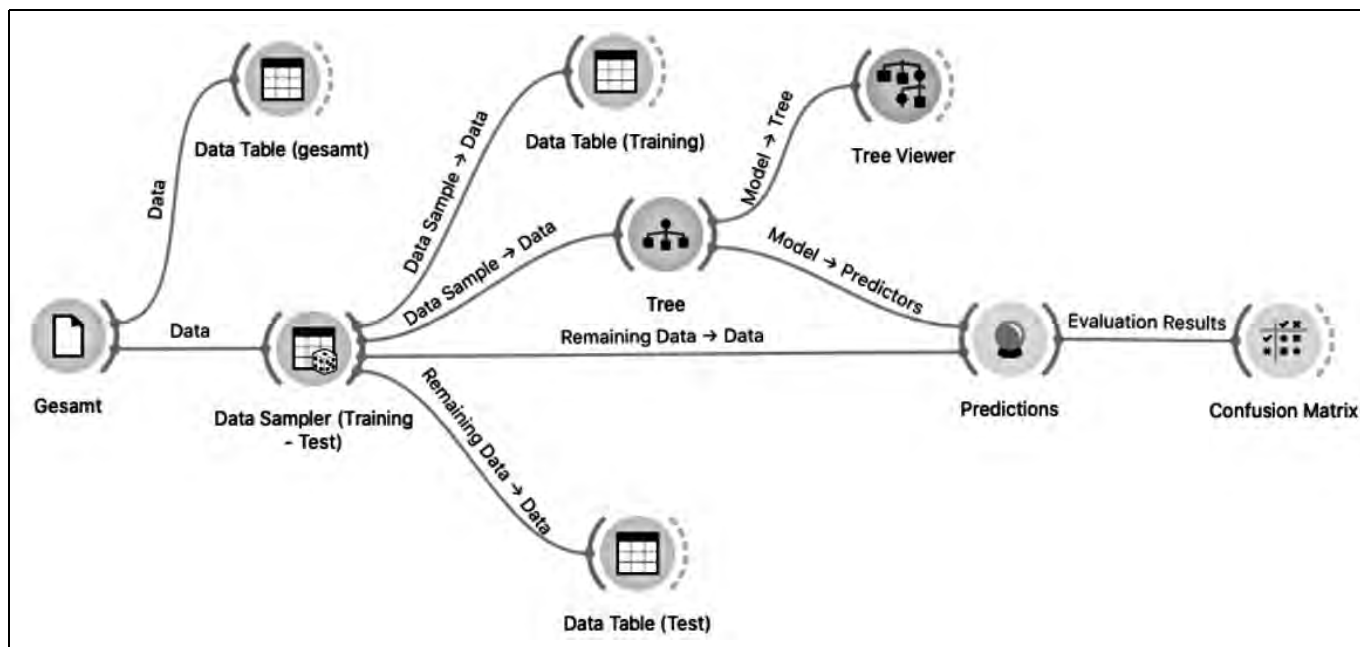
Im Bild 8 wird ein Ausschnitt aus dem erzeugten, allgemeinen Entscheidungsbaum gezeigt zusammen mit den Testaffen mit falschem Urteil. Der Affe 11 (rechts) schaut grimmig und besitzt kein Accessoire. Da sein linkes Auge weder offen noch zu ist, kann er zu keinem Blattknoten absteigen. Da zwei der drei Trainingsaffen, die diesen Knoten repräsentieren, beißen, wird auch der Affe 11 als bissig beurteilt (Urteil: ja).

• **binärer Baum:**

Jeder Ausgangszweig wird mit mehreren Attributen beschriftet. Diese werden mit einem logischen oder verknüpft. Damit legen die Eigenschaften mehrerer Trainingsaffen die Anforderungen für einen Knoten fest.

In Bild 9 ist die Situation für den binären Baum dargestellt. Der Affe 26 hat seinen Knoten erreicht, da er ein Grinsen zeigt und das linke Auge offen hält. Er kann den Knoten nicht verlassen, da ein Ausgang für sein Pflaster rechts fehlt. Da sieben der zwölf Trainingsaffen, die diesen Knoten darstellen, bissig sind, wird der Affe als bissig eingestuft. Dieses Urteil ist falsch.

Der Affe 11 erreicht seinen Knoten, da er grimmig schaut und das linke Auge ein X darstellt. In den Testdaten ist kein Affe enthalten, der diese beiden Eigenschaften vereint. Die Knoten können also



**Bild 10: Orange 3 – Spezifikation für die Projektphase.**

Überlagerungen verschiedener Trainingsaffinen darstellen.

