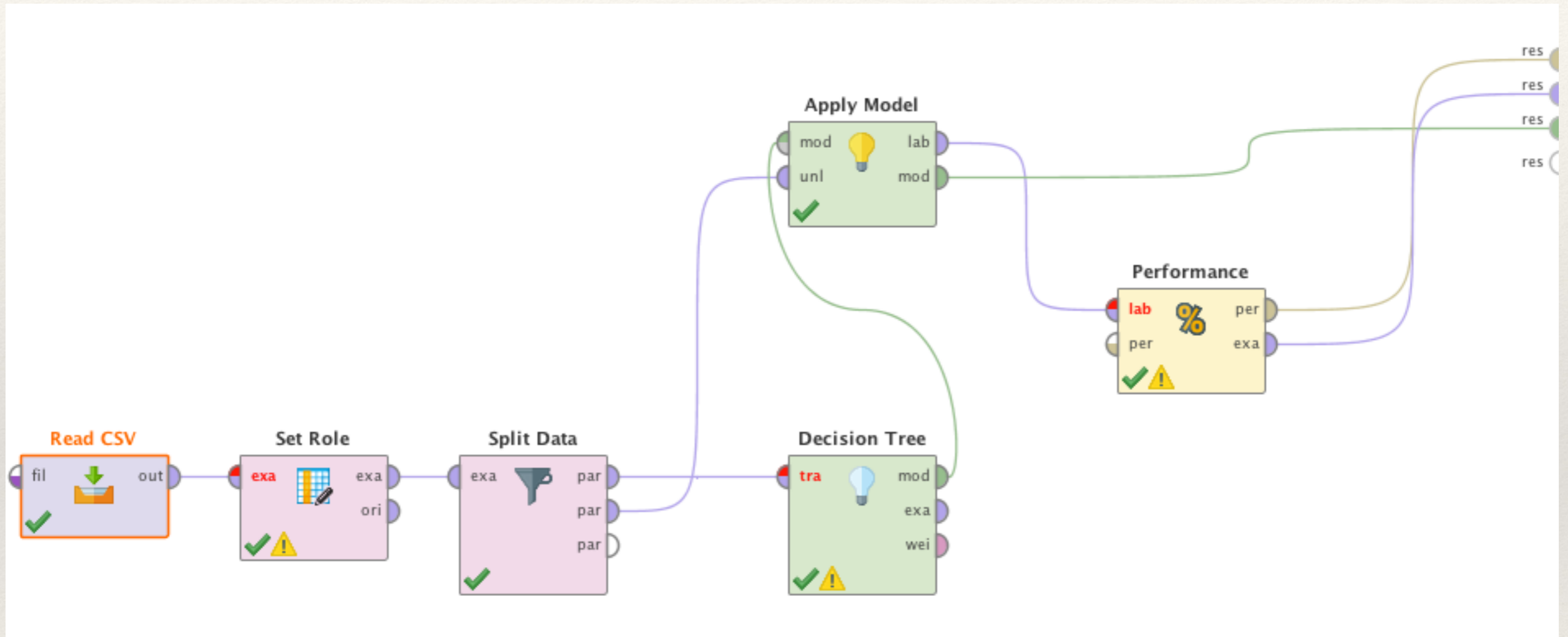


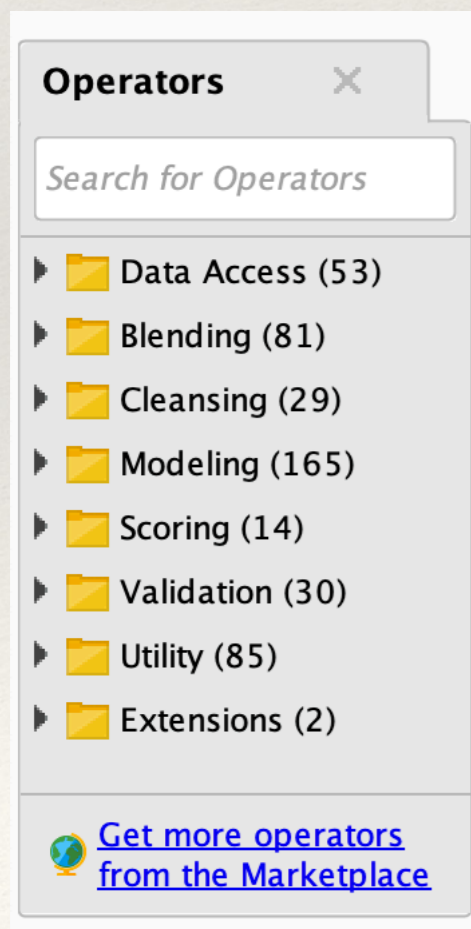
Example in RapidMiner



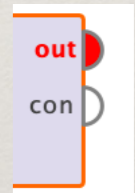
Example in RapidMiner

Start RapidMiner, and choose *File* and *New Process* from the menu. A empty process window should open.

