

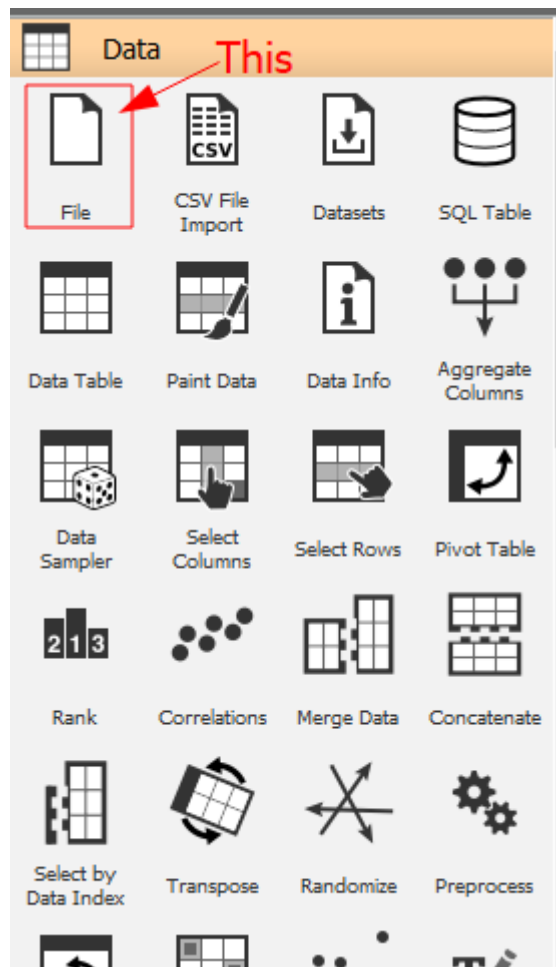
Orange Practicals

Practical 1

Aim: Classification using orange.

Procedure:

- Drag and drop a File widget from the Data section found in the left panel to the workspace.





File

- Double click the File widget and choose the **zoo.tab** dataset.

Source

☒ File: zoo.tab ... Reload

☐ URL:

Info

Zoo dataset
This dataset consists of 101 animals with various traits to describe them.

101 instance(s)
16 feature(s) (no missing values)
Classification; categorical class with 7 values (no missing values)
1 meta attribute(s)

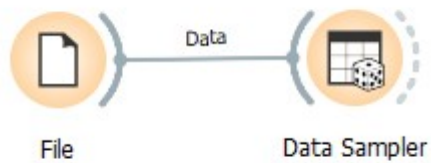
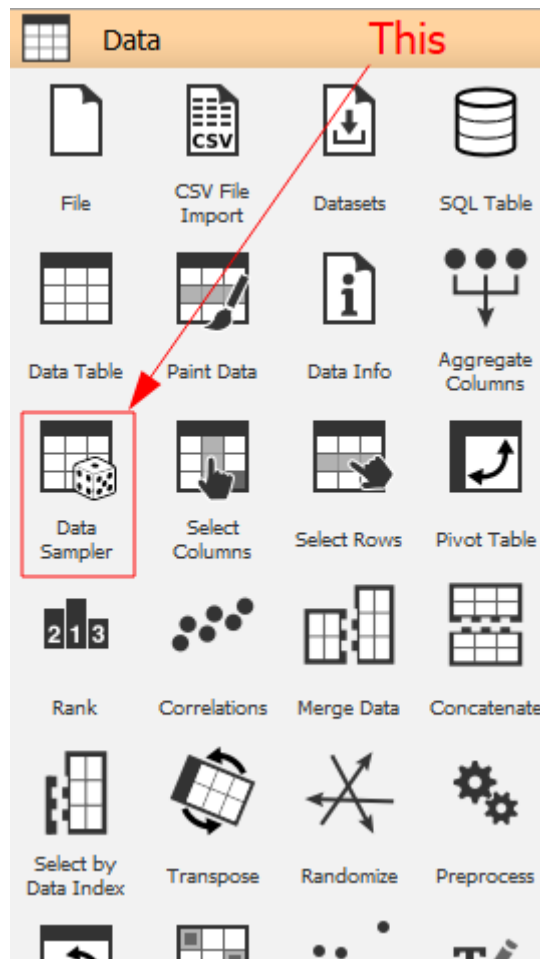
Columns (Double click to edit)

	Name	Type	Role	Values
1	hair	C categorical	feature	0, 1
2	feathers	C categorical	feature	0, 1
3	eggs	C categorical	feature	0, 1
4	milk	C categorical	feature	0, 1
5	airborne	C categorical	feature	0, 1
6	aquatic	C categorical	feature	0, 1

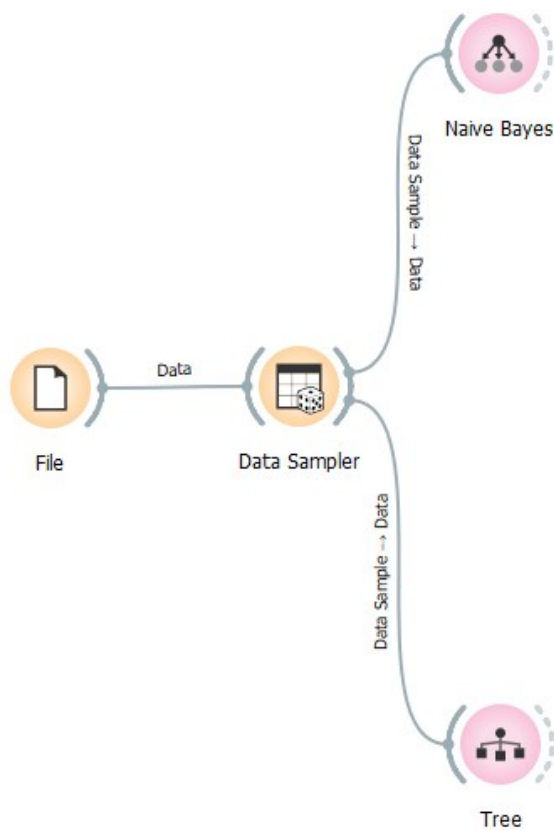
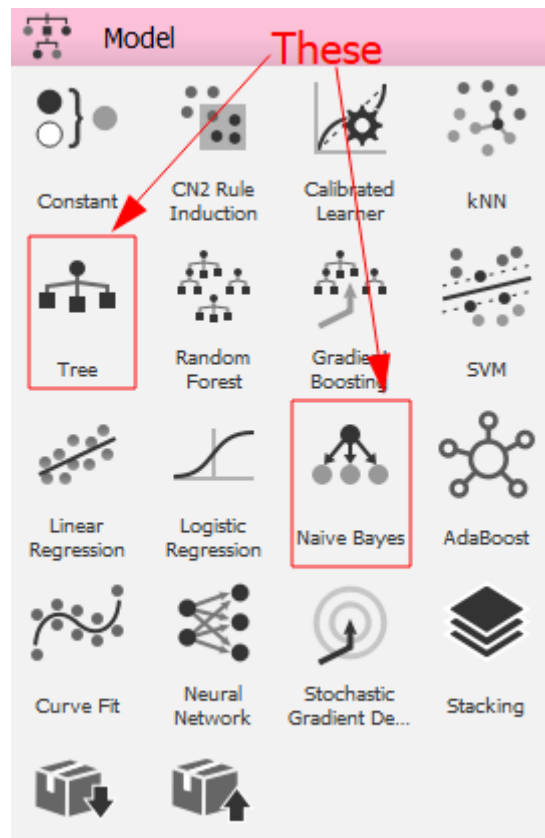
Reset Apply

Browse documentation datasets

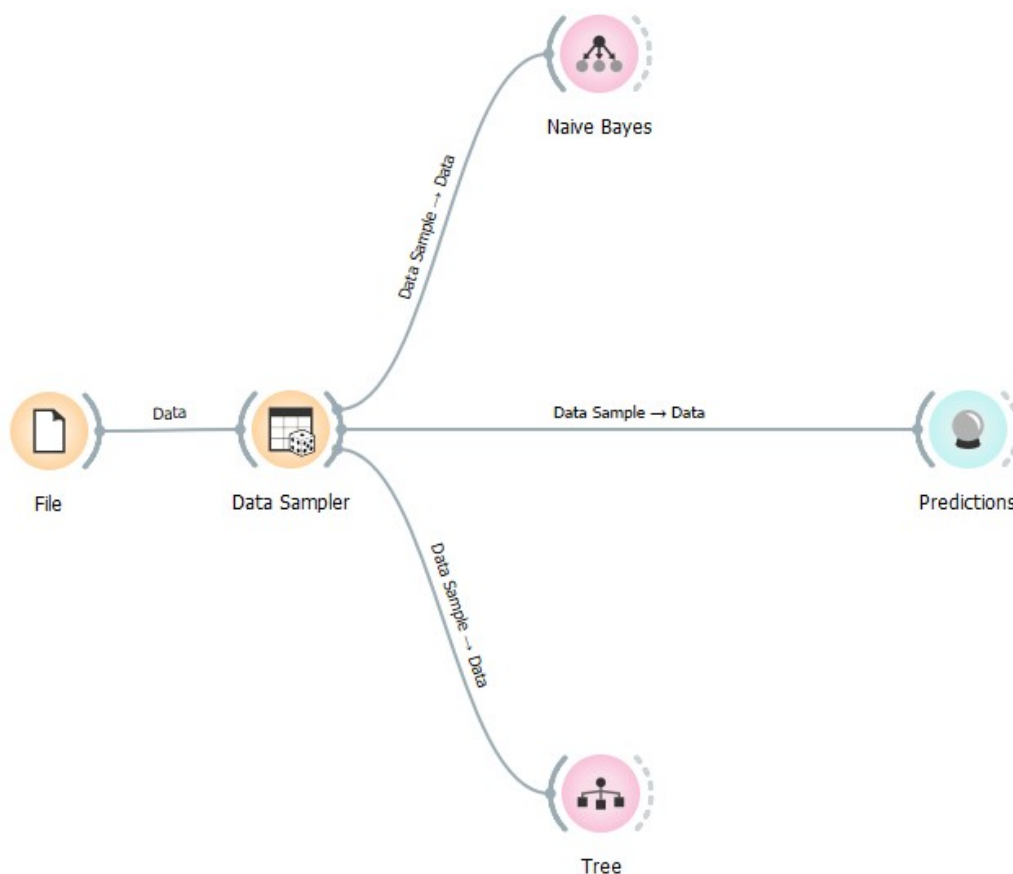
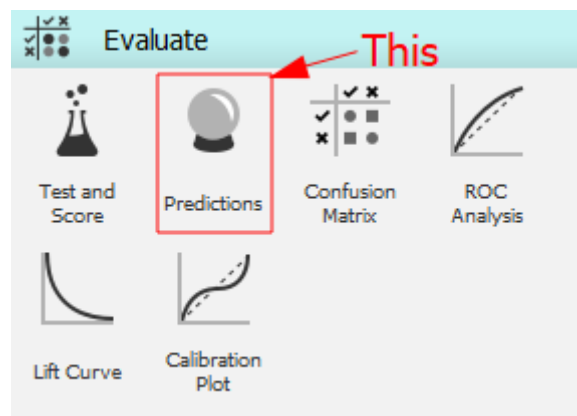
- Drag and drop a Data Sampler widget onto the workspace and connect it to the File widget. This widget is found in the Data section.



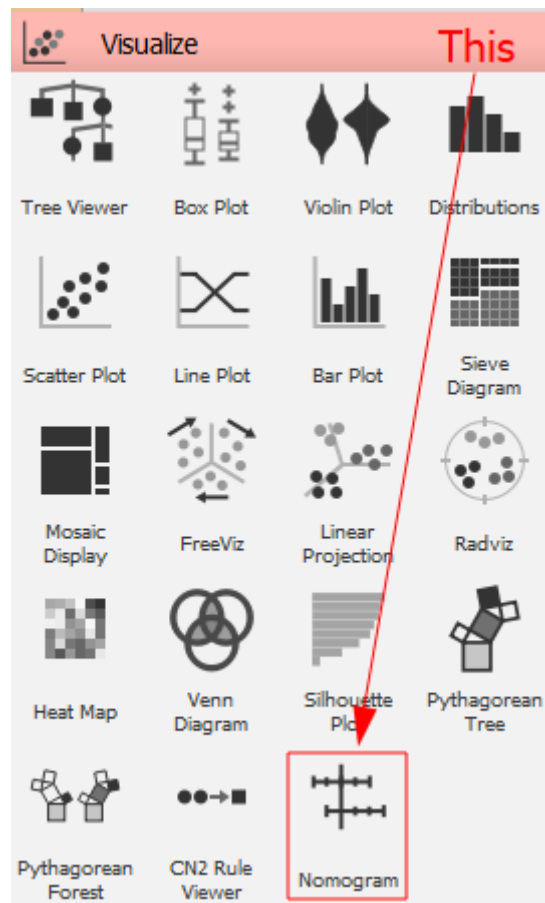
- We choose the Naïve Bayes' and Classification Tree models to classify our data. Drag and drop these widgets to the workspace. They are found under the Models section.

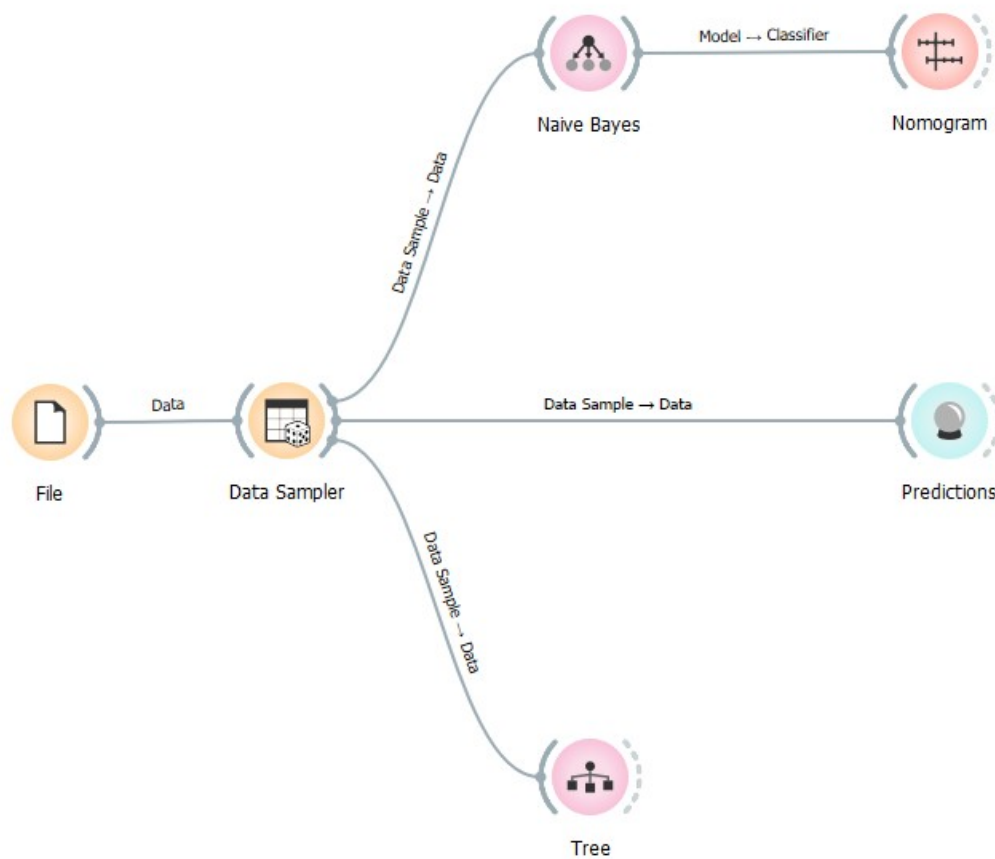


- To perform predictions on the data, we drag and drop the Predictions widget onto the workspace. This widget is found in the Evaluate section.

