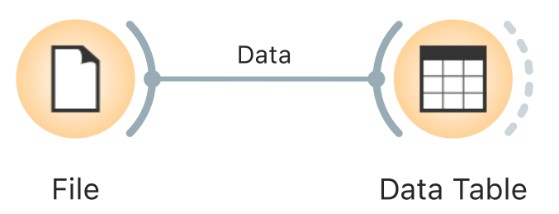
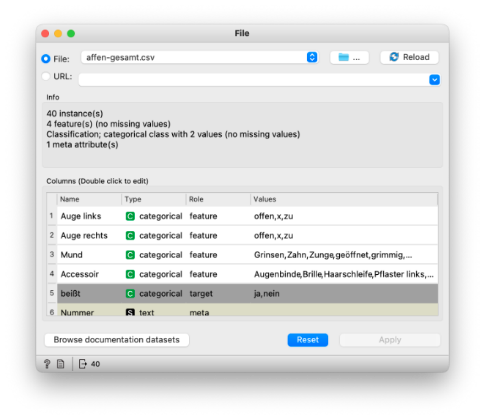
Um Daten einzulesen, gibt es das Widget *File* im Abschnitt *Data*.

Das Konfigurationsmenü erreicht man mit einem Doppelklick auf das Widget. Dort muss die Datei ausgewählt werden ( ).





**Feature**: Zu untersuchende Merkmale, um Zielklasse zu bestimmen (z.B. „Auge links“ oder „Mund“)



Nach dem Laden können die Merkmale mit ihren Datentypen, Rollen und möglichen Ausprägungen überprüft werden.

Type:

C **Categorical**: Es gibt nur endlich viele Merkmalsausprägungen

(z.B. Auge links)

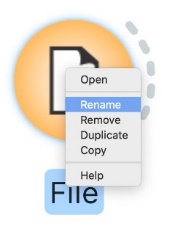
S **Text**: Zeichenketten (z.B. Name)

N **Numeric**: Zahlenwerte (z.B. Alter)

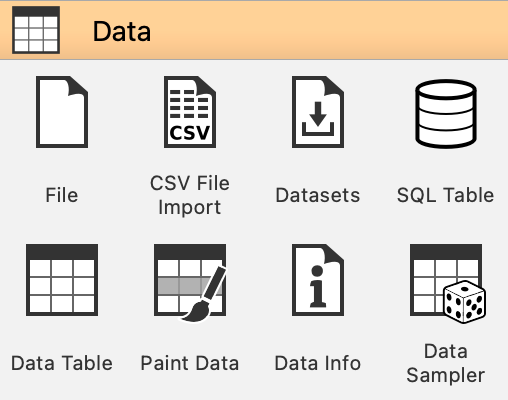
Role:

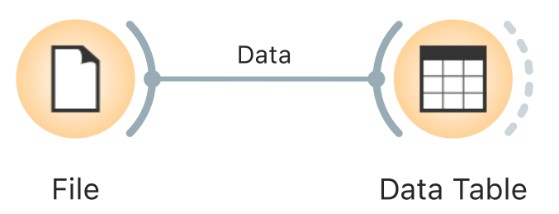
**Target**: Zu bestimmende Klasse (z.B. „beißt“ mit den Ausprägungen „ja/nein“)

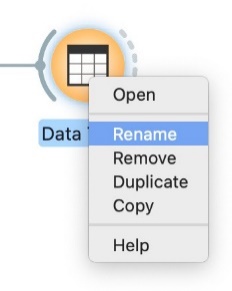
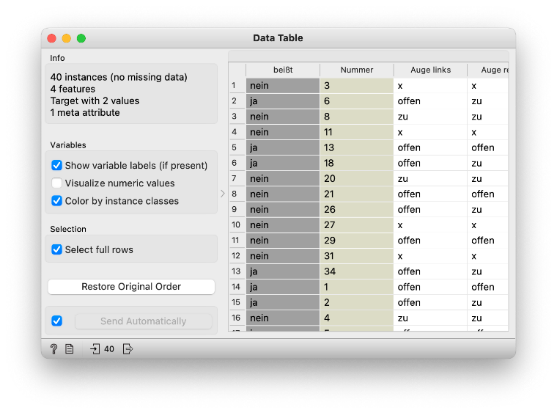
**Meta**: Zusatzinformationen, die für Untersuchung nicht wichtig sind (z.B. Äffchennummer)