Mit dieser letzten Sicht erscheint der binäre Baum flexibler als der allgemeine Baum.

▷ Lebensmittelchemie (Bestimmung der Weinqualität aus Messungen):

- Daten aus Messungen von Weinproben (Säuregehalt, Chloride, pH-Wert, ...)
- gesuchtes Attribut: Qualität des Weins

Die Schülerinnen und Schüler sollten aus den gegebenen Datensätzen die Attribute zur Analyse auswählen und entsprechend vorbereiten. Bei der Analyse sollten auch allgemeine und binäre Entscheidungsbaummodelle verglichen werden.

Grundsätzlich konnten alle Gruppen feststellen, dass beide Entscheidungsbaumvarianten Ergebnisse in ähnlicher Qualität erzeugen. Die Vorhersage hatte in der Regel auch Fehler.

Die Gruppen mit dem Wirtschaftsthema waren darüber überrascht, dass in den amerikanischen Datensätzen auch das Attribut Rasse gespeichert wurde. Eine Gruppe zum Thema *Schule* irritierte die Möglichkeit der Integration vermeintlich sachfremder Attribute zur Notenvorhersage, wie das Attribut Bildungsabschluss der Mutter.

## Big Data Projekte

Die Projekte werden mit einer leicht veränderten *Orange 3*-Struktur bearbeitet (siehe Bild 10). Dabei wird nur eine große Datei (Gesamt) eingelesen und die Datensätze werden von einem Data Sampler in Trainings- und Testdaten (Data Sample und Remaining Data) aufgeteilt. Diese Trainings- und Testdaten werden dann, wie bekannt, analysiert.

Die vier Projekte waren (wie von Grillenberger vorgeschlagen):

- ▷ Schule (Notenvorhersage für fehlende Schülerinnen und Schüler):
  - Schüler- und Elterndaten von portugiesischen Schülerinnen und Schüler
  - gesuchtes Attribut: Note einer Klassenarbeit
- ▷ Medizin (Identifizierung gutartiger und bösartiger Krebszellen):
  - Informationen über Zellkerne, die durch bildgebende Verfahren gewonnen wurden
  - gesuchtes Attribut: Einstufung der Krebszellen
- ▷ Wirtschaft (Vorhersage des Einkommens anhand von Zensusdaten):
  - Umfangreiche Zensusdaten (Alter, Geschlecht, Bildung, Beschäftigung, ...)
  - gesuchtes Attribut: Einkommensklassen

## Und was zeigt uns das eingesteckte Affenspiel?

Da das Affenspiel ausgedacht ist, kann es zwischen den Zeichnungen nur Korrelationen, aber keine Kausalitäten geben. Dies unterstreicht die Tatsache, dass der Entscheidungsbaum nur Korrelationen zwischen den Attributen und dem Urteil abbilden kann und Kausalitäten nicht betrachtet. Es gilt also:

- ▷ Überwachtes Lernen untersucht Korrelationen in den Trainingsdatensätzen und keine Kausalitäten.

Bereits dies ist nicht allen Schülerinnen und Schülern klar, da sie es gewohnt sind, im Ursache-Wirkungszusammenhang zu denken.

Alle weiteren Aussagen können nicht streng wissenschaftlich aus diesem einen Beispiel abgeleitet werden, plausibilisieren aber einige typische Aussagen über überwachtetes Lernen.

▷ Urteile können fehlerhaft sein:

Die beiden erzeugten Bäume urteilen bereits nicht vollständig richtig. Der Eindruck, dass die Ergebnisse nicht immer stimmen, wird durch die Projekte nochmals unterstrichen.

▷ Modelle können nicht manuell verbessert werden:

Das Werkzeug *Orange 3* besitzt keine Möglichkeit, den erzeugten Baum manuell zu verändern. Die einzigen Handlungsoption, die einem gegeben ist:

- erneutes Training mit
  - veränderten Parametern und/oder
  - veränderten Trainingsdaten.

▷ Die Entwicklerin bzw. der Entwickler benötigt Erfahrung bei der Wahl der Parameter:

Die verschiedenen Parameter für die Baumerzeugung haben Auswirkungen auf die Struktur des Baums und unter Umständen auf die Güte der Vorhersage. Die verschiedenen Beschreibungen der zwei Baumarten lassen mal den einen, mal den anderen als vorteilhaft erscheinen:

- Baum ohne Einschränkung (allgemeiner Baum) / Baum mit Einschränkung des Verzweigungsgrads (binärer Baum)
- ein Knoten kann durch mindestens einen Trainingsaffen dargestellt werden (allgemeiner Baum) / ein Knoten kann durch eine Überlagerung mehrerer Trainingsaffen dargestellt werden (binärer Baum).