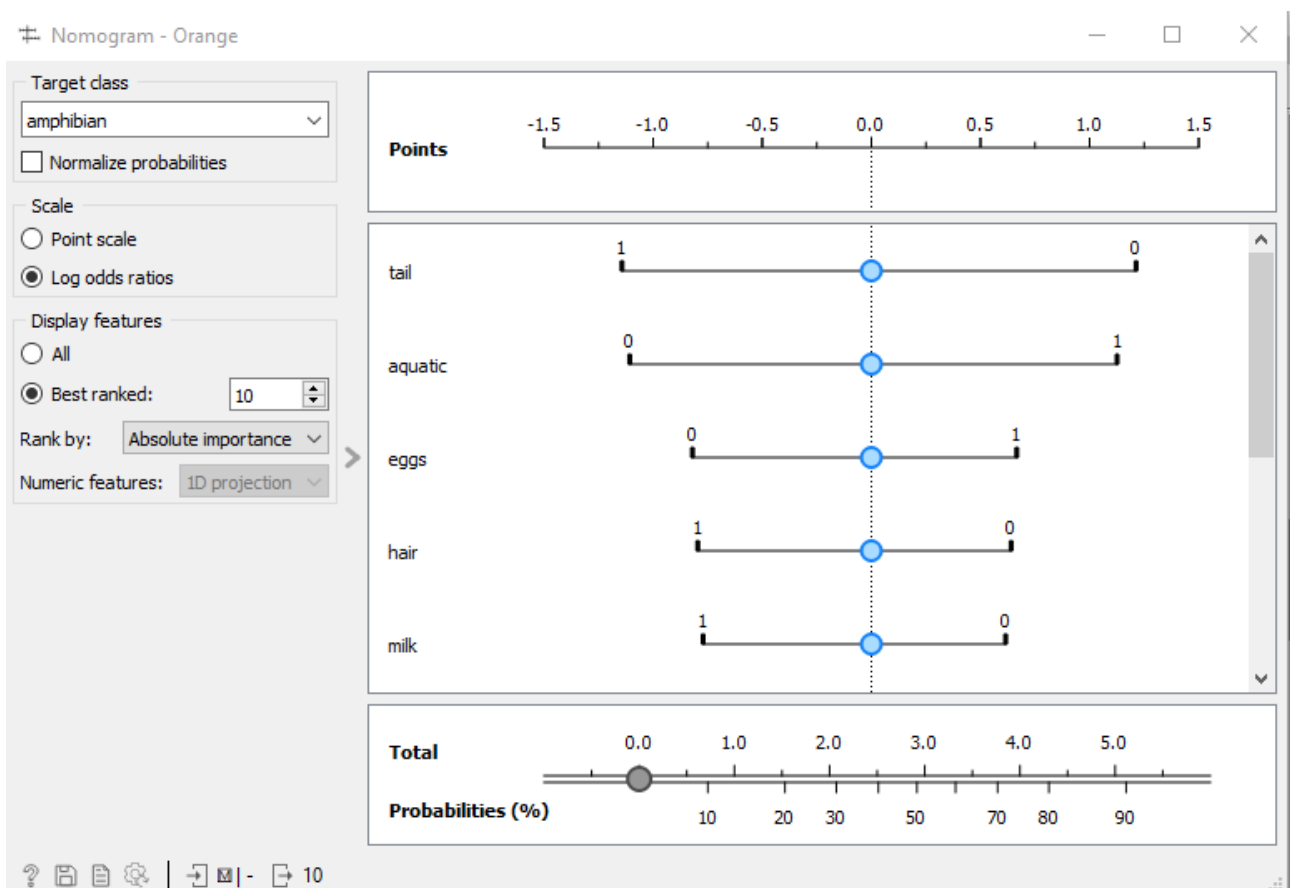
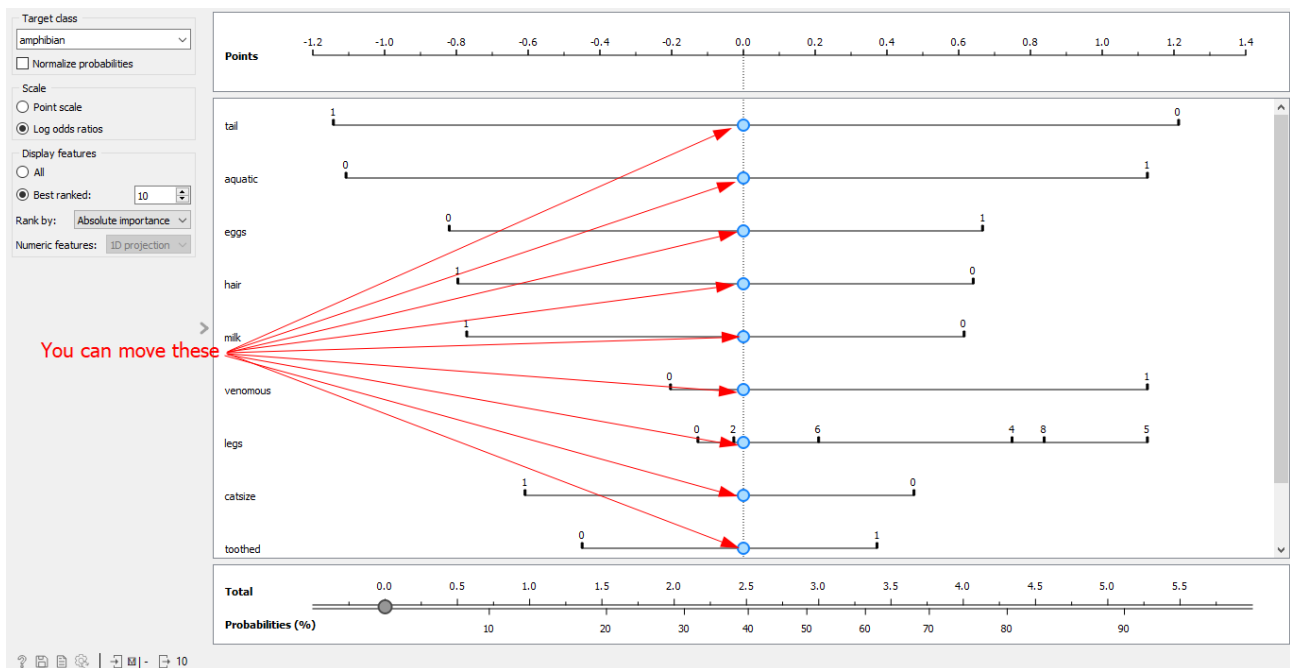


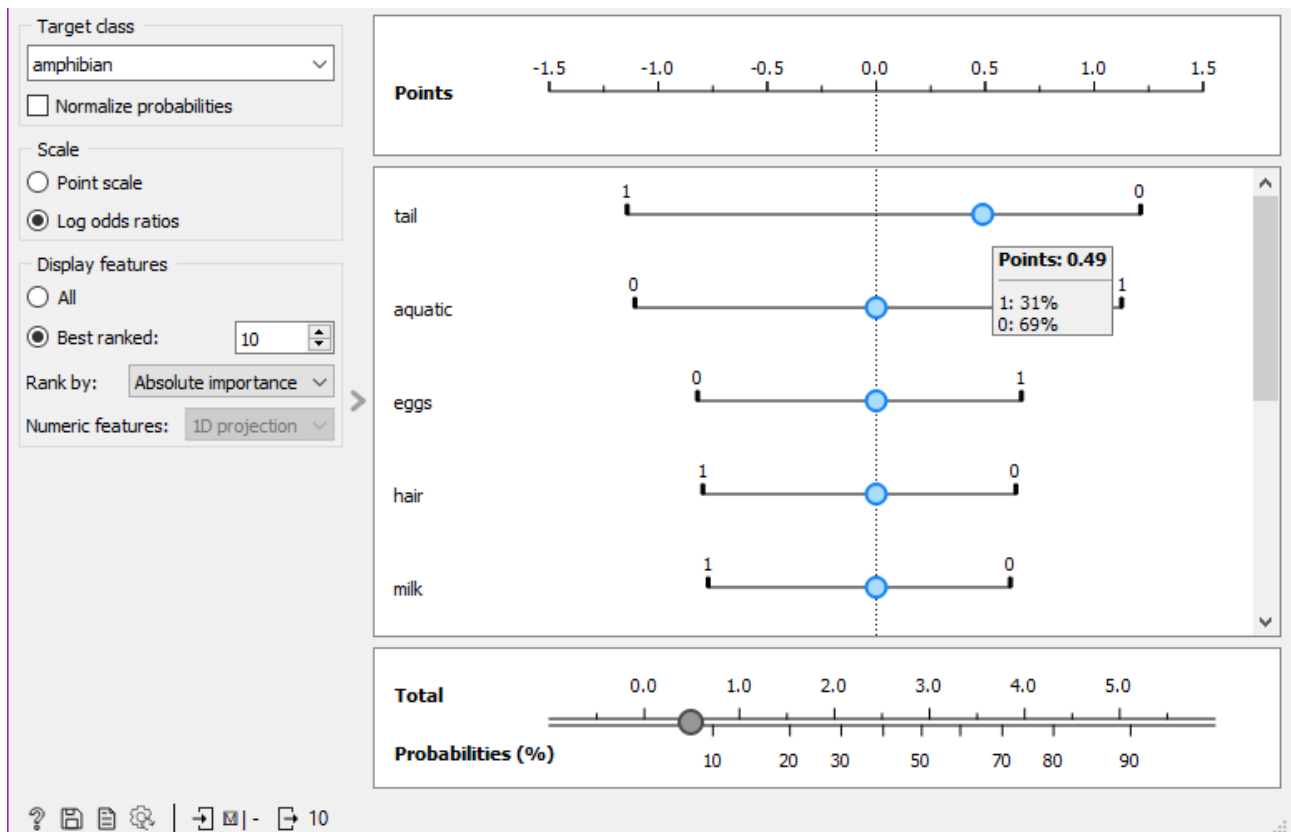
- A Nomogram is useful to view data from a Naïve Bayes' model. Drag and drop this widget onto the workspace. It can be found in the Visualise section.



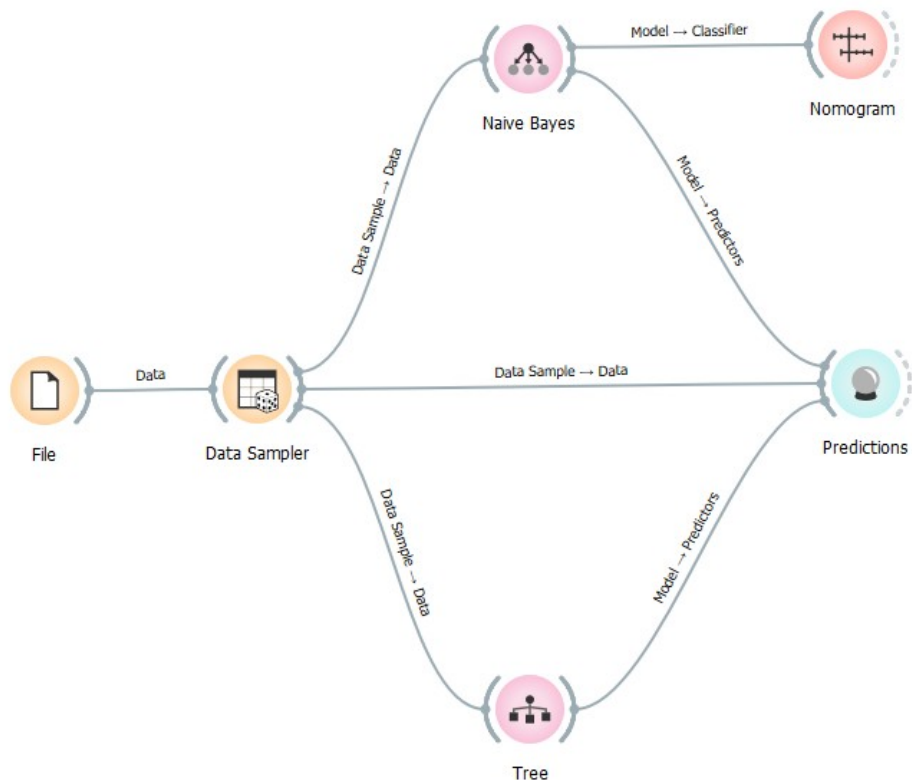


- You can move the points in the Nomogram to see the probabilities of a particular class. Here 1 indicates favourable probability while 0 indicates unfavourable probability.





- Connect the Naive Bayes' widget and Tree widget to the Predictions widget to perform predictions.

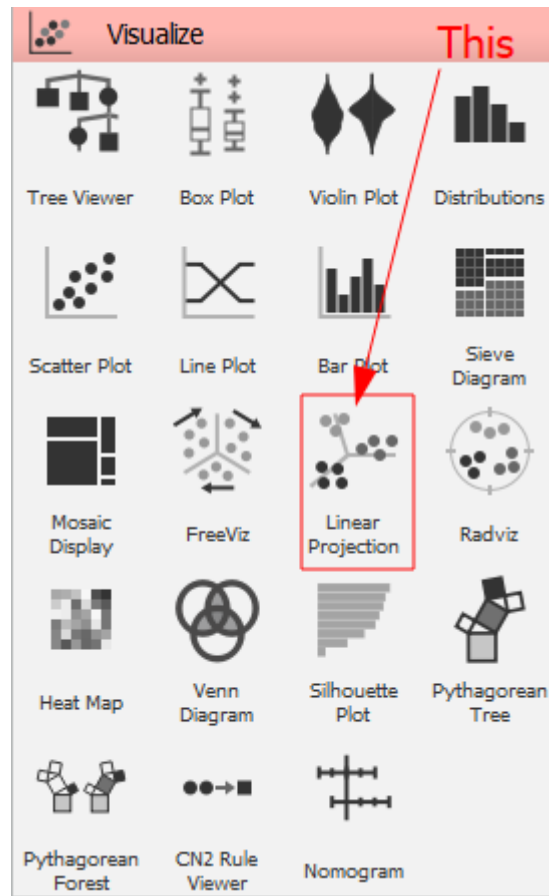


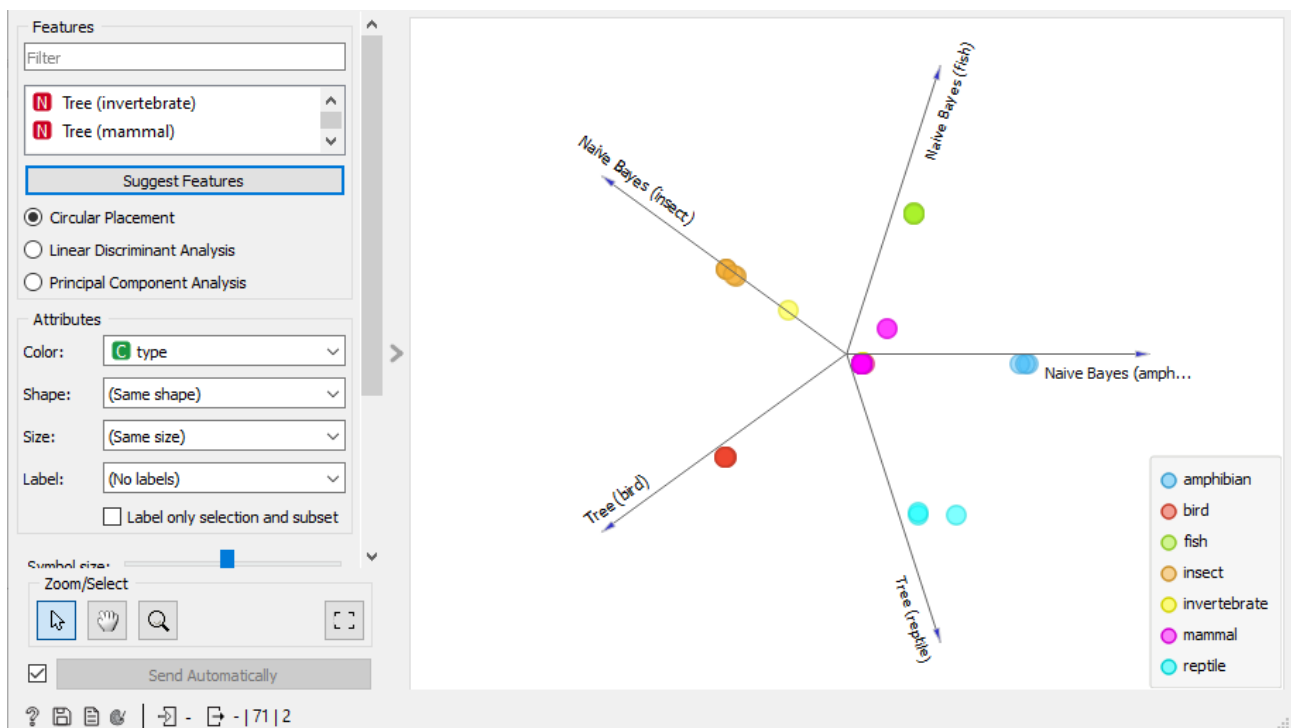
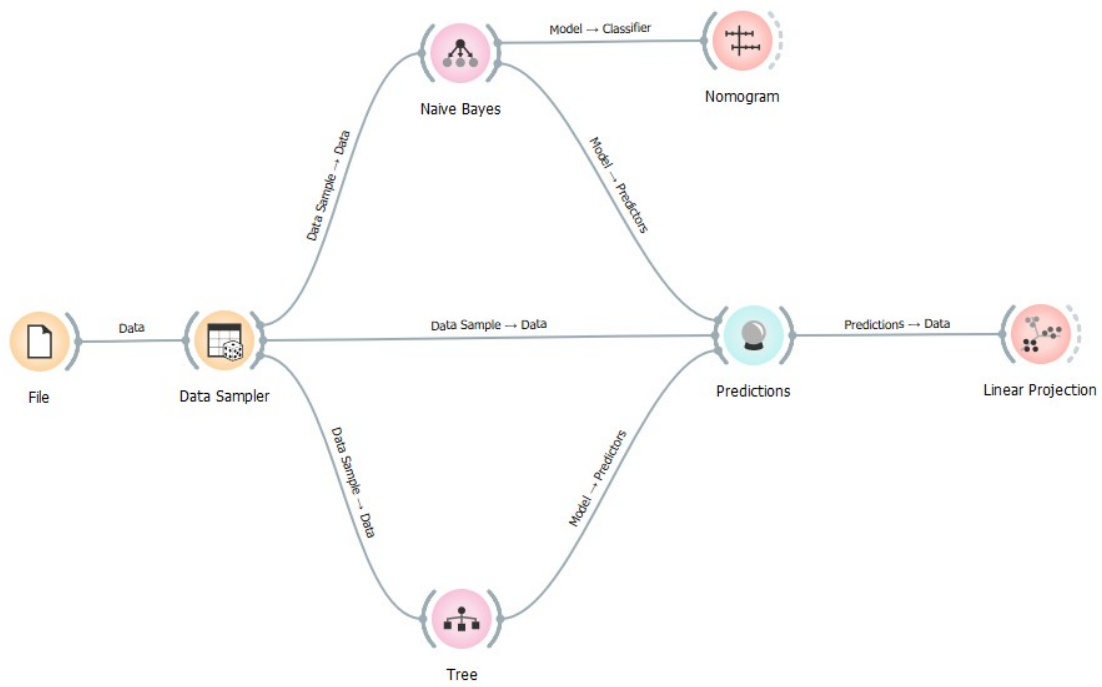
➤ Double click the Predictions widget to see the predictions.

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	1.000	0.944	0.948	0.967	0.944
Tree	0.999	0.986	0.986	0.988	0.986

type	name	hair	feathers	eggs	milk
mammal	squirrel	1	0	0	1
mammal	oryx	1	0	0	1
mammal	porpoise	0	0	0	1
mammal	puma	1	0	0	1
mammal	lion	1	0	0	1
mammal	honeybee	1	0	1	0
mammal	elephant	1	0	0	1
mammal	leopard	1	0	0	1
mammal	cheetah	1	0	0	1
mammal	aardvark	1	0	0	1
fish	dogfish	0	0	1	0
insect	gnat	0	0	1	0
insect	wasp	1	0	1	0
bird	gull	0	1	1	0
invertebrate	seawasp	0	0	1	0
mammal	boar	1	0	0	1
mammal	vampire	1	0	0	1
bird	skimmer	0	1	1	0
fish	chub	0	0	1	0
mammal	goat	1	0	0	1
reptile	seasnake	0	0	0	0
amphibian	toad	0	0	1	0
amphibian	frog	0	0	1	0
insect	fishbird	0	0	1	0

- We will use the Linear Projection widget to visualize the data. Drag and drop this widget from the Visualise section.