Mit einem Rechtsklick auf das Widget kann dieses umbenannt werden (Rename).

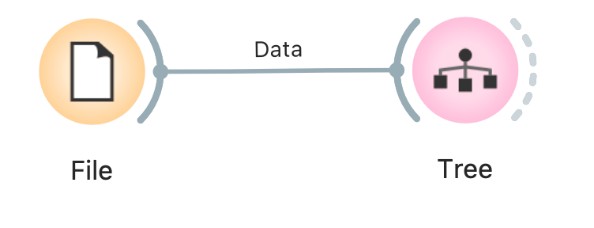
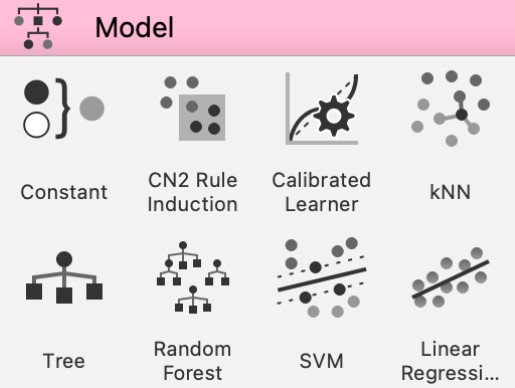
Um sich den Inhalt einer Tabelle anzusehen, gibt es das Widget *Data Table* im Abschnitt *Data*.

Du musst eine Datenquelle an das neue *Widget* anschließen. Die Quelle ist das *File*-Widget. Die Verbindung wird mit klicken und ziehen von der Quelle (File) zur Senke (Data Table) hergestellt.



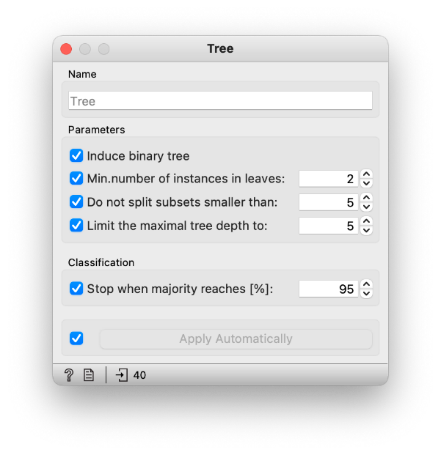
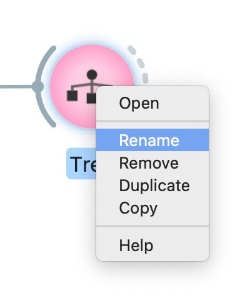
Mit einem Doppelklick auf das Widget kann die Tabelle betrachtet werden.

Mit einem Rechtsklick auf das Widget kann dieses umbenannt werden (Rename).

Für das Erzeugen von Entscheidungsbäumen, die Entscheidungen in Form hierarchischer Regeln abbilden, gibt es das Widget *Tree* im Abschnitt *Model*.

Du musst eine Datenquelle an das neue Widget anschließen. Die Quelle ist unser *File*-Widget mit den Trainingsdaten. Die Verbindung geschieht mit klicken und ziehen von der Quelle (File) zur Senke (Tree).

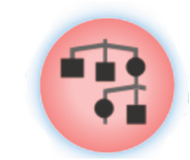
Mit einem Doppelklick auf das Widget können Eigenschaften des zu erzeugenden Baums festgelegt werden.

