Calibrated Learner

Wraps another learner with probability calibration and decision threshold optimization.

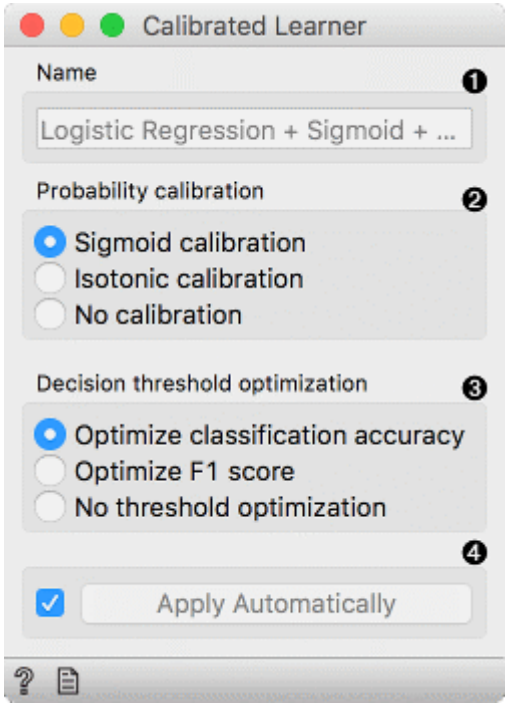
Inputs

- Data: input dataset
- Preprocessor: preprocessing method(s)
- Base Learner: learner to calibrate

Outputs

- Learner: calibrated learning algorithm
- Model: trained model using the calibrated learner

This learner produces a model that calibrates the distribution of class probabilities and optimizes decision threshold. The widget works only for binary classification tasks.



1. The name under which it will appear in other widgets. Default name is composed of the learner, calibration and optimization parameters.
2. Probability calibration:

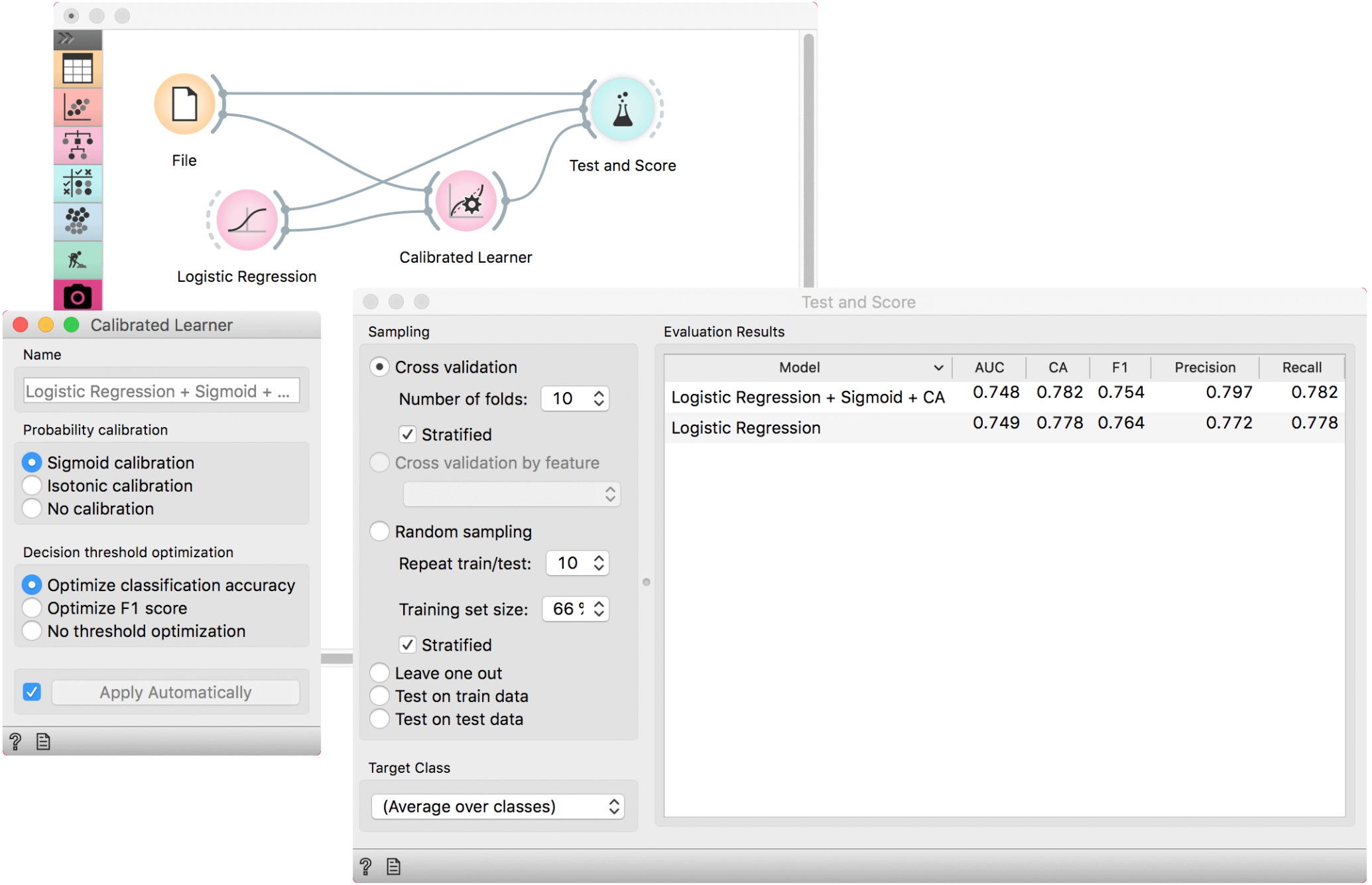
    - Sigmoid calibration
    - Isotonic calibration
    - No calibration
3. Decision threshold optimization:
    - Optimize classification accuracy
    - Optimize F1 score
    - No threshold optimization
4. Press *Apply* to commit changes. If *Apply Automatically* is ticked, changes are committed automatically.

Example

A simple example with **Calibrated Learner**. We are using the *titanic* data set as the widget requires binary class values (in this case they are ‘survived’ and ‘not survived’).

We will use Logistic Regression as the base learner which will we calibrate with the default settings, that is with sigmoid optimization of distribution values and by optimizing the CA.

Comparing the results with the uncalibrated **Logistic Regression** model we see that the calibrated model performs better.


File
Test and Score
Logistic Regression
Calibrated Learner
Calibrated Learner
Name
Logistic Regression + Sigmoid + ...
Probability calibration
Sigmoid calibration
Isotonic calibration
No calibration
Decision threshold optimization
Optimize classification accuracy
Optimize F1 score
No threshold optimization
Apply Automatically
Test and Score
Sampling
Cross validation
Number of folds: 10
Stratified
Cross validation by feature
Random sampling
Repeat train/test: 10
Training set size: 66 %
Stratified
Leave one out
Test on train data
Test on test data
Target Class
(Average over classes)
Evaluation Results
Model AUC CA F1 Precision Recall
Logistic Regression + Sigmoid + CA 0.748 0.782 0.754 0.797 0.782
Logistic Regression 0.749 0.778 0.764 0.772 0.778