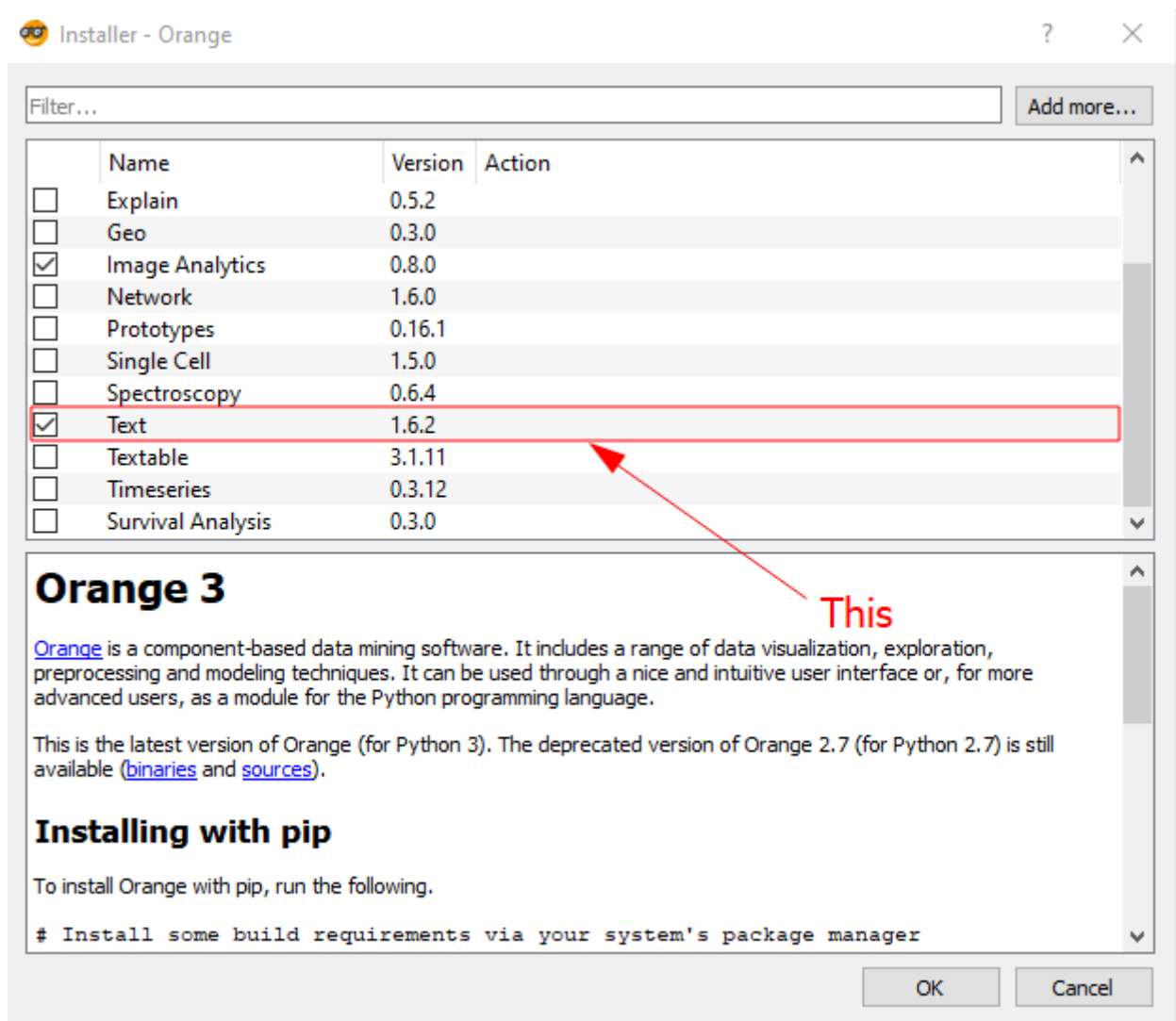


Practical 2

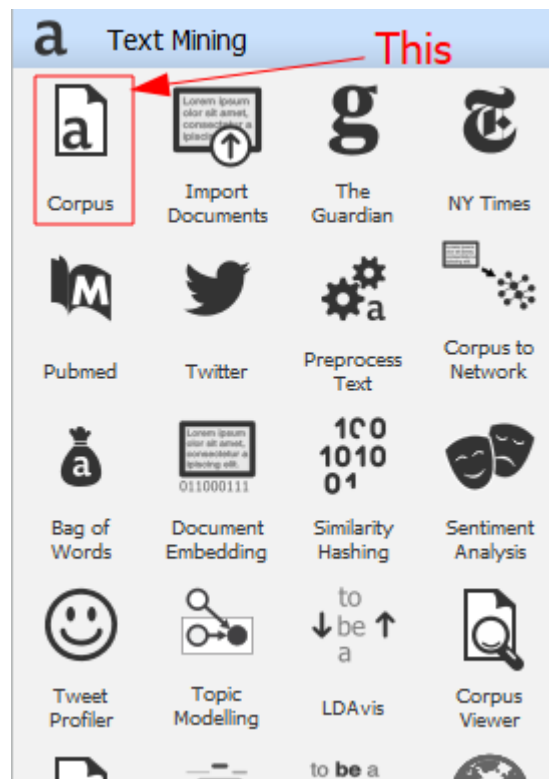
Aim: Text classification using Orange to determine whether the story has magic or is related to animals

Procedure:

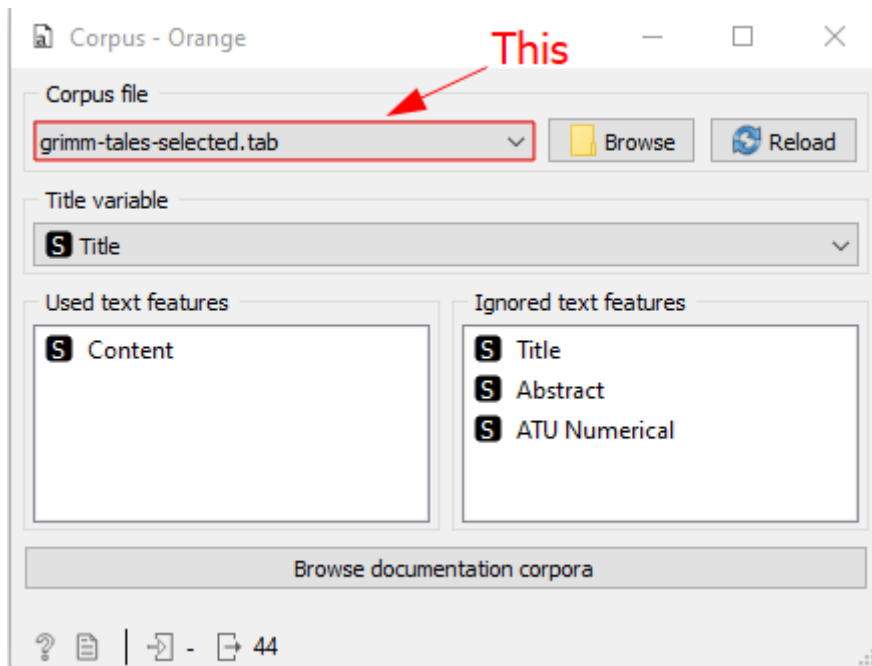
- Before starting, we need to install the **Text** add-on. Install it by navigating to **Options > Add ons...**. You will be prompted to restart Orange after the installation completes. If it is already installed, skip this step.



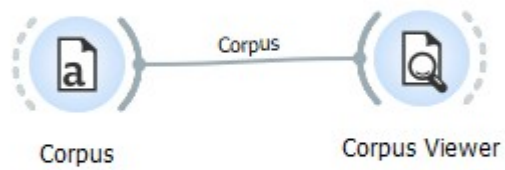
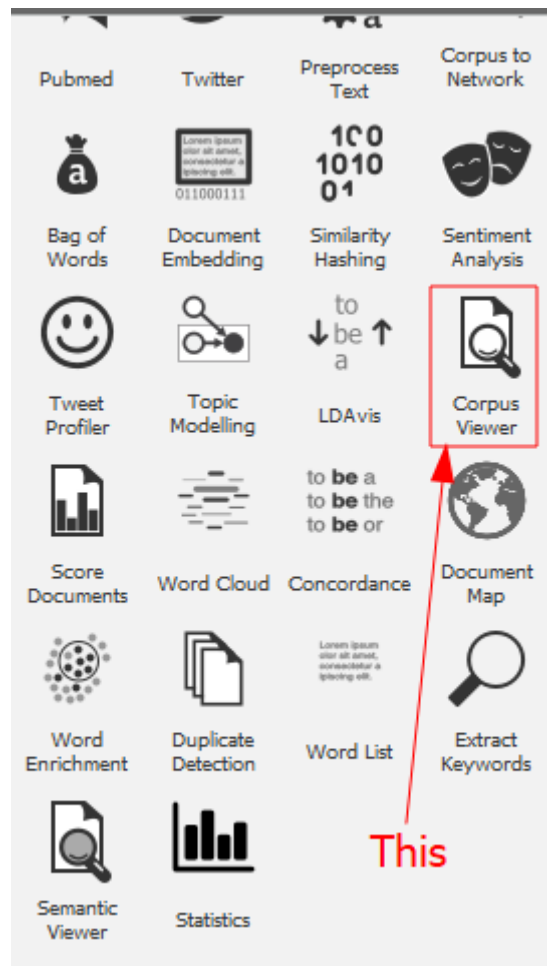
- Drag and drop a Corpus widget to the workspace. It can be found in the newly added Text Mining section.



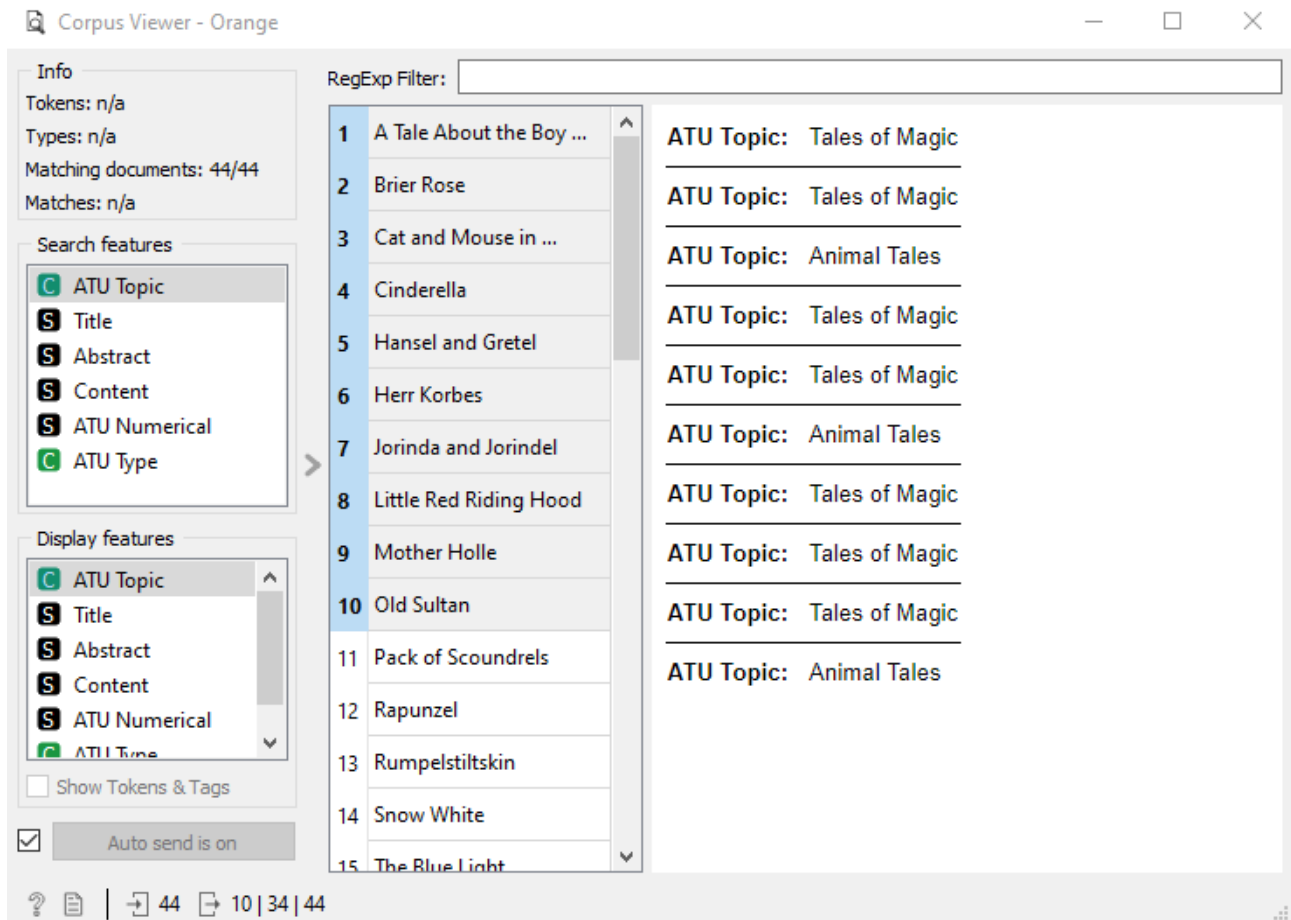
- Double click the Corpus widget and select the `grimm-tales-selected.tab` corpus file.



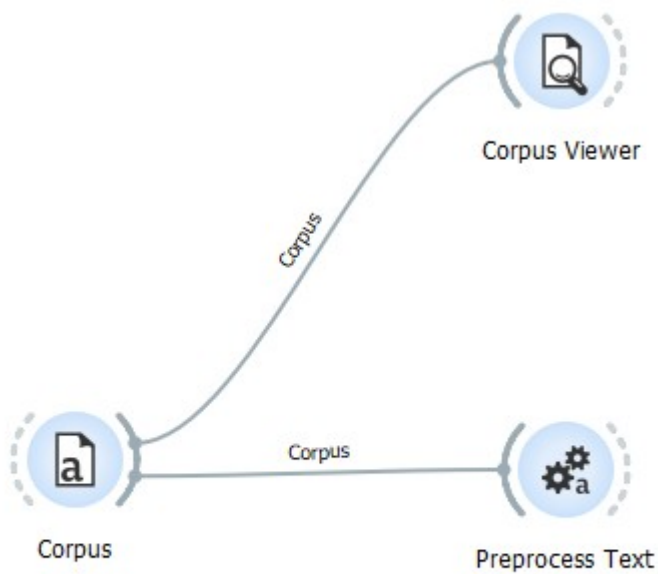
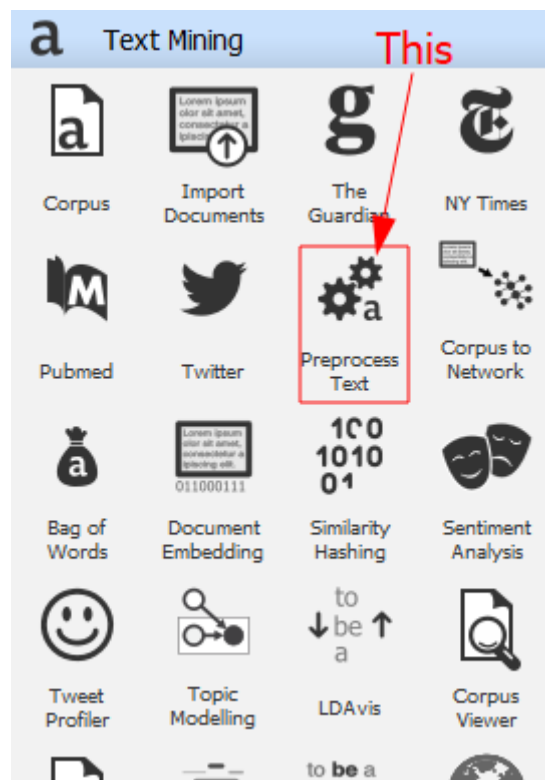
- Drag and drop a Corpus Viewer widget and connect it to the Corpus widget. The Corpus Viewer widget can be found in the Text Mining section.



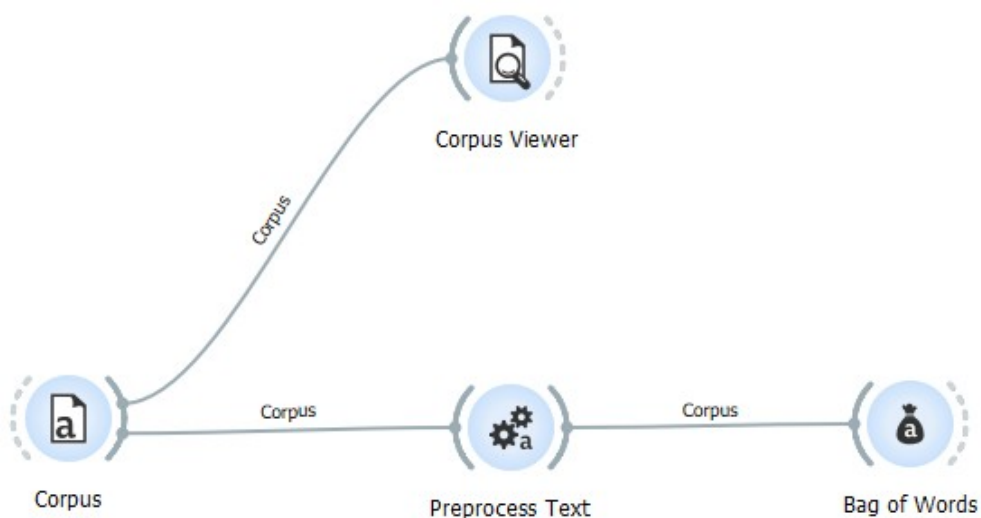
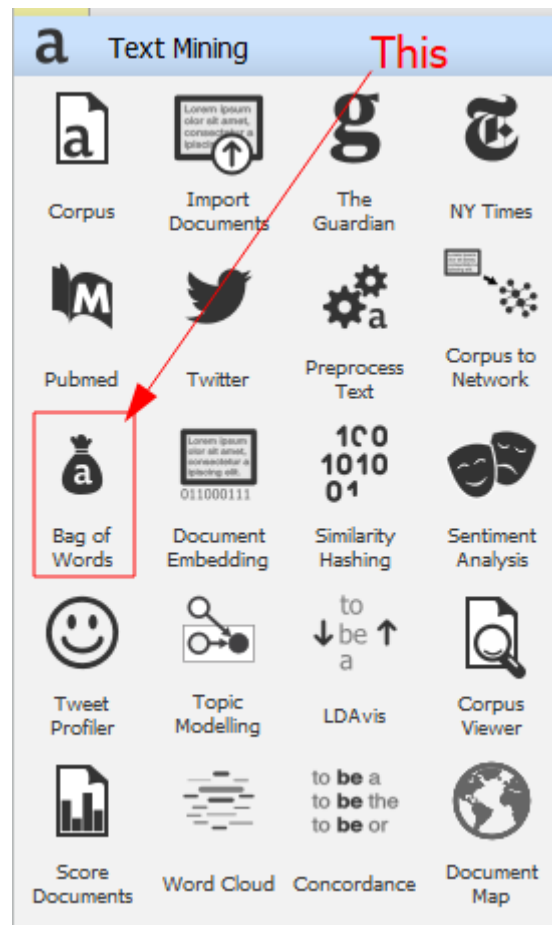
- Double click the **Corpus Viewer** widget to visualize the corpus in a tabular format. Select the first 10 entries to use as training set.