The operators shown in the previous slide can be found in the *Operators* tab. Find each operator using the *Search* function and drag them into the empty process area.



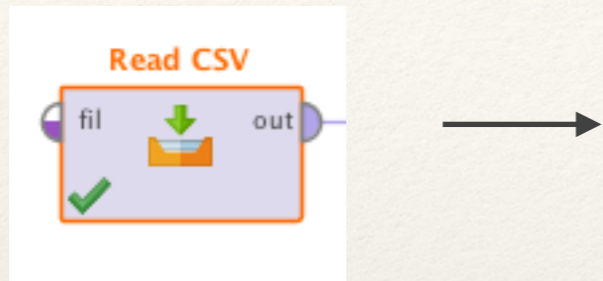
Join operators by clicking on the semi-circles (connections) on the edges of the operator and dragging to the appropriate connections on other operators.



Operators that are not currently installed may be download by clicking on the *Get more operators from the Marketplace* link.

Connections between operators that were created by mistake can be removed by clicking on the connection and pressing Delete.

Example in RapidMiner



These are the parameters for the Read CSV operator.

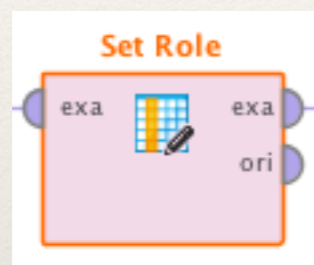
The parameters tab is normally located to the right of the RapidMiner window.

Set the parameters for this operator as shown here.


The image shows the 'Parameters' dialog box for the 'Read CSV' operator. The dialog has a title bar with 'Parameters' and a close button. Below the title bar, there is a tab labeled 'Read CSV' with a green downward arrow icon. A button labeled 'Import Configuration Wizard...' is located below the tab. The main area of the dialog contains several parameter fields and checkboxes. The 'csv file' field is set to 'iris.csv' and has a folder icon to its right. The 'column separators' field is set to ','. The 'use quotes' checkbox is checked. The 'quotes character' field is set to '"'. The 'skip comments' checkbox is unchecked. The 'parse numbers' checkbox is checked. The 'decimal character' field is set to '.'. The 'grouped digits' checkbox is unchecked. The 'date format' field is a dropdown menu set to 'Enter value...' and has a calendar icon to its right. The 'first row as names' checkbox is checked. At the bottom of the dialog, there are two links: 'Show advanced parameters' and 'Change compatibility (8.1.001)'.

Parameter	Value
csv file	iris.csv
column separators	,
use quotes	<input checked="" type="checkbox"/>
quotes character	"
skip comments	<input type="checkbox"/>
parse numbers	<input checked="" type="checkbox"/>
decimal character	.
grouped digits	<input type="checkbox"/>
date format	Enter value...
first row as names	<input checked="" type="checkbox"/>

Example in RapidMiner



Parameters ✕

 **Set Role**

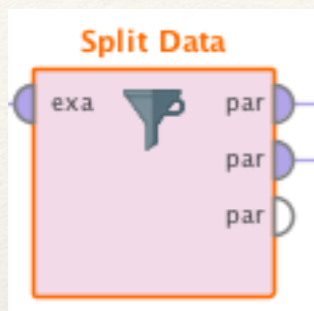
attribute name ⓘ

target role ⓘ

set additional roles ⓘ

✓ [Change compatibility \(8.1.001\)](#)

Example in RapidMiner




Parameters ×

Split Data





partitions Edit Enumeration (2)... ⓘ


sampling type automatic ⓘ

Edit Parameter List: partitions

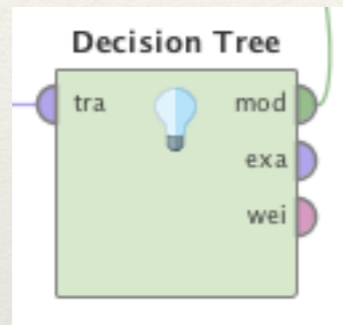
 **Edit Parameter List: partitions**
The partitions that should be created.

ratio
0.7
0.3

 **Add Entry**  **Remove Entry**  **OK**  **Cancel**

 [Show advanced parameters](#)

Example in RapidMiner

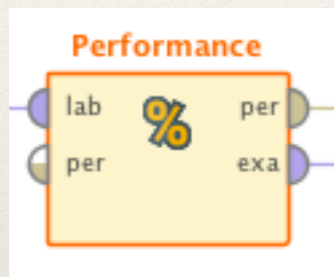


Parameters ✕

Decision Tree

criterion	information_gain ▼ ⓘ
maximal depth	5 ⓘ
<input checked="" type="checkbox"/> apply pruning	ⓘ
confidence	0.25 ⓘ
<input type="checkbox"/> apply prepruning	ⓘ