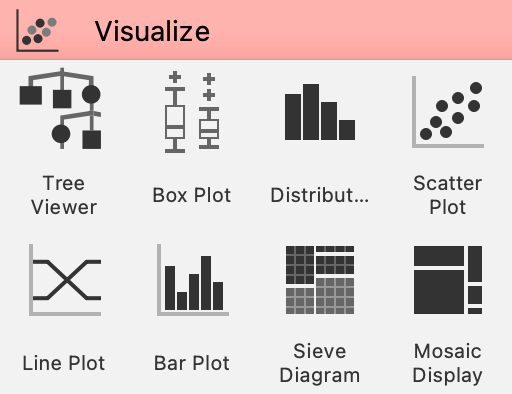


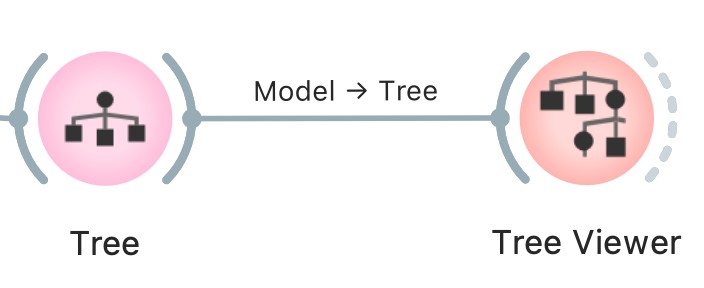
Im Beispiel rechts wurde festgelegt, dass der zu erzeugende Entscheidungsbaum ein binärer Baum sein soll

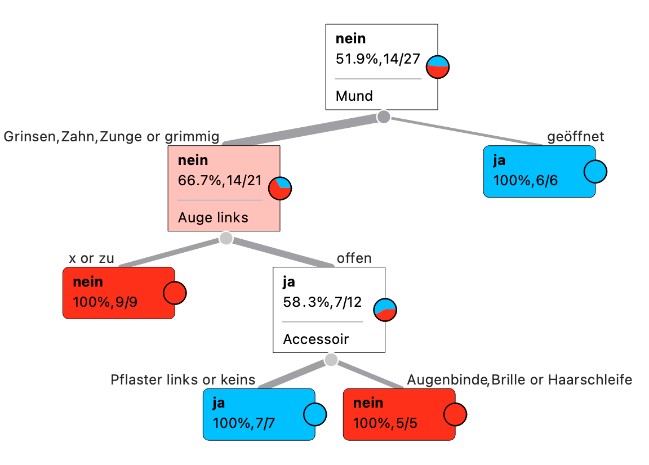
( Induce binary tree), d.h. jeder Knoten verfügt über maximal zwei nachfolgende Knoten.

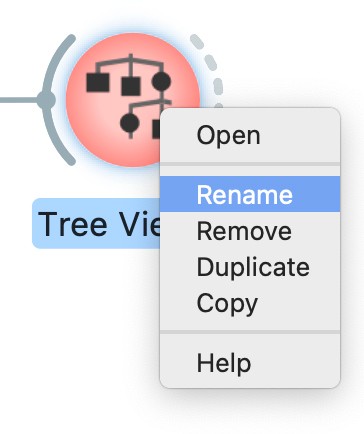
Mit einem Rechtsklick auf das Widget kann dieses umbenannt werden (Rename).



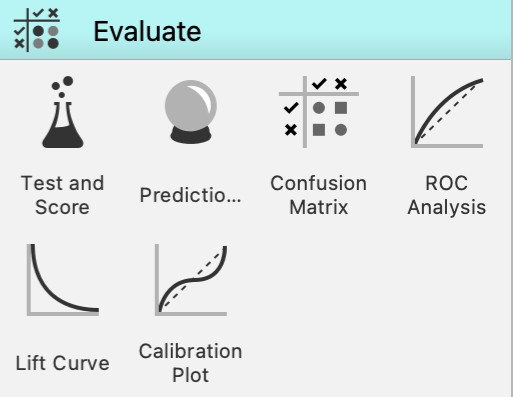
Zur Darstellung eines Entscheidungsbaums gibt es das Widget *Tree Viewer* im Abschnitt *Visualize*.