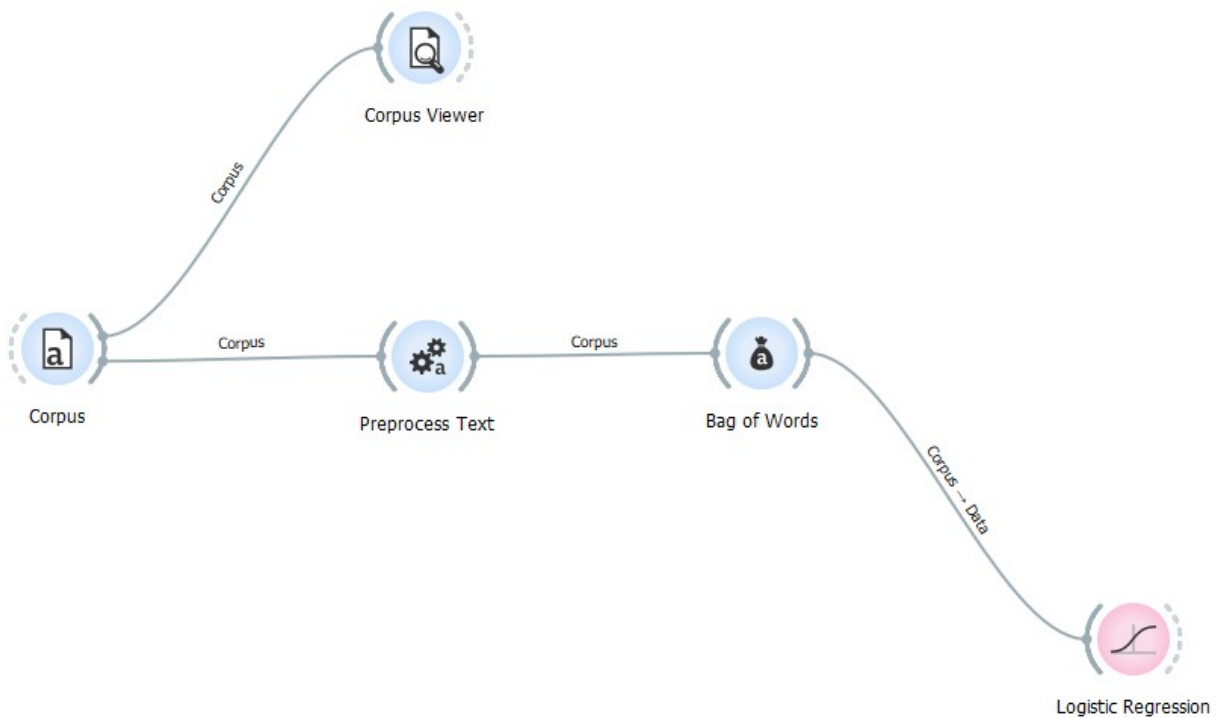
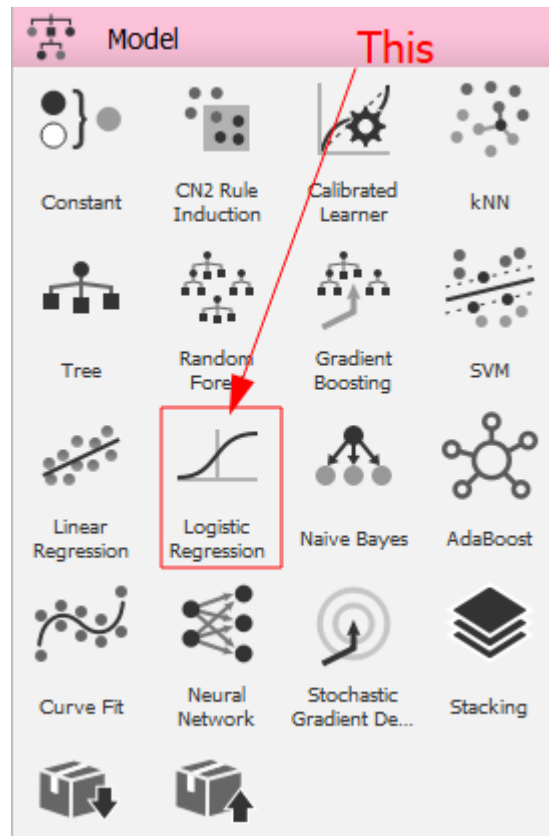
- Add a **Preprocess Text** widget to the workspace. This widget can also be found in the **Text Mining** section.



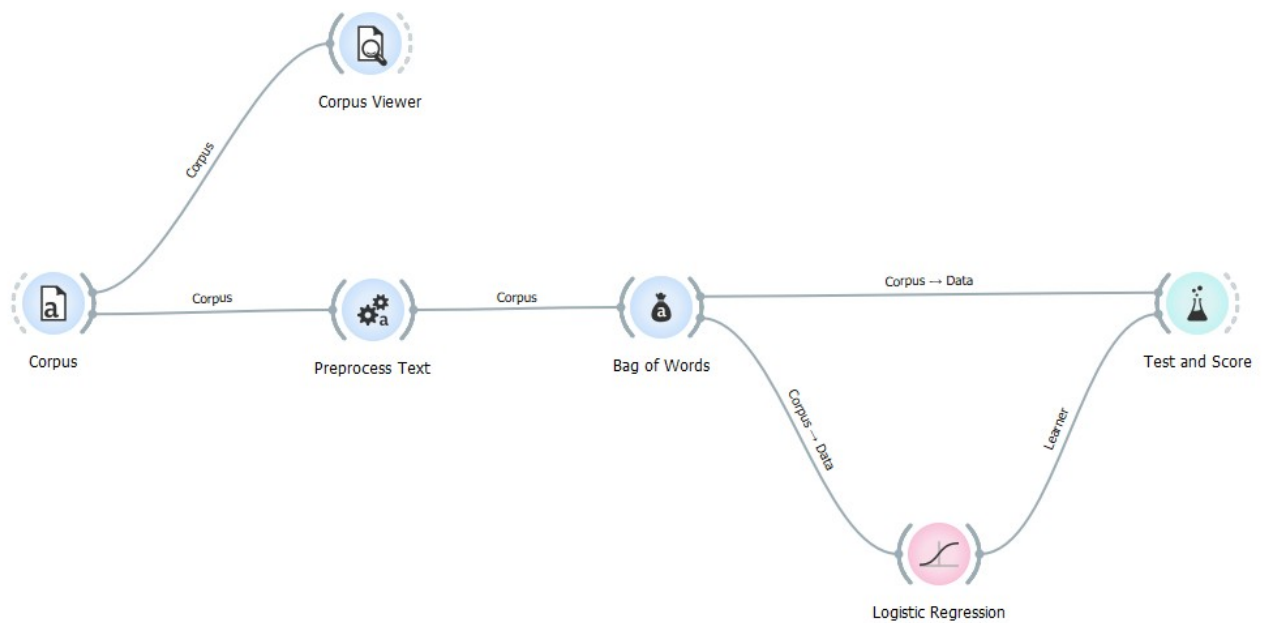
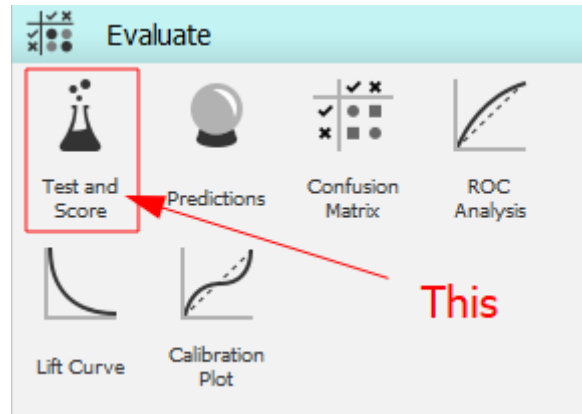
- Add a Bag Of Words widget to the workspace and connect it to the Preprocess Text widget. This widget is also found in the Text Mining section.



- We will use **Logistic Regression** to model our data. Drag and drop this widget from the **Model** section onto the workspace.



- We will use the Test and Score widget to check our model. Drag and drop this widget from the Evaluate section onto the workspace and connect it to the Logistic Regression model and Bag of Words widget.



- Double-click the Test and Score widget after it finishes processing. It will provide you with data such as the accuracy of the model etc.

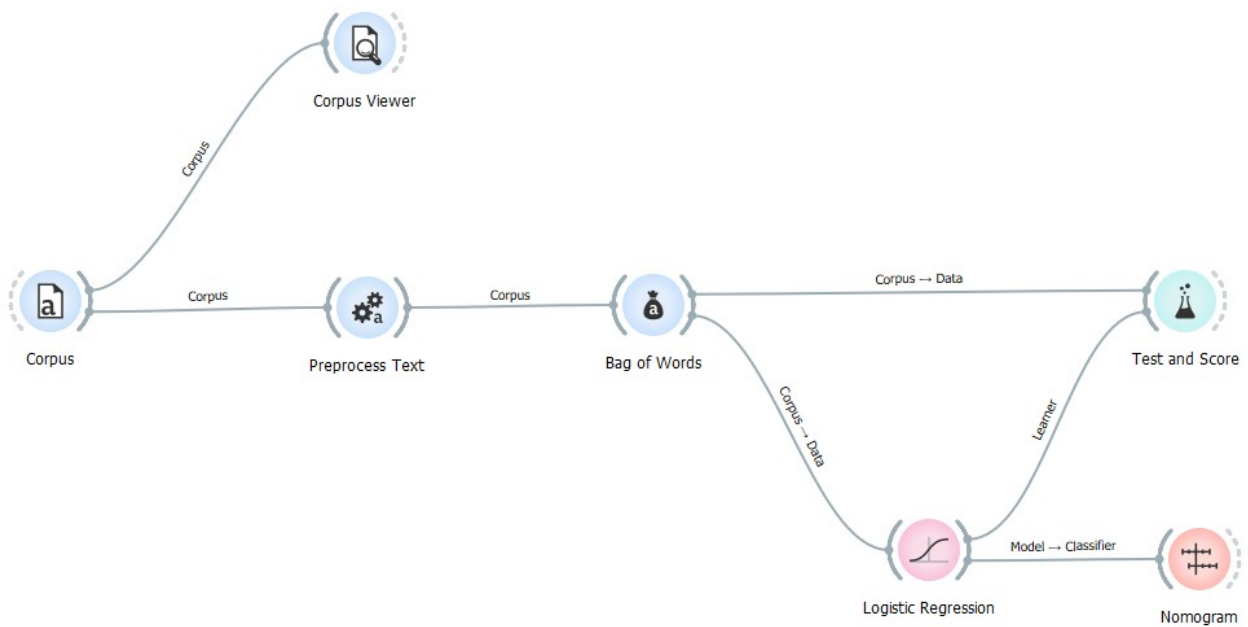
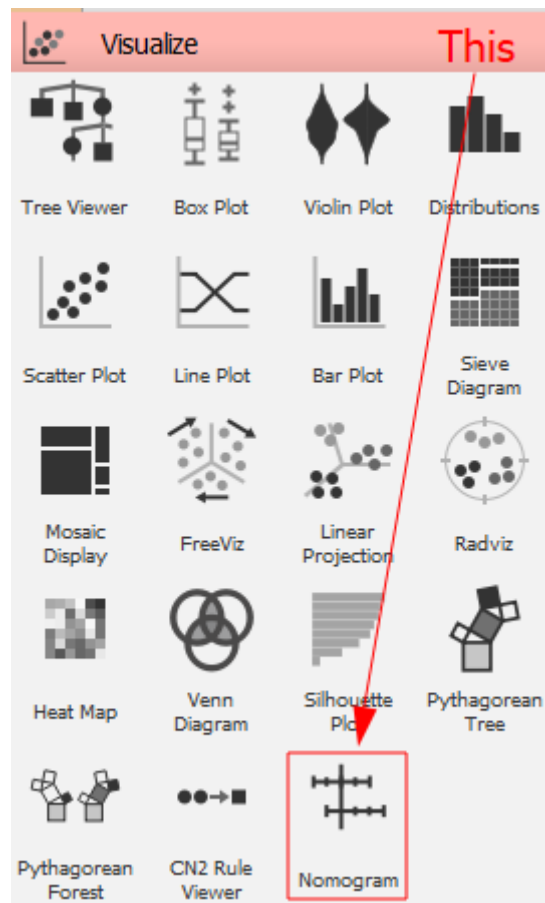
The screenshot displays the Orange3 software interface. On the left, the 'Sampling' section is active, showing 'Cross validation' selected with 10 folds and stratified sampling. The 'Target Class' is set to '(Average over classes)'. The 'Model Comparison' section shows 'Area under ROC curve' as the metric, with a 'Negligible difference' threshold of 0.1. The 'Evaluation Results' table on the right shows the performance of the Logistic Regression model. Below this, the 'Model Comparison by AUC' section shows a comparison between the Logistic Regression model and itself, with a probability of 1.0 for the model to be higher than itself.

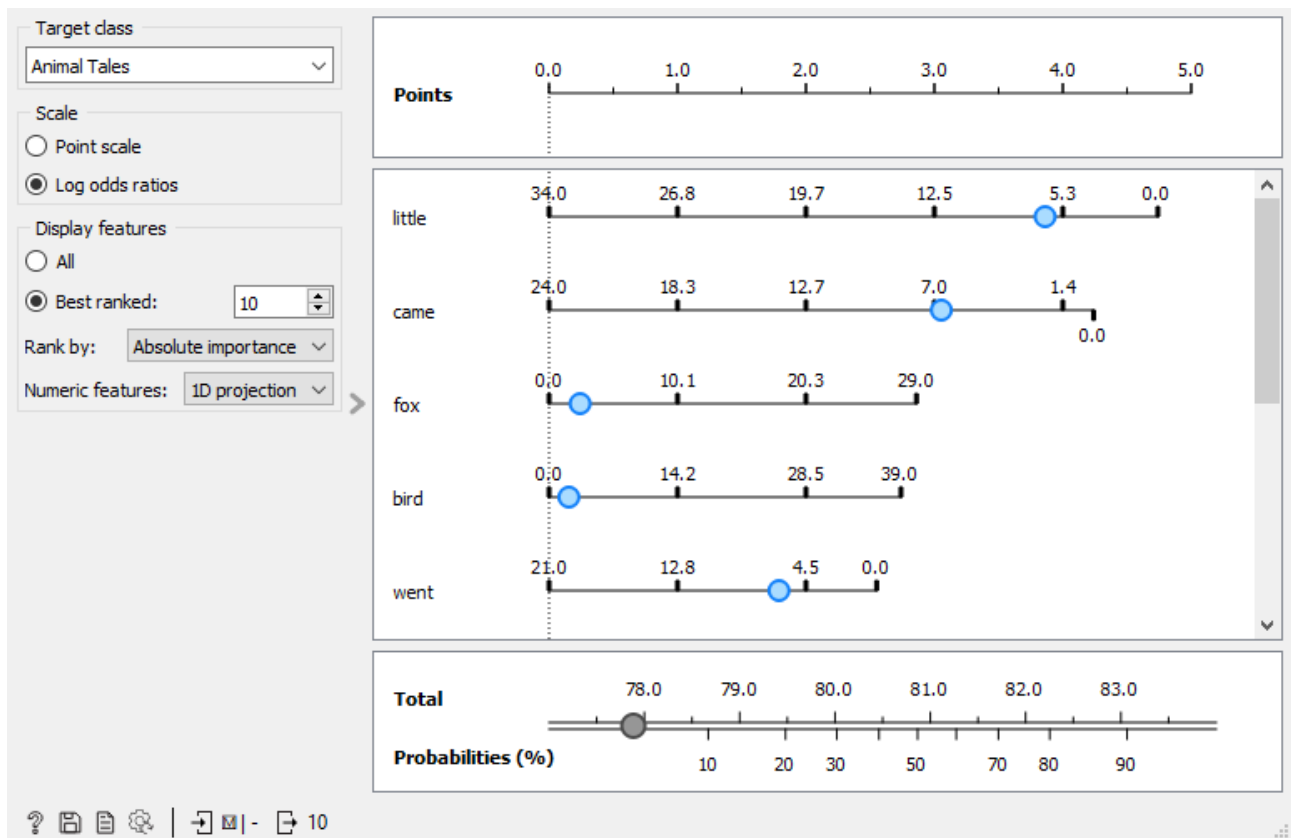
Model	AUC	CA	F1	Precision	Recall
Logistic Regression	0.968	0.909	0.910	0.926	0.909

Model	Logistic ...
Logistic Regression	1.0

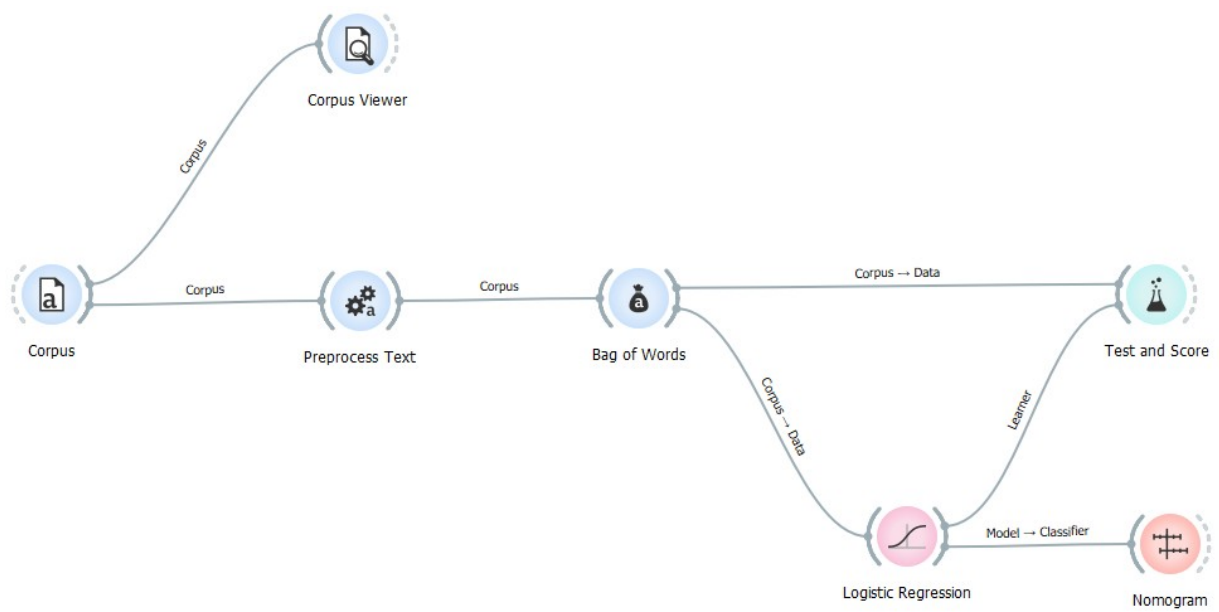
Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

- We will use a Nomogram to visualise our model. Drag and drop it from the Visualise section onto the workspace and connect it to the Linear Regression model.





