Apply Domain

Given dataset and template transforms the dataset.
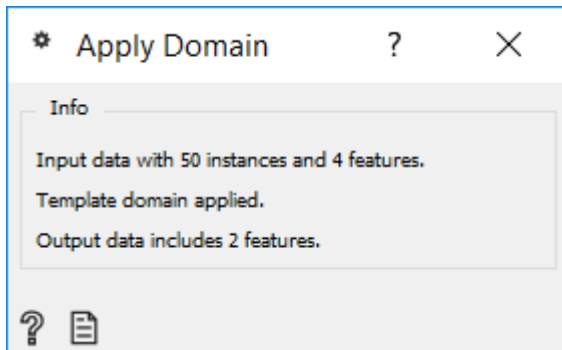
Inputs

- Data: input dataset
- Template Data: template for transforming the dataset

Outputs

- Transformed Data: transformed dataset

Apply Domain maps new data into a transformed space. For example, if we transform some data with PCA and wish to observe new data in the same space, we can use Apply Domain to map the new data into the PCA space created from the original data.

The widget receives a dataset and a template dataset used to transform the dataset.

Example

We will use iris data from the File widget for this example. To create two separate data sets, we will use Select Rows and set the condition to *iris is one of iris-setosa, iris-versicolor*. This will output a data set with a 100 rows, half of them belonging to iris-setosa class and the other half to iris-versicolor.

We will transform the data with PCA and select the first two components, which explain 96% of variance. Now, we would like to apply the same preprocessing on the 'new' data, that is the remaining 50 iris virginicas. Send the unused data from **Select Rows** to **Apply Domain**. Make sure to use the *Unmatched Data* output from **Select Rows** widget. Then add the *Transformed data* output from **PCA**.

Apply Domain will apply the preprocessor to the new data and output it. To add the new data to the old data, use Concatenate. Use *Transformed Data* output from **PCA** as *Primary Data* and *Transformed Data* from **Apply Domain** as *Additional Data*.

Observe the results in a Data Table or in a Scatter Plot to see the new data in relation to the old one.

Transform

Info

Input data with 50 instances and 4 features.
Preprocessor PCA applied.
Output data includes 2 features.

Matching Data → Data

Transformed data → Primary Data

PCA

Preprocessor

Transformed Data → Additional Data

Concatenate

Data

Data Table

Data

Scatter Plot

File

Data

Select Rows

Unmatched Data → Data

Transform

PCA

Components Selection

Components: 2

Variance covered: 96%

0.963

0.204

Decomposition

○ PCA
○ Truncated SVD

Options

☑ Normalize data
Show only first 20

☑ Apply automatically

Apply

Proportion of variance

Principal Components

Scatter Plot

Axis x: [N] PC1

Axis y: [N] PC2

Find Informative Projections

Color: [C] iris

Shape: (Same shape)

Size: (Same size)

Label: (No labels)

☐ Label only selected points

Symbol size:

Opacity:

Jittering:

☐ Jitter numeric values

☐ Show color regions
☑ Show legend
☐ Show gridlines
☑ Show all data on mouse hover
☐ Show regression line

Zoom/Select

☑ Send Automatically

PC2

PC1

● Iris-setosa
● Iris-versicolor
● Iris-virginica