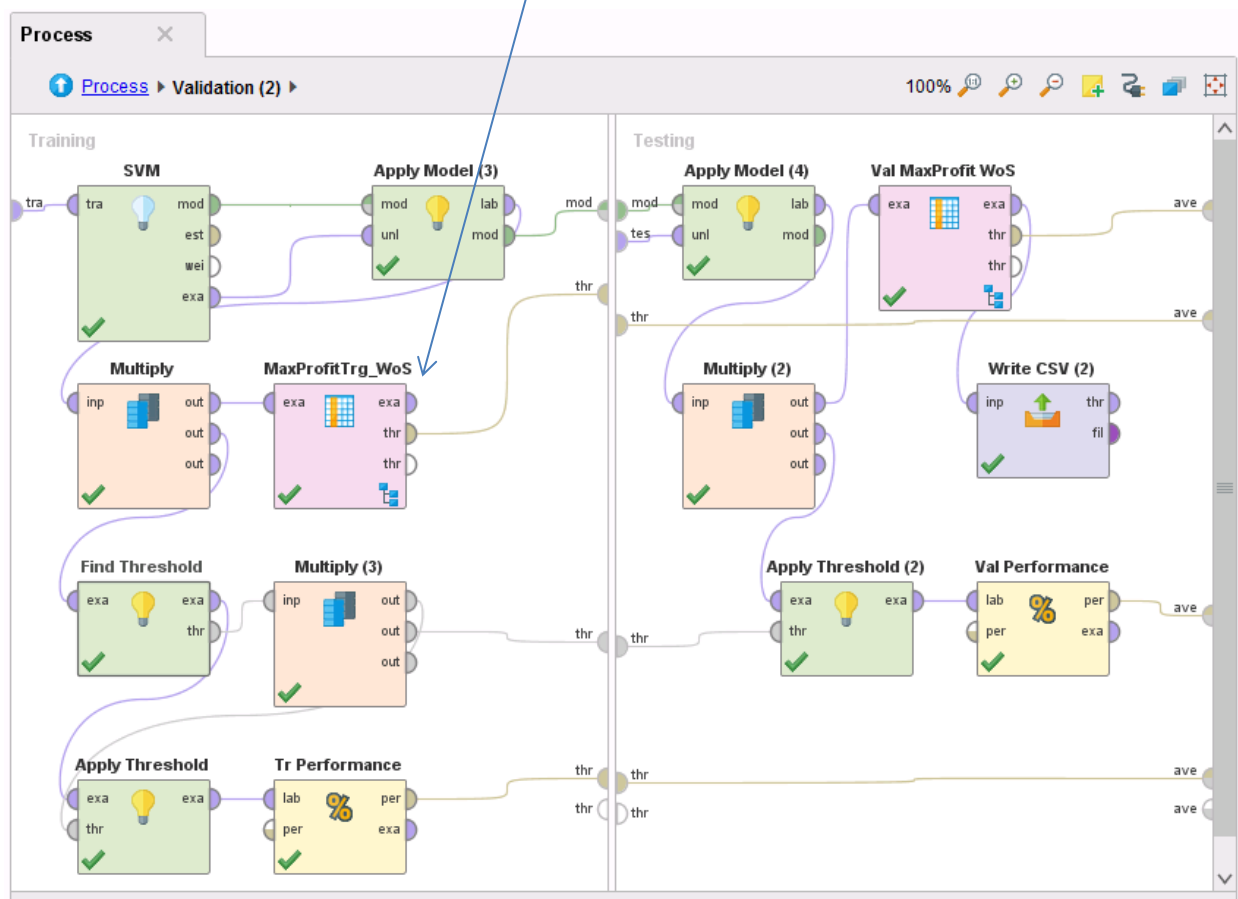


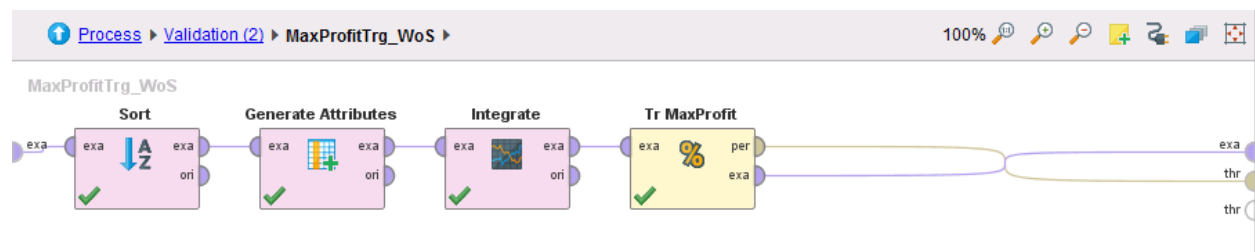
RapidMiner process – Assignment 3, maxProfit operator



The **maxProfit** operator helps to determine the maximum net profit for any model (without having to export the model scores to calculate this in Excel).


The calculation of max-profit can be conveniently packages inside a Operate on Subset operator. Steps are:

1. Sort the data on confidence(1) values
2. Generate a profit attribute. If the case is an actual donor, profit is \$12.32, else it is -0.68. (Need to use scaled values to reflect profit for original data with 5.1% donors. Values given in the process are not the accurate ones, you need to determine the scaled profit values to use).
3. Get the cumulative profit.
4. Get the maximum value of the cumulative profit.



The process above also shows a FindThreshold operator – to find the best threshold to minimize misclassification costs. This is based on costs for misclassifying a non-donor/donor. To get misclassification-cost values reflecting the 5.1 % donor rate in the original data, we should scale the \$12.32 and \$0.68 values for donors and non-donors.

Parameters ×

 **Find Threshold**

☒ *define labels* ⓘ

first label

0

ⓘ

second label

1

ⓘ

misclassification costs first

1.076

ⓘ

misclassification costs second

1.571

ⓘ