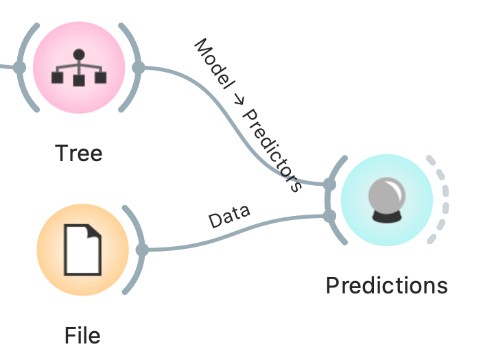
Du musst einen Entscheidungs- baum als Quelle an das neue Widget anschließen. Die Quelle ist ein *Tree*-Widget. Die Verbindung wird mit klicken und ziehen von der Quelle (Tree) zur Senke (Tree Viewer) hergestellt.

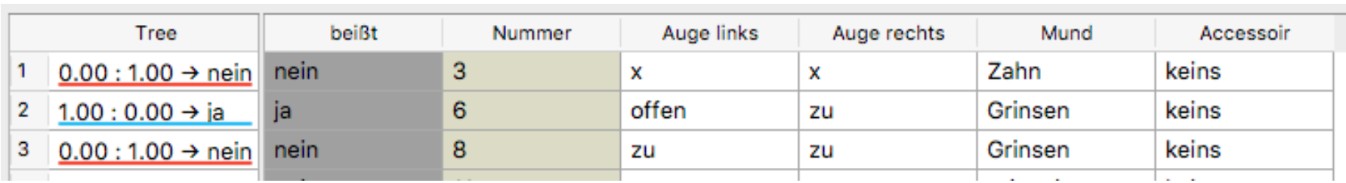
Mit einem Doppelklick auf das Widget kann man den Entscheidungsbaum betrachten. Beachte dabei auch die Details, wie Einfärbung der Knoten und Linienstärke der Kanten.

In den Knoten sind auch die Anzahl an Exemplaren (engl. instance) vermerkt, für die der Knoten steht. So vertritt der Wurzelknoten hier 27 Affen.

Mit einem Rechtsklick auf das Widget kann dieses umbenannt werden (Rename).

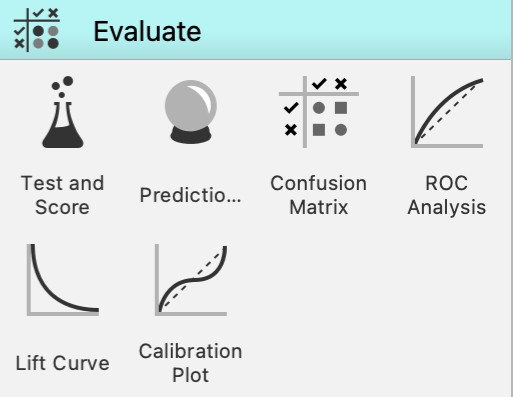
Zur Nutzung eines Entscheidungsbaums für Vorhersagen (also das Anwenden auf Datensätze) gibt es das Widget *Predictions* im Abschnitt *Evaluate*.

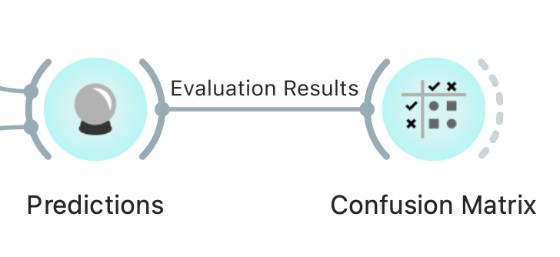
Du musst eine Datenquelle und ein Modell an das neue Widget anschließen. Die Datenquelle ist ein *File*-Widget, das Modell unser Baum. Die Verbindung geschieht mit klicken und ziehen von der Quelle (File sowie Tree) zur Senke (Predictions).