Example in RapidMiner



Parameters

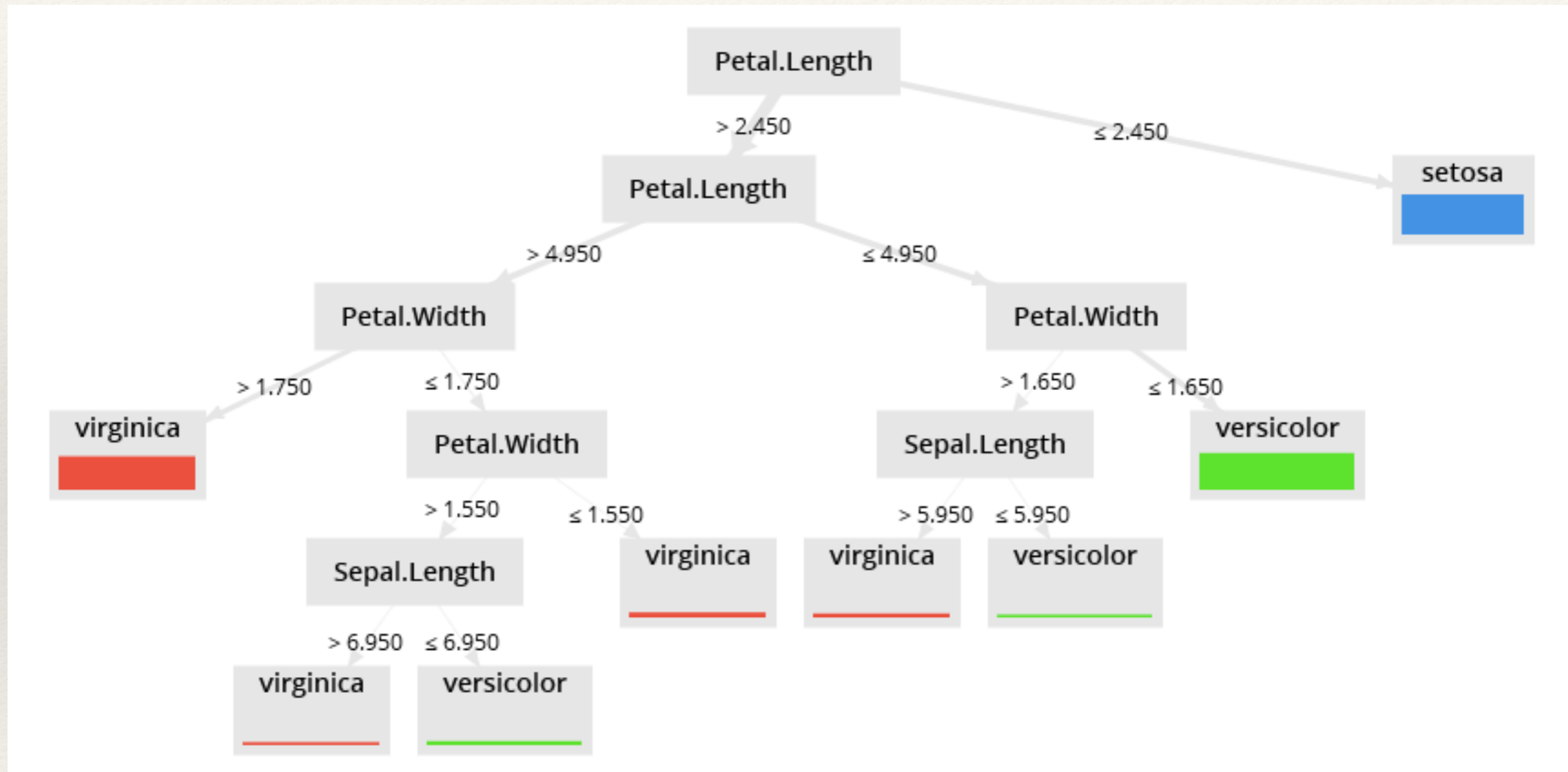
% Performance (Performance (Classification))

main criterion

- accuracy
- classification error
- kappa
- weighted mean recall
- weighted mean precision
- spearman rho
- kendall tau
- absolute error
- relative error
- relative error lenient
- relative error strict
- normalized absolute error
- root mean squared error

[Show advanced parameters](#)

Example in RapidMiner - Results



Example in RapidMiner - Results

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

accuracy: 95.56%

	true setosa	true versicolor	true virginica	class precision
pred. setosa	15	0	0	100.00%
pred. versicolor	0	15	2	88.24%
pred. virginica	0	0	13	100.00%
class recall	100.00%	100.00%	86.67%	

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall is also called sensitivity, particularly in medical applications