- To see whether the model works as intended, we create a new Corpus widget and set the file to `andersen.tab`.



Corpus (1)

Corpus file

andersen.tab

Browse

Reload

Title variable

Title

Used text features

Content

Ignored text features

Title

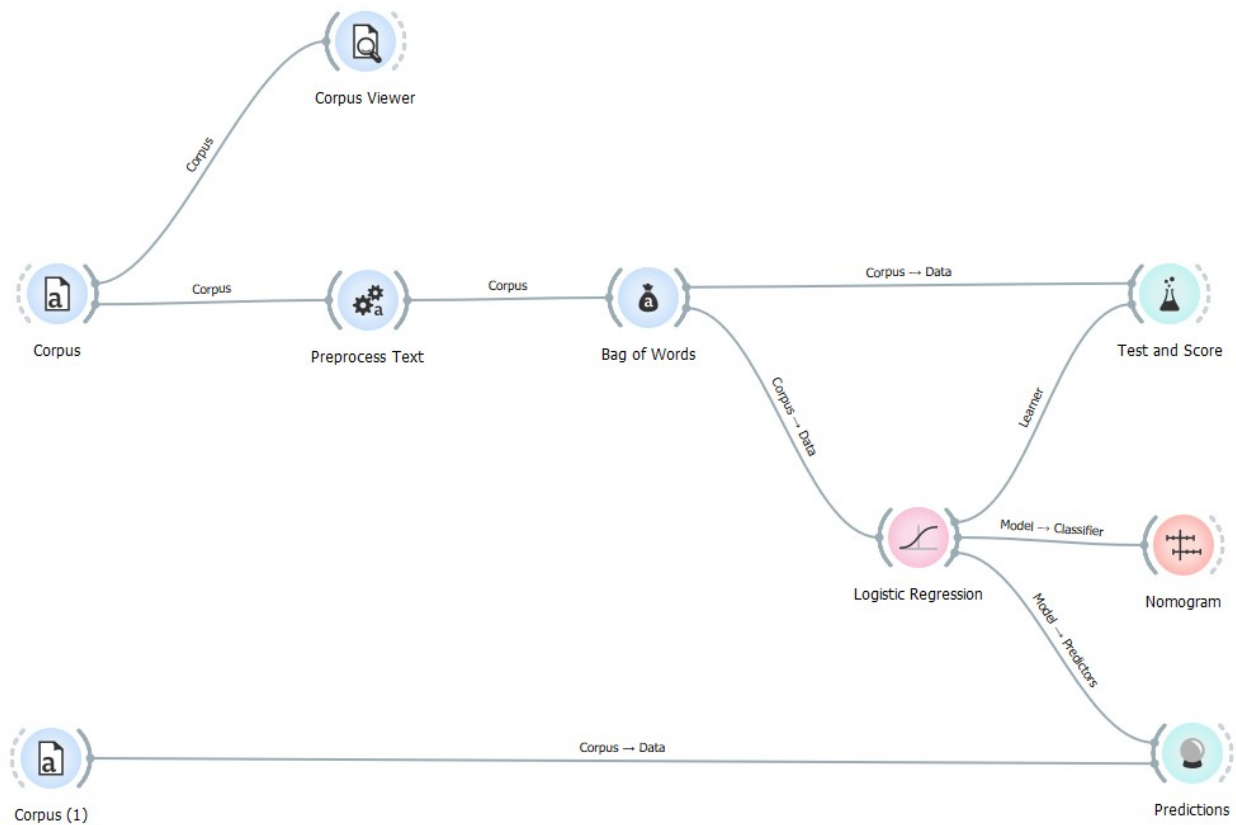
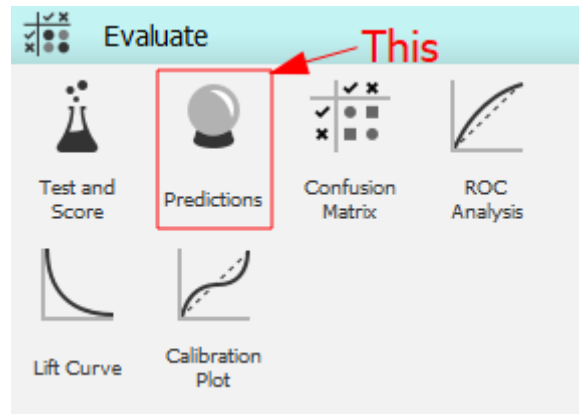
Browse documentation corpora

?

-

3

- Now drag and drop a Predictions widget from the Evaluate section onto the workspace. Connect it to the model as well as the new Corpus widget to visualise the results.



Show probabilities for

Animal Tales

Tales of Magic

Restore Original Order

Logistic Regression

1 0.18 : 0.82 → Tales of Ma...

2 0.00 : 1.00 → Tales of Ma...

3 0.00 : 1.00 → Tales of Ma...

Title

Content

The Little Matc...It was terribly c...

The Philosophie...Far away towar...

The Ugly Duckli...It was lovely su...

?

|

→ 3

|

↩ 3

|

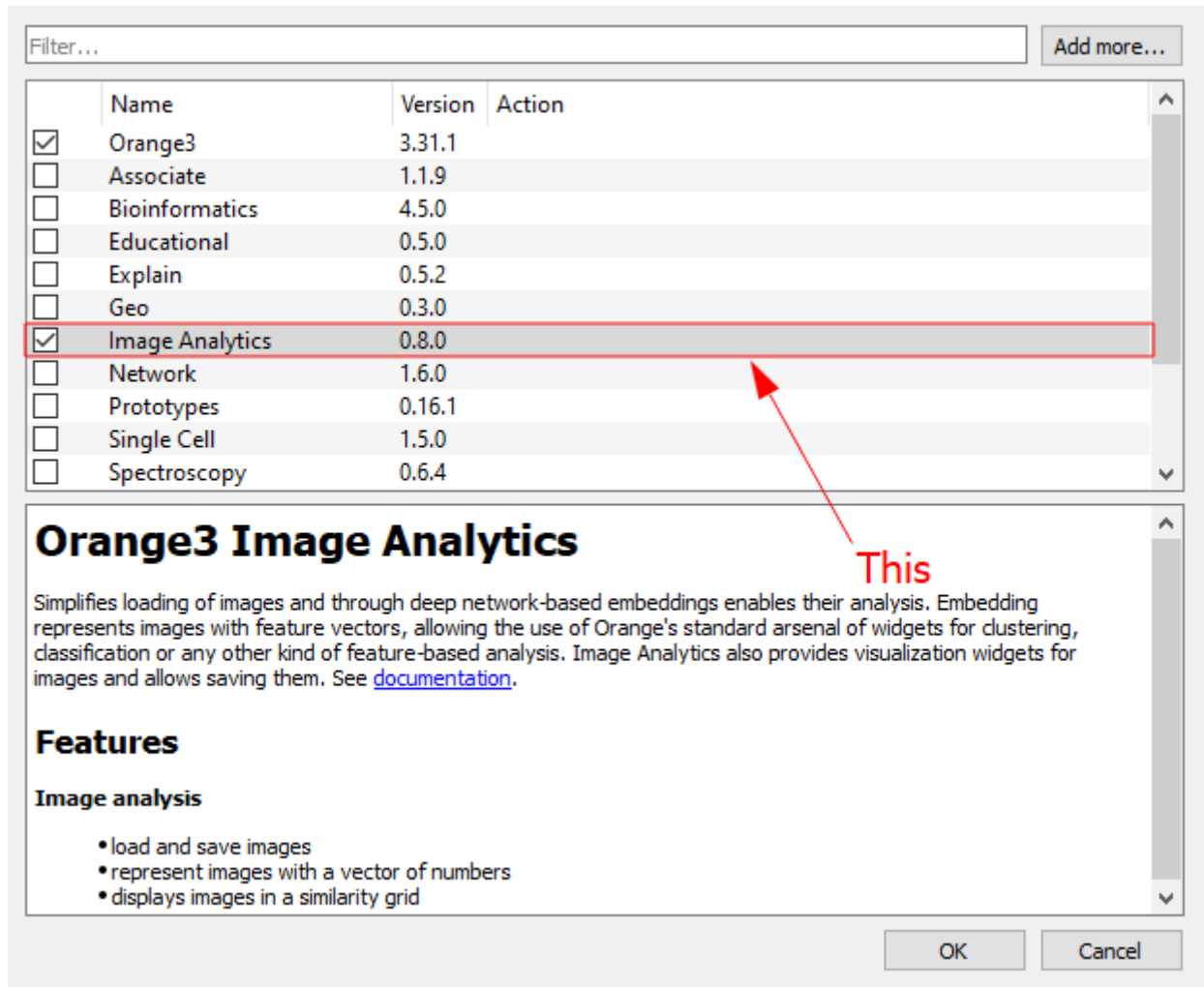
-

Practical 3

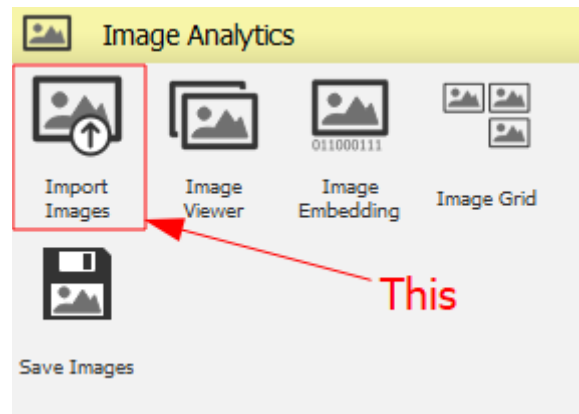
Aim: Image classification using orange.

Procedure:

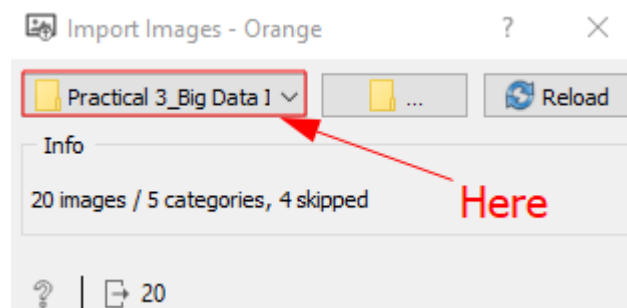
- Before starting, we need to install the **ImageProcessing** add-on. Install it by navigating to **Options > Add ons...**. You will be prompted to restart Orange after the installation completes. If it is already installed, skip this step.



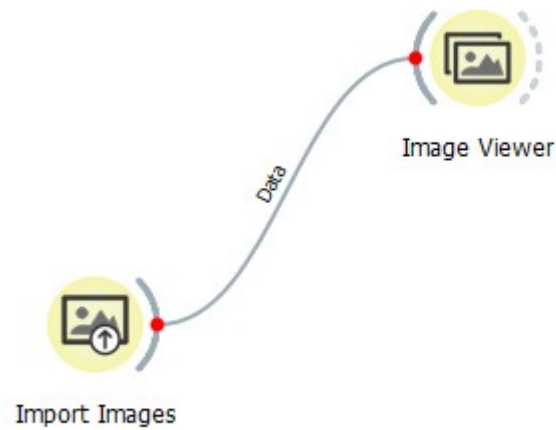
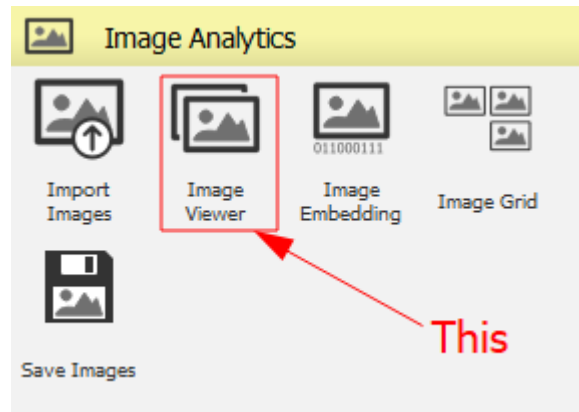
- Add a Import Images widget from the newly added Image Analytics section to the workspace.

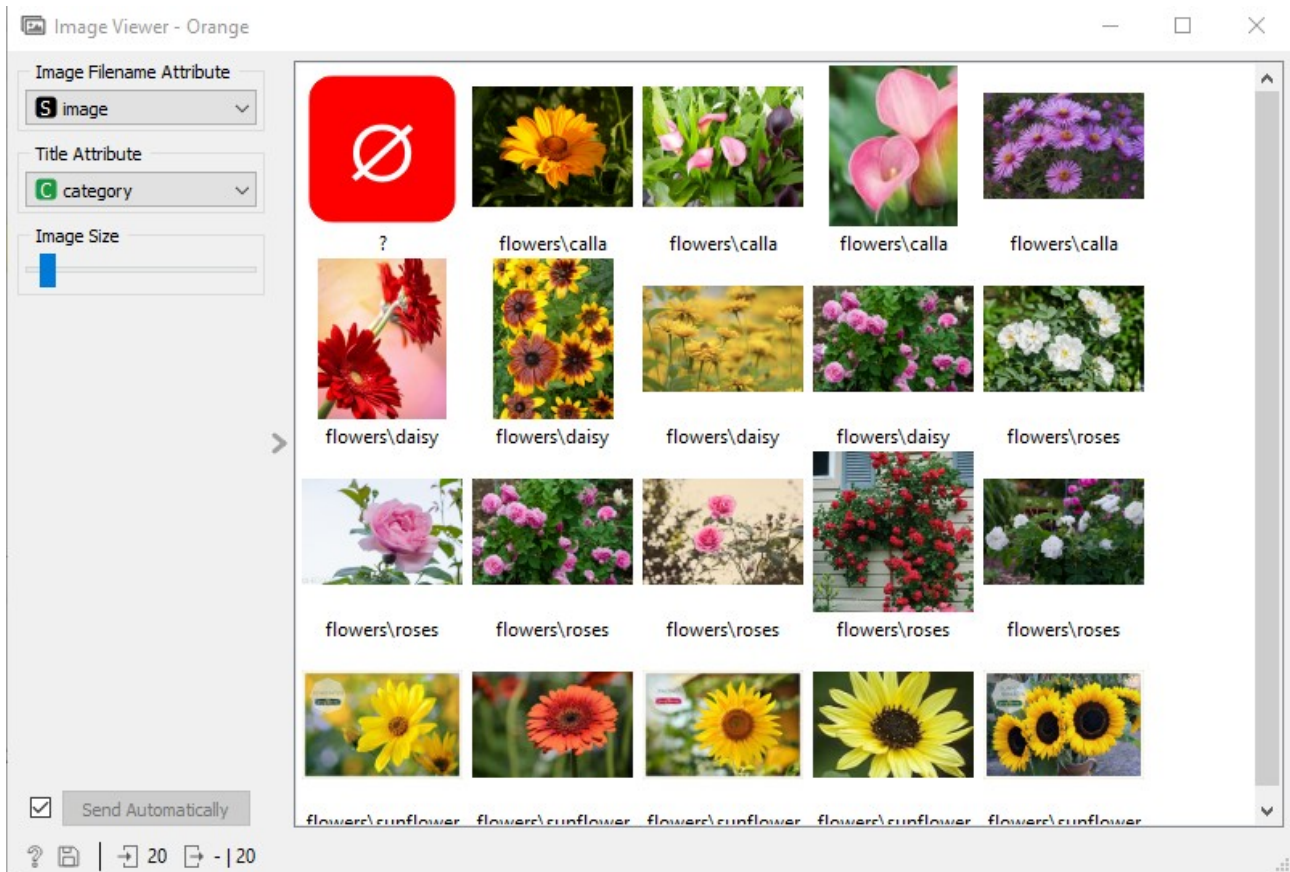


- Load the directory containing the images in the newly added Import Images widget by double clicking the widget.