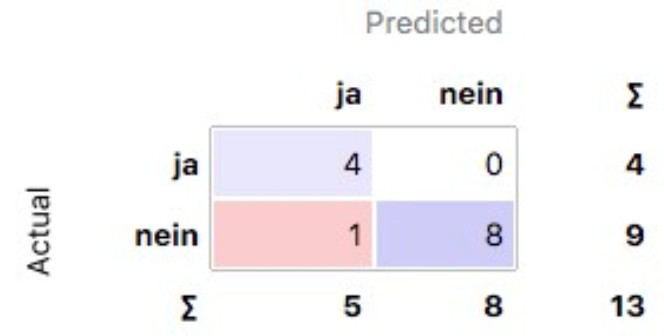


Mit einem Doppelklick kann die Prognose direkt mit den Daten verglichen werden. Für eine bessere Übersicht kann die Trennung zwischen der Baumvorhersage, links, und den Daten, rechts, mit der Maus verschoben werden (so wie oben geschehen).

Hier kann man beispielsweise sehen, dass die Äffchen 3, 6 und 8 richtig vorhergesagt wurden. Äffchen 3 und 8 beißen nicht (nein) und Äffchen 6 beißt (ja).

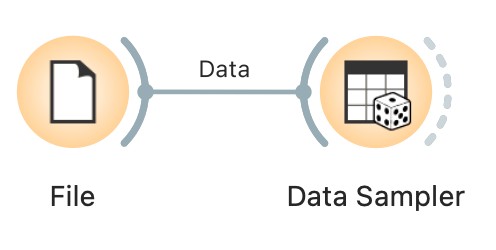
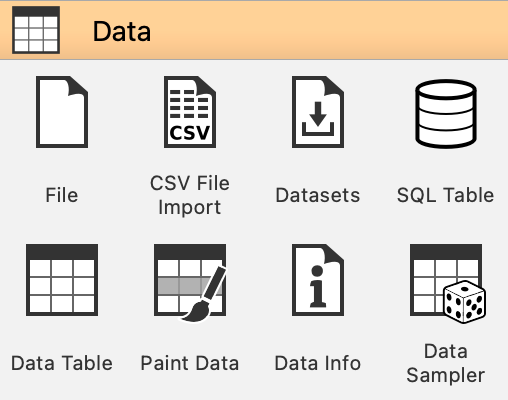
Zur Darstellung einer Konfusionsmatrix gibt es das Widget *Confusion Matrix* im Abschnitt *Evaluate*.

Du musst die Vorhersage als Quelle an das neue Widget anschließen. Die Quelle ist ein *Predictions*-Widget. Die Verbindung geschieht mit klicken und ziehen von der Quelle (Predictions) zur Senke (Confusion Matrix).



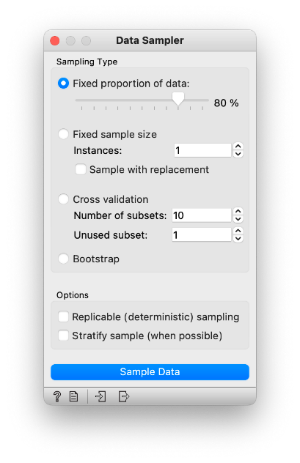
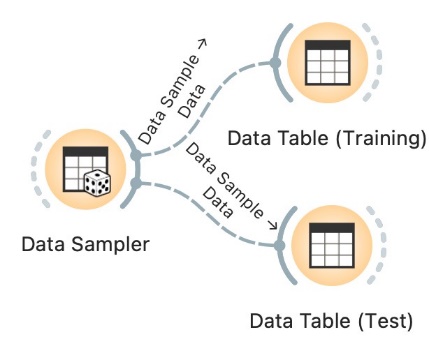
Mit einem Doppelklick auf das Widget kann die Konfusionsmatrix betrachtet werden.



