

Aufgabe 2: Entscheidungsbaum erstellen

Situation: Sie haben einen Datensatz mit 14 Fischen, von denen bekannt ist, ob diese friedlich (grünes Blatt) oder feindselig (Fischgräte) sind (siehe Darstellung unten). Für die beiden Fische rechts ist zu entscheiden, ob Sie unbedenklich ins Aquarium gegeben werden können oder nicht.

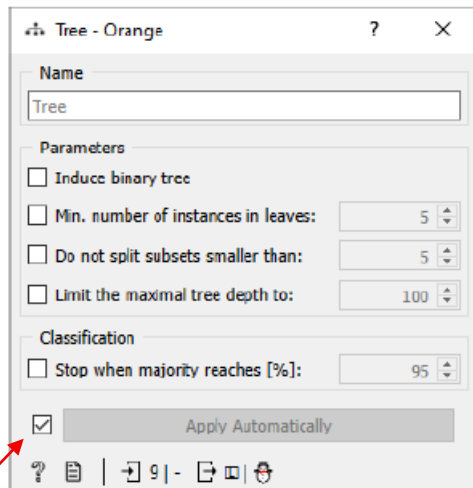


Aufgabe: Entscheidungsbaum erstellen

1. Betrachten Sie zunächst die Bilder aller 16 Fische und benennen Sie die auftretenden Fischmerkmale mit deren Ausprägungen. Achten Sie bei Ihren Bezeichnungen auf Eindeutigkeit (Farbe allein genügt nicht).
2. Stellen Sie für die beiden rechts stehenden Fische eine Vermutung an, in welche Klasse (friedlich oder feindselig) sie gehören. Begründen Sie Ihre Vermutung.
3. Lesen Sie mit dem Widget „File“ den **fischeinstiegtrainingsdaten.csv** in Orange ein und überprüfen Sie, ob das Merkmal „Label“ (friedlich/feindselig) als Zielmerkmal (= target) gesetzt wurde (erkennbar am Fettdruck).
Fehlt die Angabe eines Zielmerkmals, dann stellen Sie die Rolle (role) des Merkmals „Label“ auf „target“ um.
Überführen Sie die Daten in das Widget „Data Table“ und vergleichen Sie die Bezeichnungen der Merkmale der Fische in der Datentabelle mit denen von Ihnen beschriebenen (siehe Aufgabe 1). Notieren Sie etwaige Unterschiede.

4. Erstellen Sie mit dem Widget „Tree“ einen Entscheidungsbaum und machen Sie diesen mit dem Widget „Tree Viewer“ sichtbar.

Da in diesem Übungsbeispiel nur sehr wenige Daten verarbeitet werden, brauchen Sie bei den Hyperparametereinstellungen für das Widget „tree“ keine Einstellungen vornehmen.



Bei einem geringen Datenumfang ist es nicht nötig, die Baumtiefe oder die Anzahl Blätter (über die Zahl der Instanzen) oder die Zahl der Knoten (über die Zahl der Subsets) zu begrenzen.

Haken setzen!

5. Ermitteln Sie mithilfe Ihres Entscheidungsbaums, ob die beiden im Bild oben noch nicht klassifizierten Fische friedlich oder feindselig sind und notieren Sie deren Klasse.

Vergleichen Sie anschließend die ermittelte Klasse mit der, die Sie in der Aufgabe 2 vermutet haben. Notieren Sie ggf. Auffälligkeiten.

6. Importieren Sie mit nun den Datensatz **fischeEinstiegtestdaten.csv** in einem zweiten „File“-Widget, um den gefundenen Entscheidungsbaum mithilfe dieser Testdaten auf seine Korrekt-Klassifikationsrate zu überprüfen. Verwenden Sie dafür die Widgets „Prediction“ und „Confusion Matrix“.

Notieren Sie die KKR (= CA) des Entscheidungsbaums, die Zahl der korrekt klassifizierten Testdatenfische und umkreisen Sie den Fisch der Testdatenreihe (siehe: Seite 3), der falsch klassifiziert wurde.

Testdaten:





Hinweis: Die verwendeten Materialien der Universität Passau stammen aus einem Mebis-KI-Informatik-Kurs 11. Die Materialien wurden von Wolfgang Pfeffer, Tobias Fuchs und Ute Heuer erstellt. (Link: <https://lernplattform.bycs.de/course/view.php?id=1324361>)