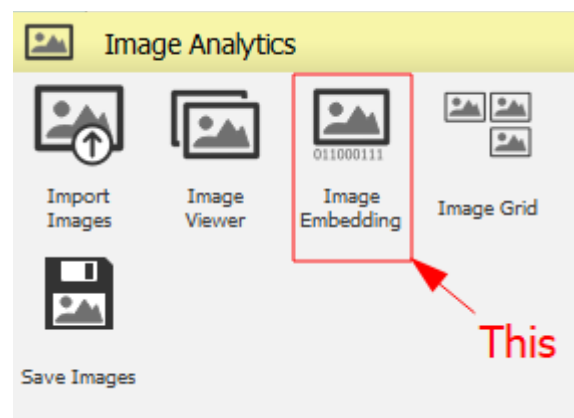


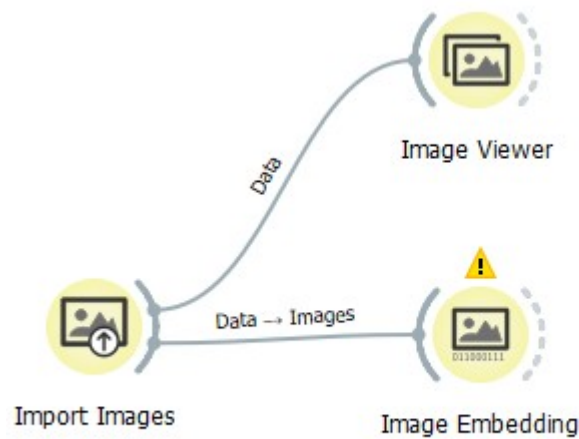
- We can view these images using the Image Viewer widget found in the Image Analytics section. Drag and drop this widget onto the workspace and connect it to the Import Images widget.



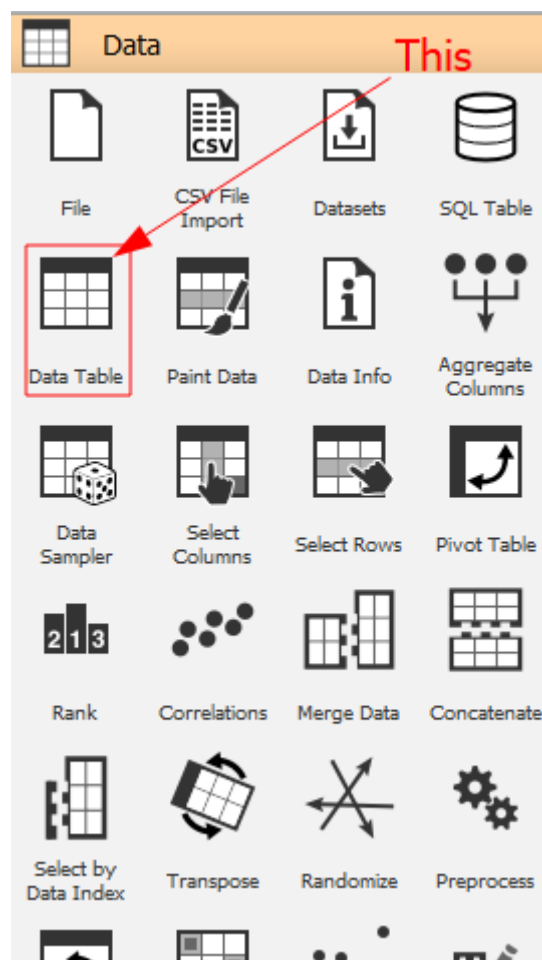


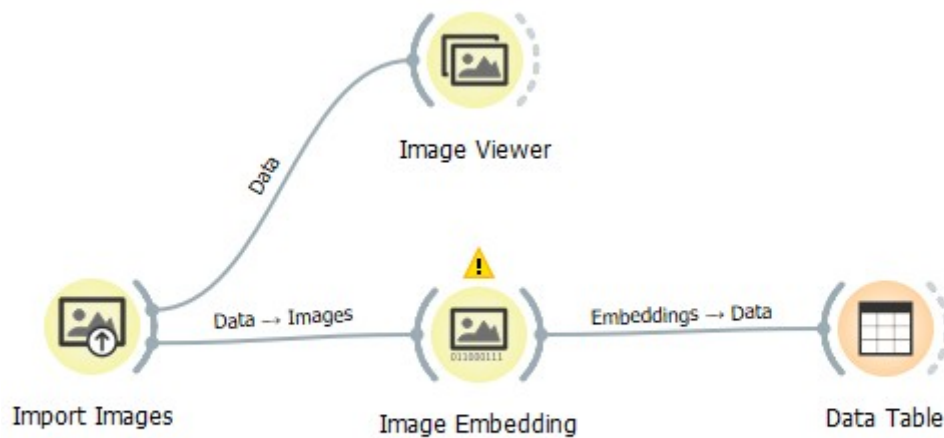
- As our models can only process numbers, we need to convert the images into numerical data. This is where Image Embedding comes into play. Drag and drop a Image Embedding widget from the Image Analytics section onto the workspace and connect it to the Import Images widget.





- We use a Data Table to visualise the tabular data generated by the Image Embedding widget. Drag and drop a Data Table widget from the Data section onto the workspace and connect it to the Image Embedding widget.





Info

18 instances (no missing data)
1000 features
Target with 5 values
5 meta attributes

Variables

☒ Show variable labels (if present)

☐ Visualize numeric values

☒ Color by instance classes

Selection

☒ Select full rows

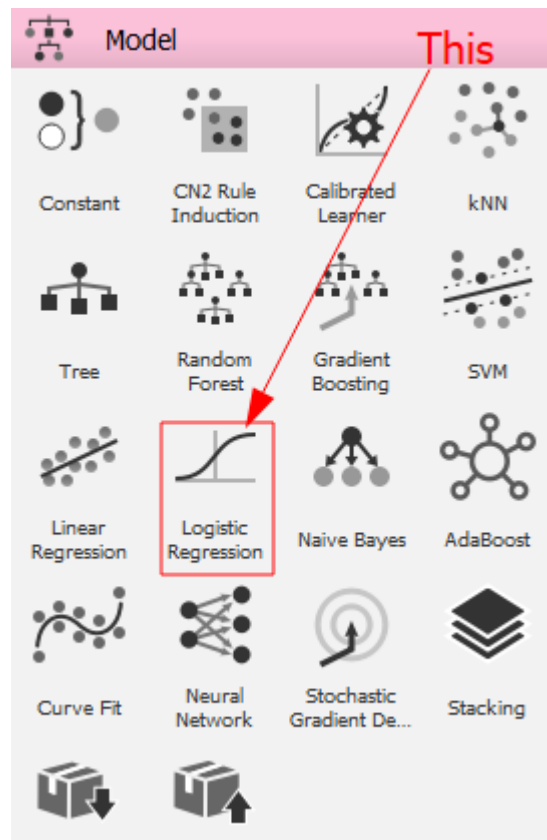
Restore Original Order

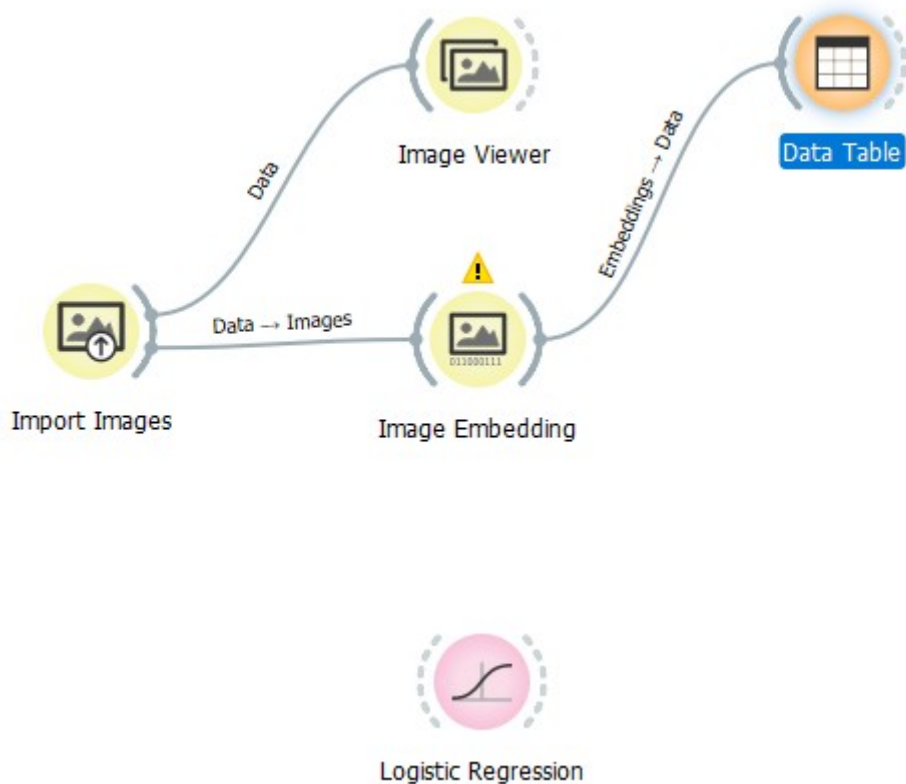
☒ Send Automatically

	category	image name	image	size	width
1	flowers\calla	calla	flowers\calla\c...	3429626	2560
2	flowers\calla	calla1	flowers\calla\c...	29712	400
3	flowers\calla	calla2	flowers\calla\c...	2890	105
4	flowers\daisy	daisy1	flowers\daisy\d...	39660	400
5	flowers\daisy	daisy4	flowers\daisy\d...	71325	375
6	flowers\daisy	daisy6	flowers\daisy\d...	3206304	2084
7	flowers\daisy	daisy8	flowers\daisy\d...	526653	933
8	flowers\roses	rose1	flowers\roses\r...	535368	933
9	flowers\roses	rose2	flowers\roses\r...	354357	933
10	flowers\roses	rose3	flowers\roses\r...	526653	933
11	flowers\roses	rose4	flowers\roses\r...	429411	933
12	flowers\roses	rose6	flowers\roses\r...	841733	700
13	flowers\roses	rose7	flowers\roses\r...	523616	933
14	flowers\sunflo...	sun1	flowers\sunflo...	6649	170
15	flowers\sunflo...	sun2	flowers\sunflo...	5481	176
16	flowers\sunflo...	sun3	flowers\sunflo...	7061	170
17	flowers\sunflo...	sun4	flowers\sunflo...	5808	177
18	flowers\sunflo...	sun5	flowers\sunflo...	222011	800

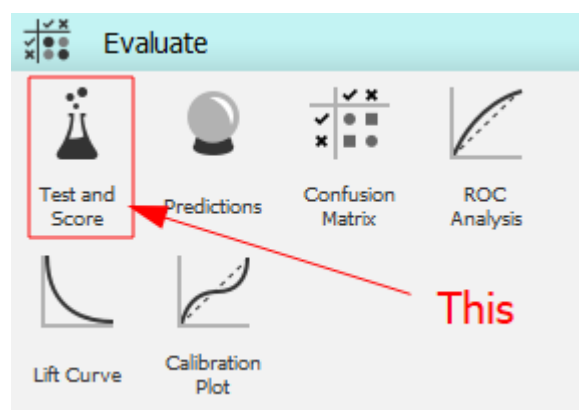
18 | 18

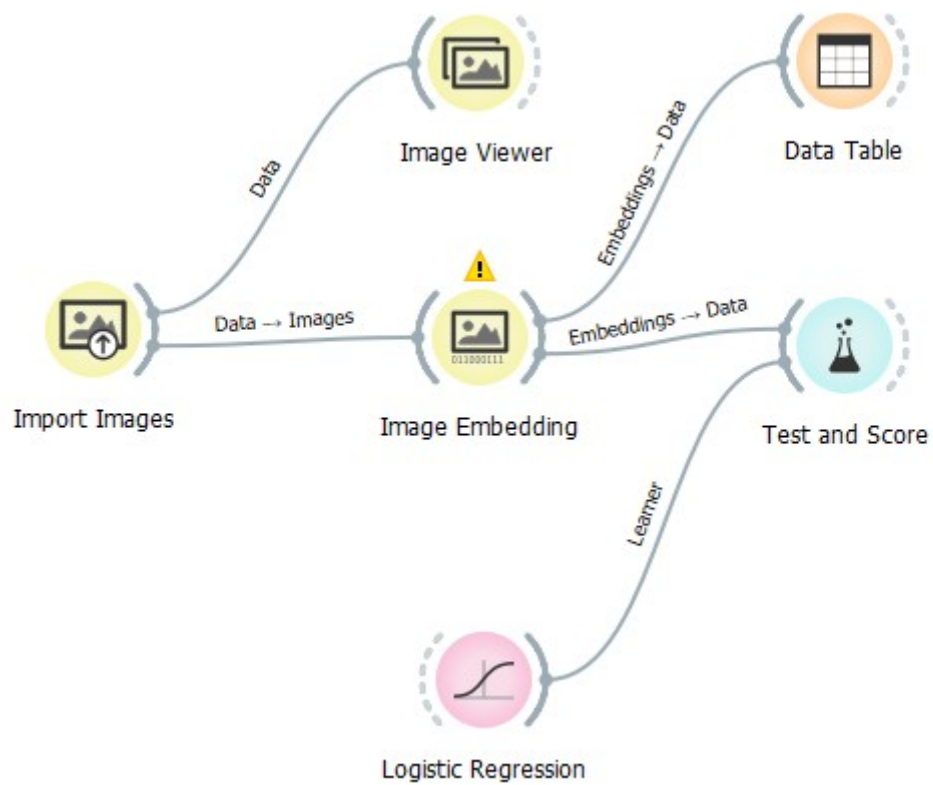
- We will use Logistic Regression as our model. Drag and drop a Logistic Regression widget from the Model section onto the workspace.





- Drag and drop a Test and Score widget from the Evaluate section onto the workspace. This widget will allow us to verify our model. Connect this widget to the Image Embedding and Logistic Regression widgets.





Sampling

☒ Cross validation

Number of folds:

☒ Stratified

☐ Cross validation by feature

☐ Random sampling

Repeat train/test:

Training set size:

☒ Stratified

☐ Leave one out

☐ Test on train data

☐ Test on test data

Target Class

Model Comparison

☐ Negligible difference:

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
Logistic Regression	0.718	0.556	0.519	0.543	0.556

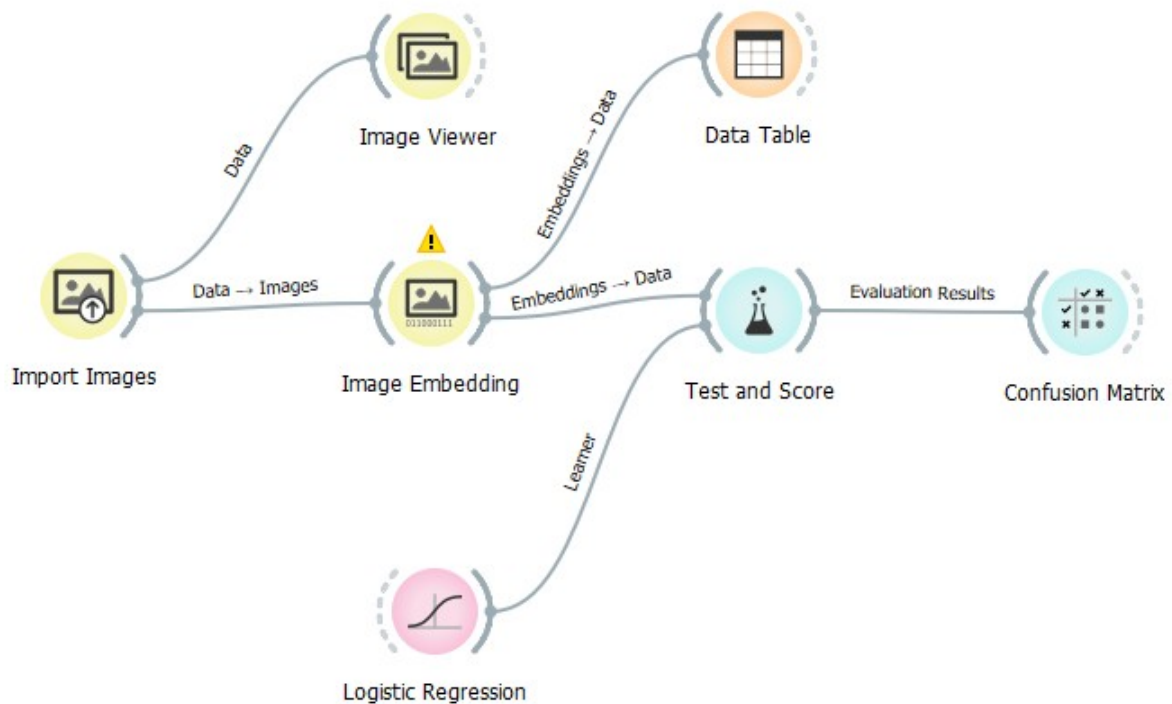
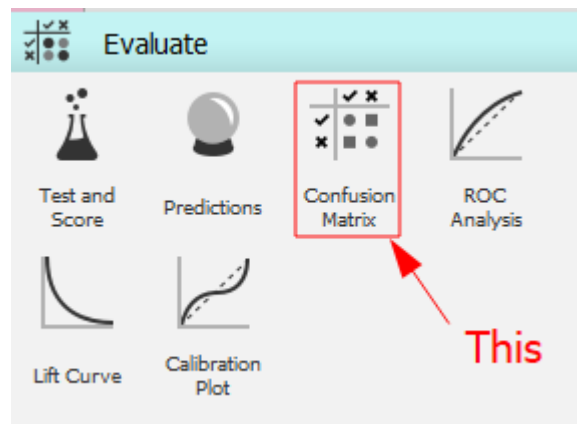
Model Comparison by AUC