Um einen großen Datensatz in zwei Teile zu zerlegen, gibt es das Widget *Data Sampler* im Abschnitt *Data*. Dies ist hilfreich, um einen Datensatz in Trainings- und Testdaten aufzuteilen.

Du musst eine Datenquelle an das neue Widget anschließen. Die Quelle ist ein *File*-Widget. Die Verbindung wird mit klicken und ziehen von der Quelle (File) zur Senke (Data Sampler) hergestellt.

Der Data Sampler erzeugt nun zwei neue Quellen (die je zwei Teile des Datensatzes enthalten):



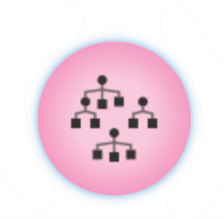
* Data Sample
* Remaining Data

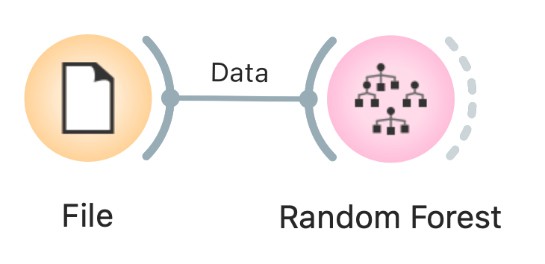
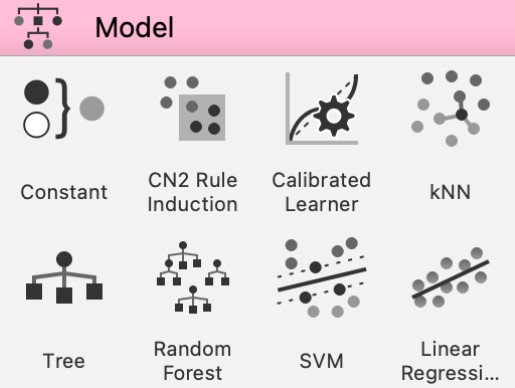
Verbindest du diese Quellen mit anderen Widgets, so musst du nachträglich kontrollieren, ob Orange die richtige Quelle gewählt hat.

Mit einem Doppelklick auf das Widget *Data Sampler* kann der Anteil der Trainingsdatensätze (*Data Sample*) an den gesamten Datensätzen festgelegt werden (z.B.

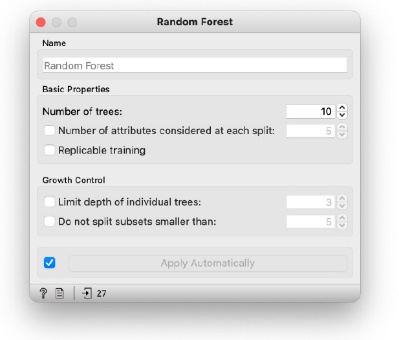
80%).

Zusätzlich kann auch ein erneutes Auswürfeln des Trainingsdaten- satzes gestartet werden (*Sample Data*).



Random Forest Modelle enthalten mehrere Entscheidungsbäume und stimmen im Mehrheitsverfahren ab. Für Random Forests gibt es das Widget Random Forest im Abschnitt Model.

Du musst eine Datenquelle an das neue Widget anschließen. Die Quelle ist ein *File*-Widget mit den Trainingsdaten. Die Verbindung wird mit klicken und ziehen von der Quelle (File) zur Senke (Random Forest) hergestellt.