Für die Wahl günstiger Parameter benötigt der Entwickler also Erfahrung.

Mit der zweiten Beschreibung der Eigenschaften der Bäume kann der allgemeine Baum aus unserem Beispiel mit typischem KI-Vokabular beschreiben werden:

Der Baum ist an die Trainingsdaten überangepasst (engl. *overfitting*).

Er hat die Trainingsbasis auswendig gelernt und zu wenig abstrahiert. Somit treten bei Übertragung auf nicht in den Trainingsdaten enthaltene Datensätze Fehler auf.

Diese Aussagen sind für dieses kleine Beispiel sicher zu hoch gegriffen, zeigen aber exemplarisch den Einsatz dieser Sprechweisen.

Und warum konnte *Orange 3* die Affen nicht so gut einschätzen wie die manuelle Lösung?

Nun, das liegt an der Modellierung der Affengesichter. Die Festlegung der Attribute und Attributwerte sollte sich nicht an der manuellen Lösung orientieren. Deshalb wurde das linke vom rechten Auge getrennt und die Accessoires wurden einzeln aufgezählt. Da die Augen in zwei Attribute behandelt werden, gibt es in einem Entscheidungsknoten die Abfrage aus der manuellen Lösung nicht: „mindestens ein Auge offen“. Möchte man diese Abfrage erreichen, so müssen das linke und rechte Auge nacheinander geprüft werden. Die einzelne Abfra-

ge beider Augen erhöht die Tiefe des Baums. Da flachere Bäume vom Algorithmus bevorzugt werden, wurde die doppelte Abfrage der Augen nicht berücksichtigt.

Dass die Modellierung für das Ergebnis wichtig ist, führt zu einer weiteren typischen Aussage über überwachtetes Lernen:

▷ Die Entwicklerin bzw. der Entwickler benötigt Erfahrung bei der Modellierung von Daten, da die Darstellung der Daten einen hohen Einfluss auf die Qualität des Verfahrens hat.

Selbstverständlich benötigt ein Entwickler auch einen Erfahrungsschatz, um sich für ein geeignetes Lernverfahren (Entscheidungsbaum, Lineare Regression, Random Forest, neuronales Netz usw.) zu entscheiden. Da in dieser Reihe nur ein Modell – der Entscheidungsbaum – behandelt wurde, fehlt hier aber ein Anknüpfungspunkt, um dies exemplarisch darzustellen.

## Fazit

Den Schülerinnen und Schülern gelang es in der Projektphase, die Entscheidungsbäume auch aus einem großen Bestand an Trainingsdaten zu erstellen. Damit konnten sie das grundsätzliche Vorgehen bei der Modellerstellung durch Trainingsdaten mit anschließendem Test selbst erleben. Hilfreich war dabei die einfache Bedienbarkeit von *Orange 3*.

Den Schülerinnen und Schülern wurde bewusst, dass sich überwachtetes Lernen typischerweise in Trainings- und Testphase gliedert, und als Einschränkung wiederholt genannt, dass ein Modell nicht in jedem Fall vollständig richtige Ergebnisse liefert.

Insgesamt waren die Schülerinnen und Schüler überrascht, wie leicht derart große Datenmengen analysiert werden können.

Dr.-Ing. Lennard Kerber  
Otto-Nagel-Gymnasium  
Schulstraße 11  
12683 Berlin-Biesdorf

E-Mail: [l.kerber@otto-nagel-gymnasium.de](mailto:l.kerber@otto-nagel-gymnasium.de)

## Literatur und Internetquellen

Lindner, A.; Seegerer, St.: AI Unplugged – Künstliche Intelligenz ohne Computer unterrichten. In: LOG IN, 40. Jg. (2020a), Nr. 193/194, S. 71–74.

Lindner, A.; Seegerer, St.: AI Unplugged – Wir ziehen künstlicher Intelligenz den Stecker. Aktivitäten und Unterrichtsmaterial zu künstlicher Intelligenz ohne Strom. 2020b.  
<https://www.aiunplugged.org/>

Orange – Data Mining. 1996 ff.  
<https://orange.biolab.si/>

Alle Internetquellen wurden zuletzt am 18. Februar 2021 geprüft und können auch aus dem Service-Bereich des LOG IN Verlags (<https://www.log-in-verlag.de/>) heruntergeladen werden.