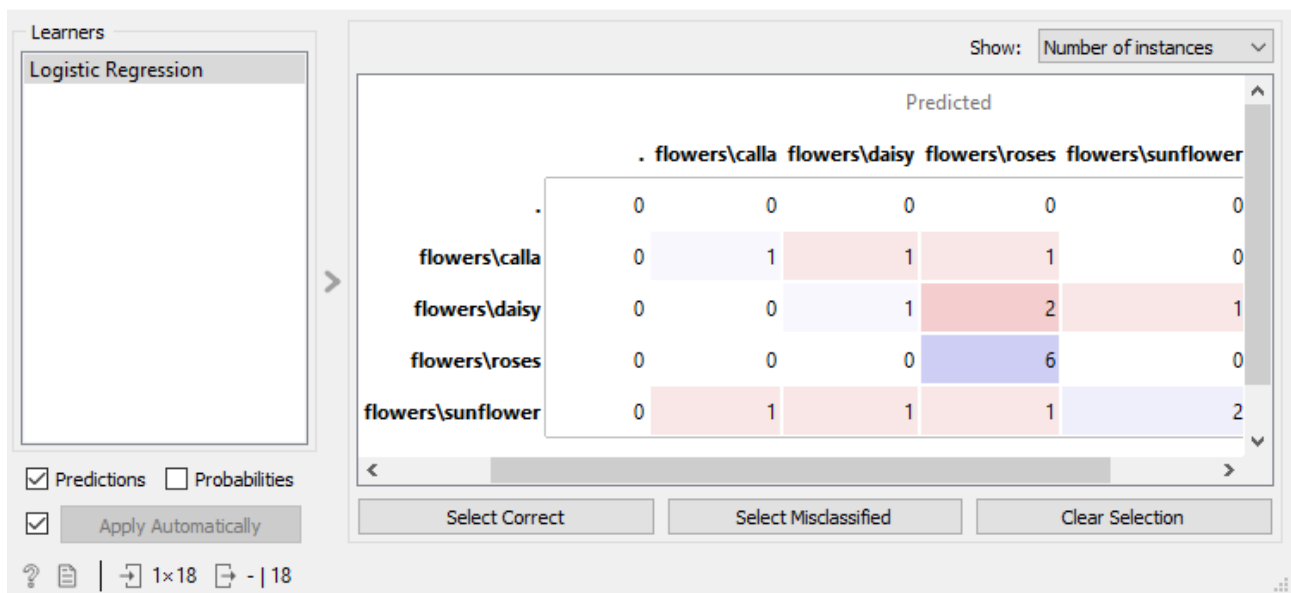
	Logistic ...
Logistic Regression	

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

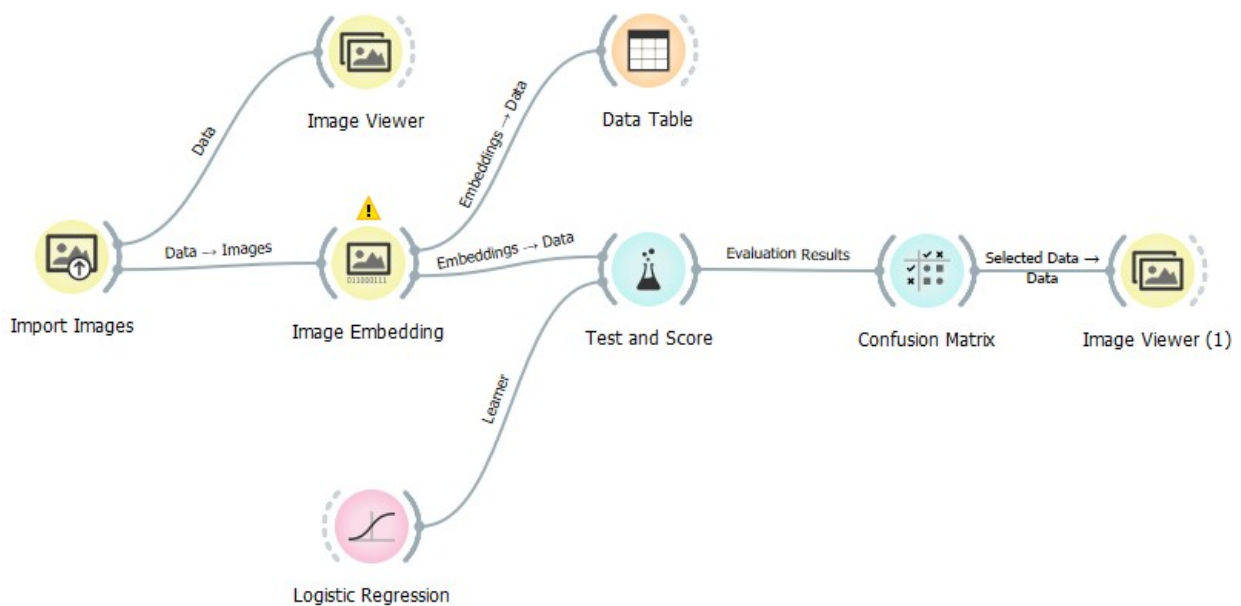
18 | 18 | 1x18

- We now drag and drop a Confusion Matrix widget from the Evaluate section onto the workspace and connect it to the Test and Score widget.





- Drag and drop another Image Viewer widget to view the selected cell(s) from the Confusion Matrix.



Learners

Logistic Regression

Show: Number of instances

Predicted

	flowers\calla	flowers\daisy	flowers\roses	flowers\sunflower
flowers\calla	0	1	1	0
flowers\daisy	0	0	1	1
flowers\roses	0	0	0	6
flowers\sunflower	0	1	1	1

☒ Predictions
 ☐ Probabilities

☒ Apply Automatically

Select Correct
 Select Misclassified
 Clear Selection

? | 1x18 | 6 | 18

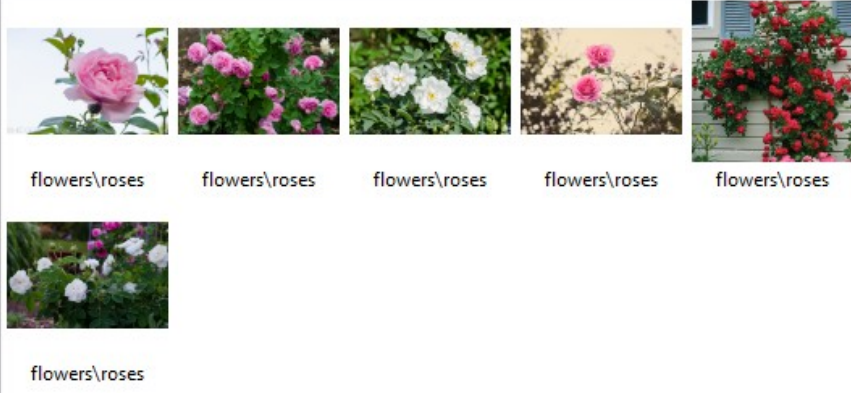
Image Viewer (1) - Orange

Image Filename Attribute: image

Title Attribute: category

Image Size: [Slider]

☒ Send Automatically



flowers\roses
 flowers\roses
 flowers\roses
 flowers\roses
 flowers\roses

flowers\roses

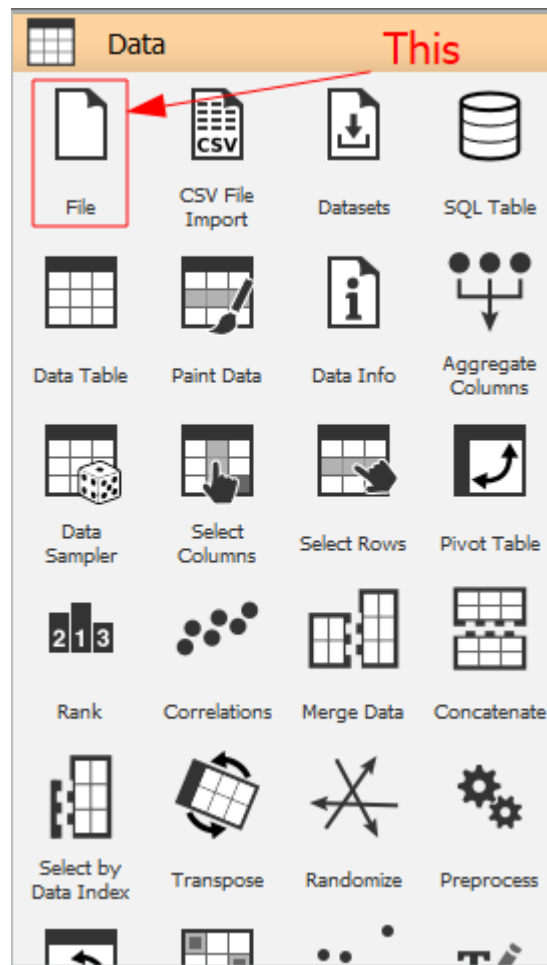
? | 6 | - | 6

Practical 4

Aim: Hierarchical clustering using orange.

Procedure:

- Drag and drop a File widget from the Data section onto the workspace.





➤ Double click the File widget and set the name to **iris.tab**.

Source

☒ File: **iris.tab** ☐ URL:

Info

Iris flower dataset
Classical dataset with 150 instances of Iris setosa, Iris virginica and Iris versicolor.

150 instance(s)
4 feature(s) (no missing values)
Classification; categorical class with 3 values (no missing values)
0 meta attribute(s)

Columns (Double click to edit)

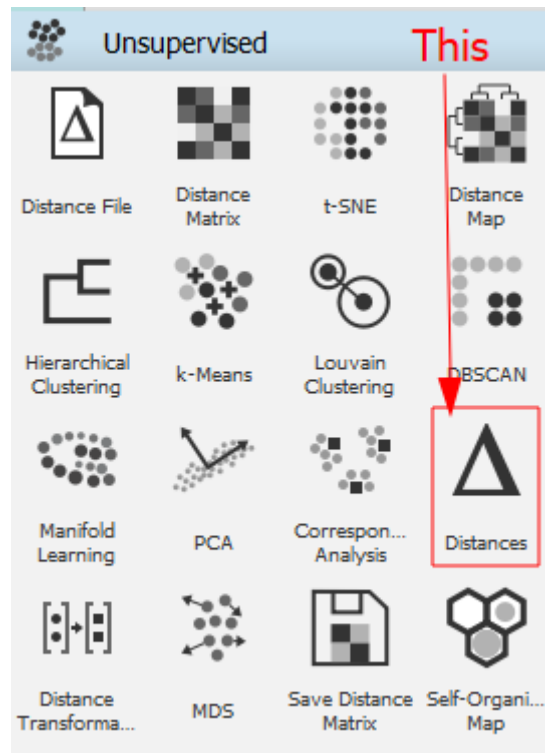
	Name	Type	Role	Values
1	sepal length	N numeric	feature	
2	sepal width	N numeric	feature	
3	petal length	N numeric	feature	
4	petal width	N numeric	feature	
5	iris	C categorical	target	Iris-setosa, Iris-versicolor, Iris-virginica

Reset Apply

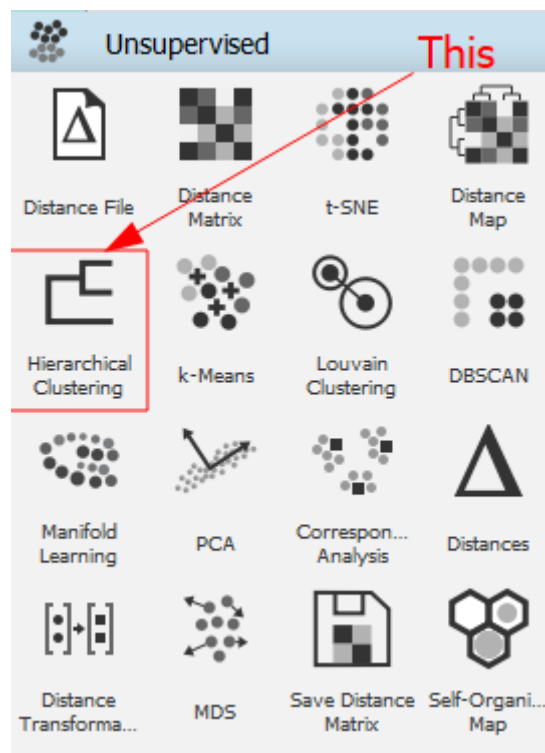
Browse documentation datasets

? | 150

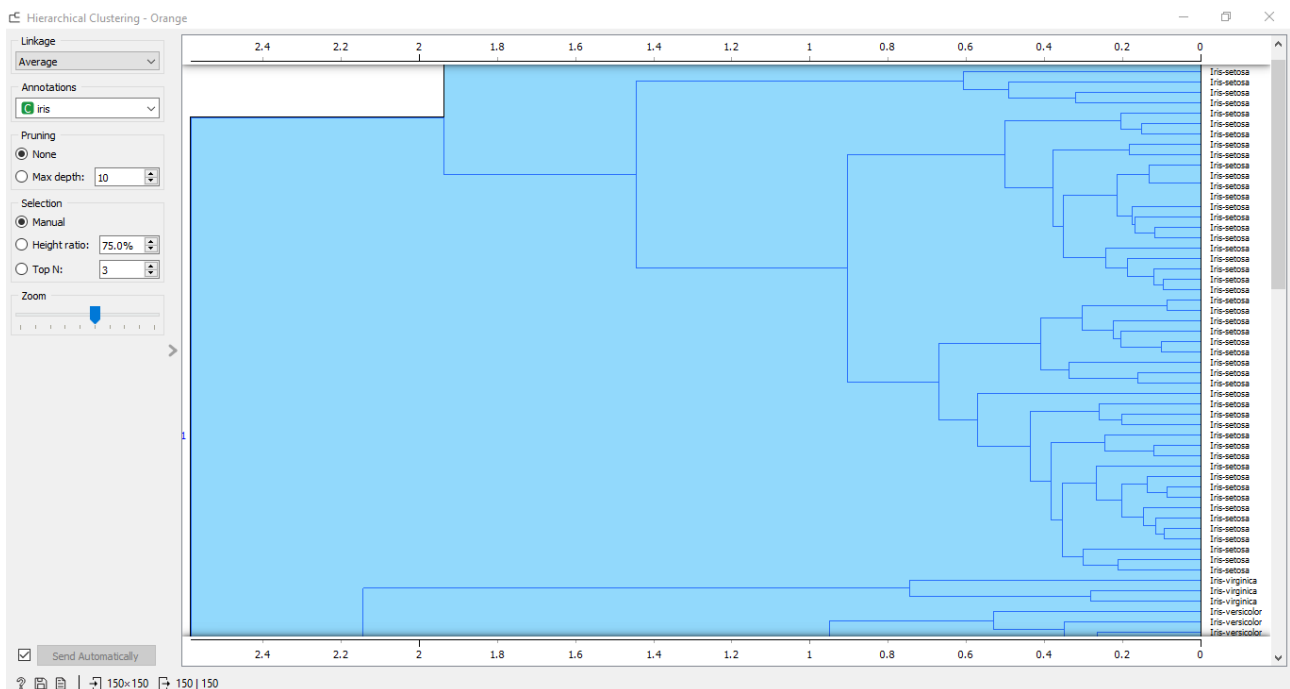
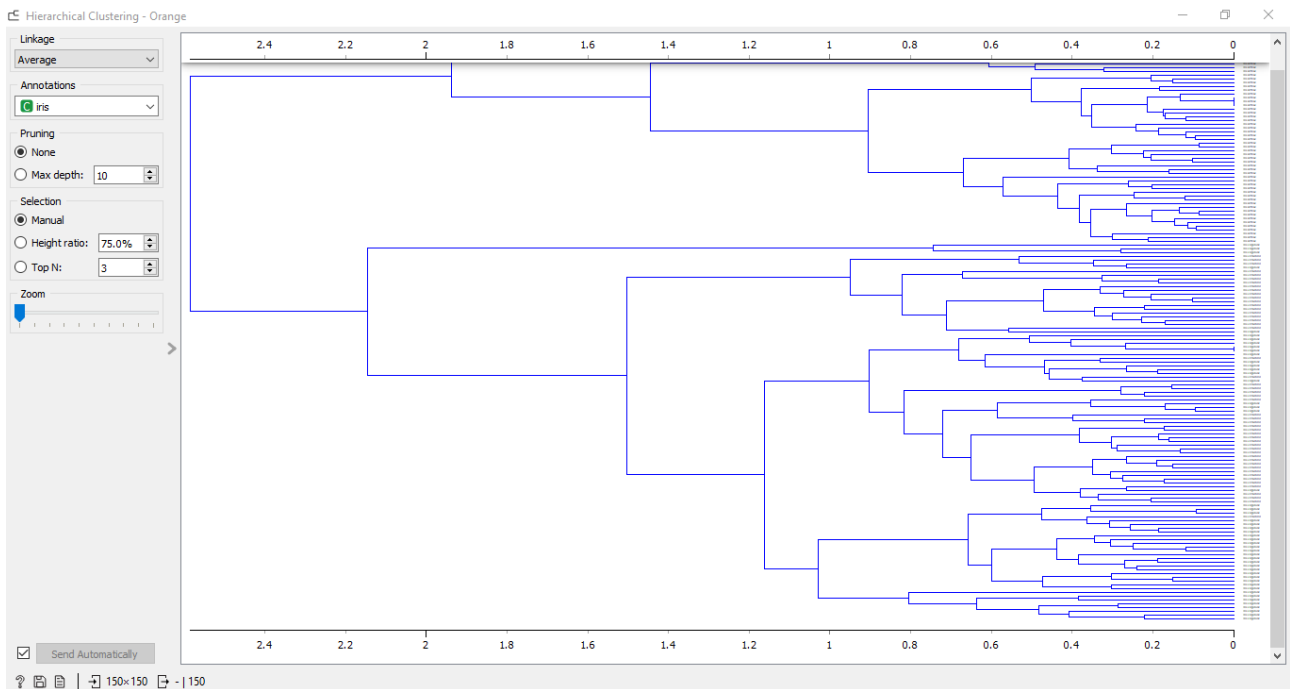
- Drag and drop the Distances widget from the Unsupervised section onto the workspace and connect it to the File widget.