Mit einem Doppelklick auf das

Widget können Eigenschaften des

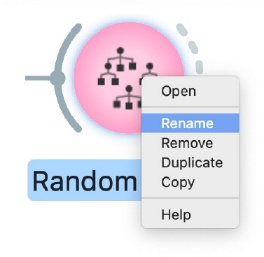
zu erzeugenden *Random Forest*

festgelegt werden.

Im Beispiel rechts wurde

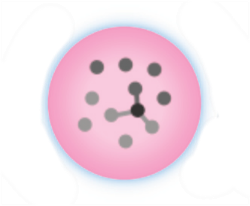
festgelegt, dass dieser aus 10

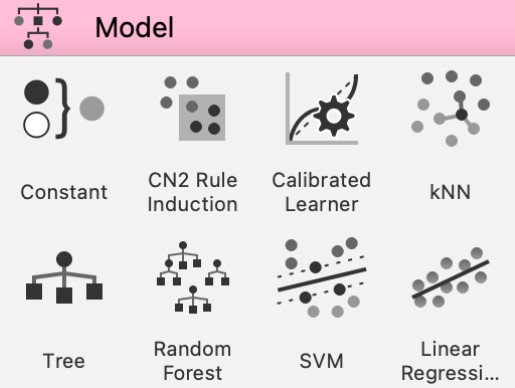
Bäumen bestehen soll.

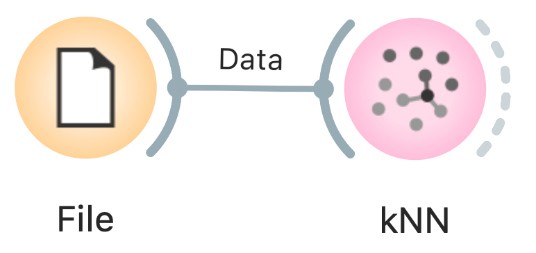


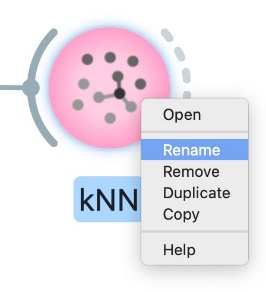
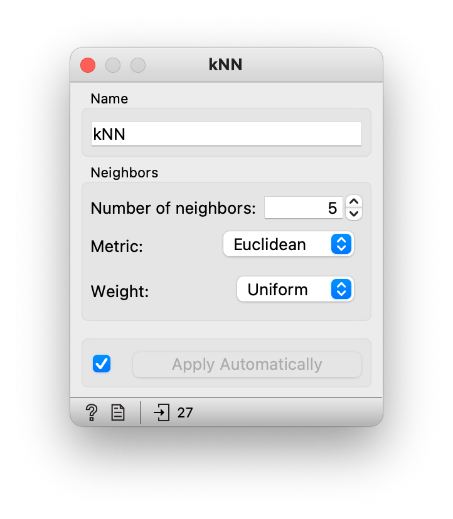
Mit einem Rechtsklick auf das

Widget kann dieses umbenannt   
werden (Rename).



Das k-Nächste-Nachbarn-Modell entscheidet auf Basis der Zielmerkmale der Datenpunkte mit den ähnlichsten Merkmalsausprägungen (die nächsten Nachbarn). Für ein k- Nächste-Nachbarn-Modell gibt es das Widget *kNN* im Abschnitt *Model*.

Du musst eine Datenquelle an das neue Widget anschließen. Die Quelle ist ein File-Widget mit den Trainingsdaten. Die Verbindung wird mit klicken und ziehen von der Quelle (File) zur Senke (kNN) hergestellt.



Mit einem Doppelklick auf das Widget können Eigenschaften des zu erzeugenden kNN-Modells festgelegt werden.

Im Beispiel rechts wurde festgelegt, dass zur Klassifikation stets die k=5 nächsten Nachbarn berücksichtigt werden sollen.

Mit einem Rechtsklick auf das Widget kann dieses umbenannt werden (Rename).