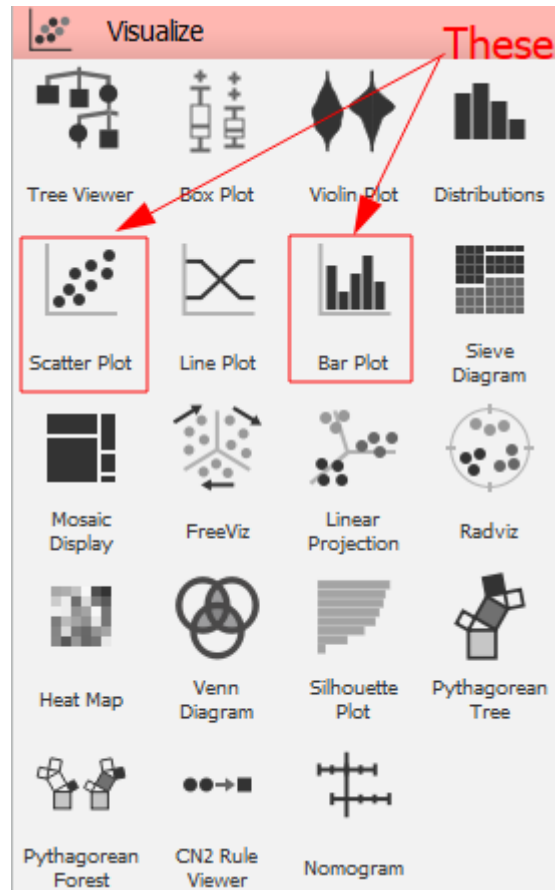
- Drag and drop a Hierarchical Clustering widget from the Unsupervised section to the workspace and connect it to the Distances widget.

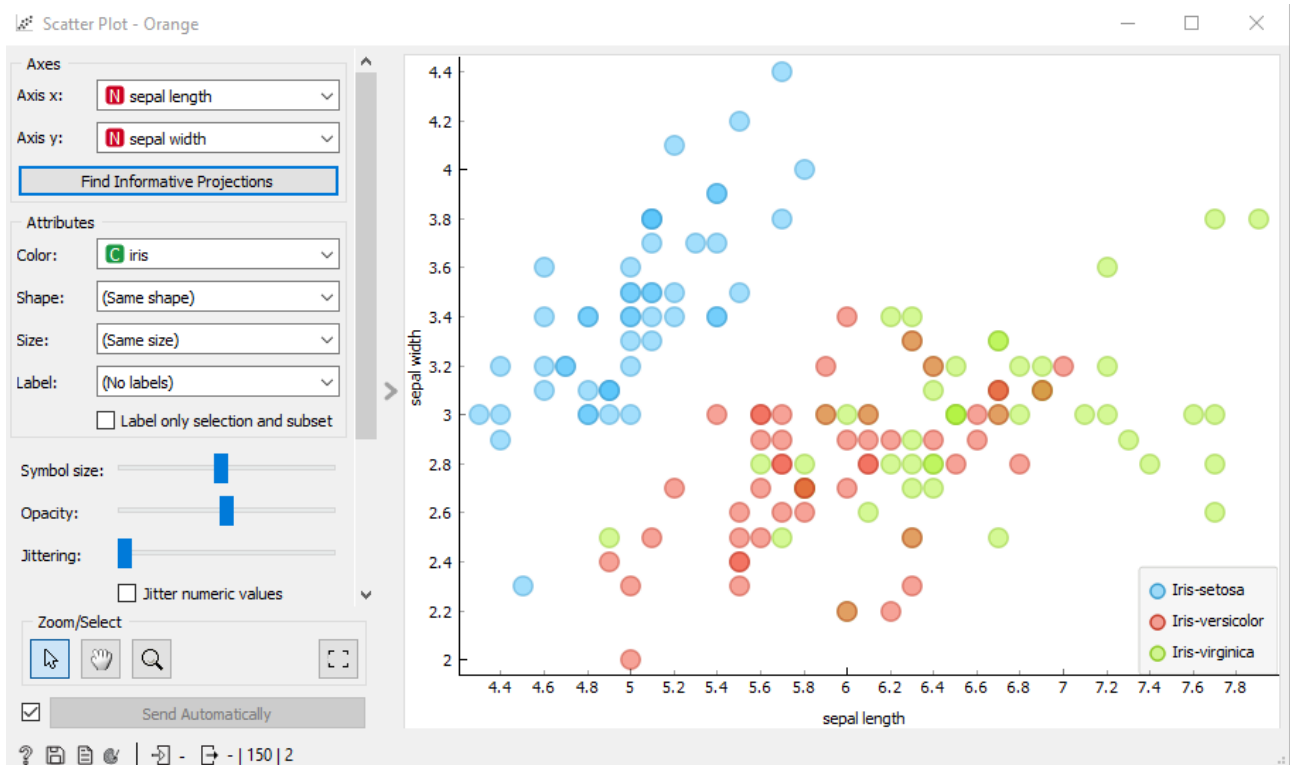
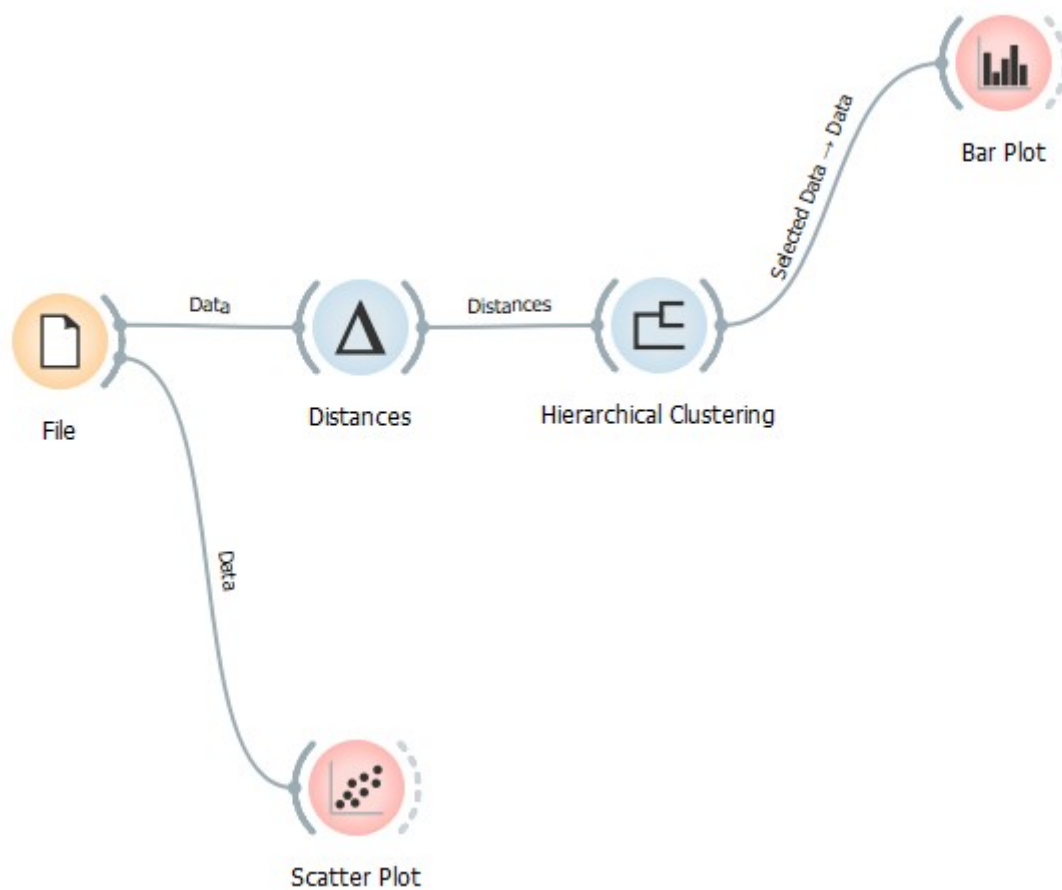


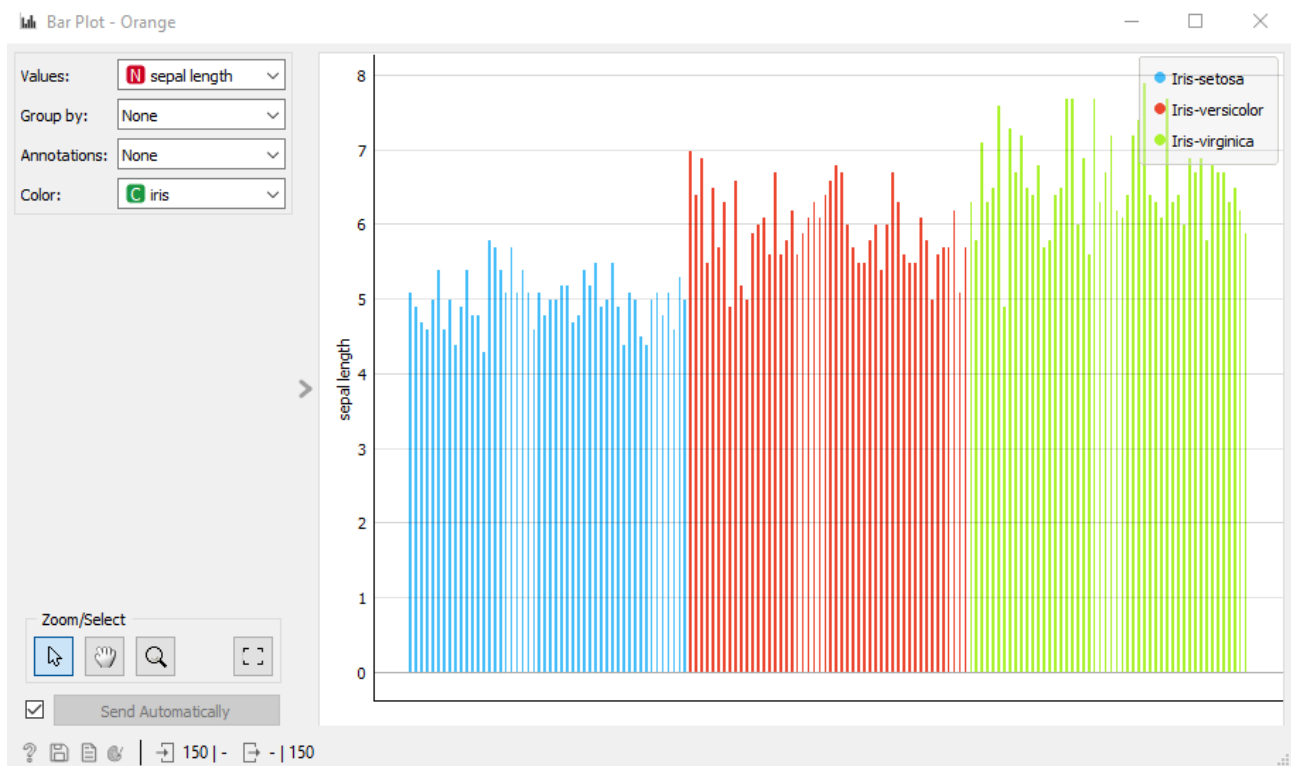
- Double click the Hierarchical Clustering widget to view the dendrogram. Select any sub-cluster as per convenience.



- Choose any visualization method of your choice. This example assumes a Bar Plot for the sub-cluster and a Scatter Plot for the source dataset. You can find both of these widgets in the Visualise section. Connect the Bar Plot to the Hierarchical Clustering widget and Scatter Plot to the File widget.





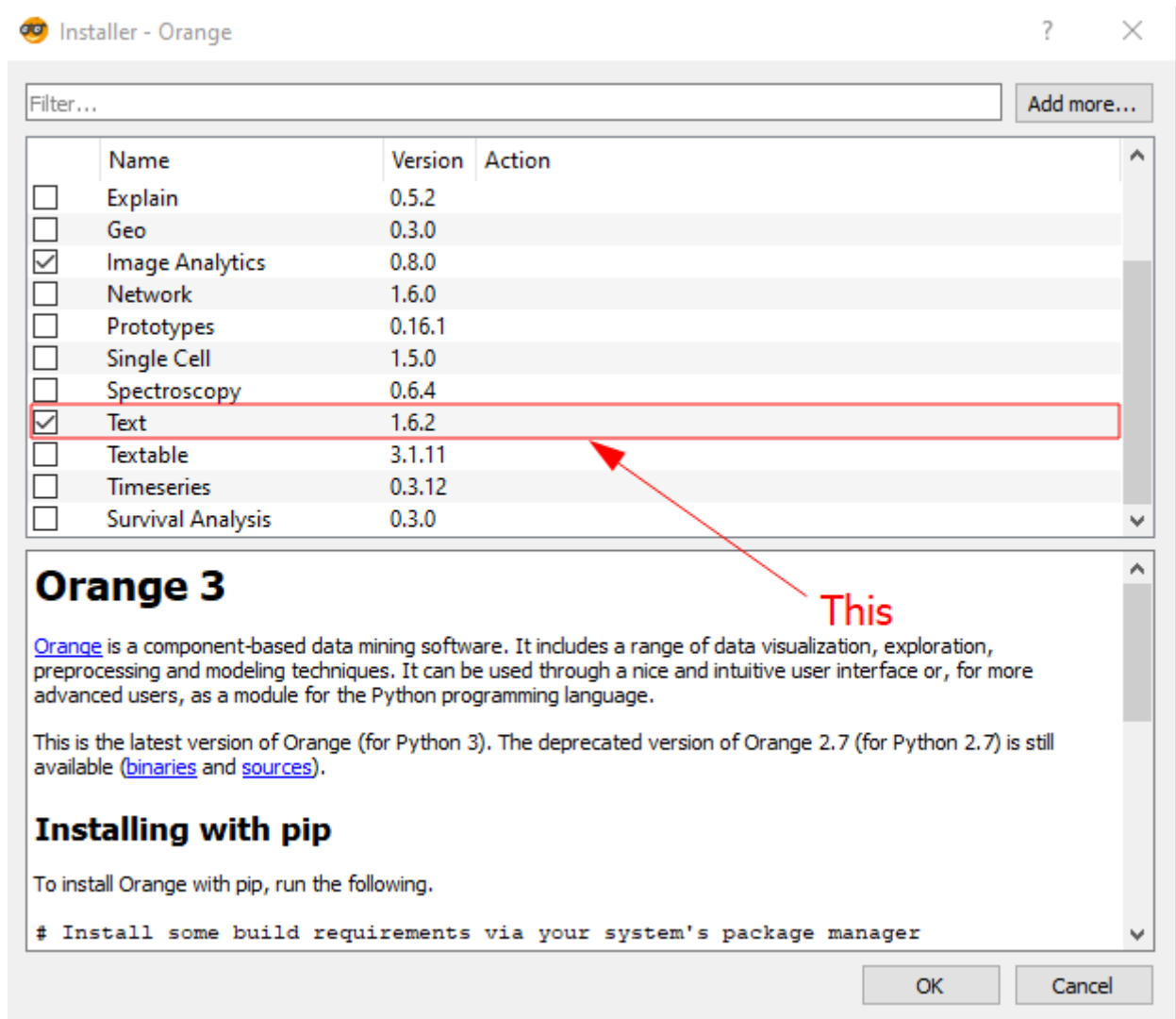


Practical 5

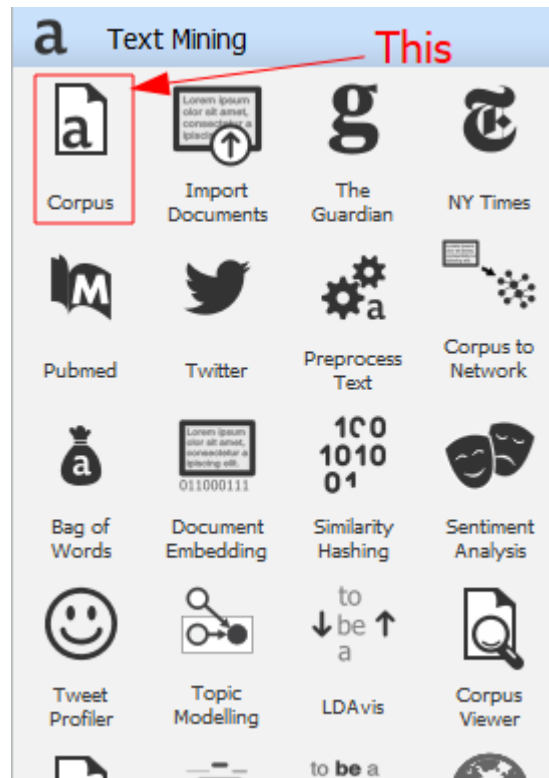
Aim: Hierarchical clustering using orange.

Procedure:

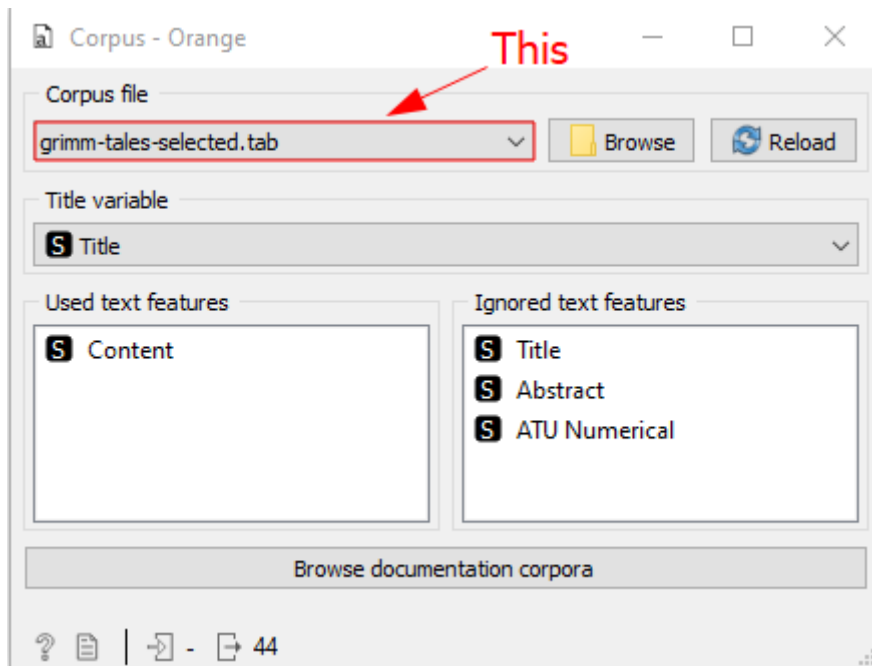
- Before starting, we need to install the **Text** add-on. Install it by navigating to **Options > Add ons...**. You will be prompted to restart Orange after the installation completes. If it is already installed, skip this step.



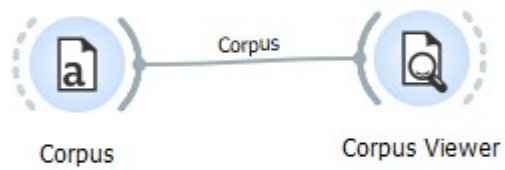
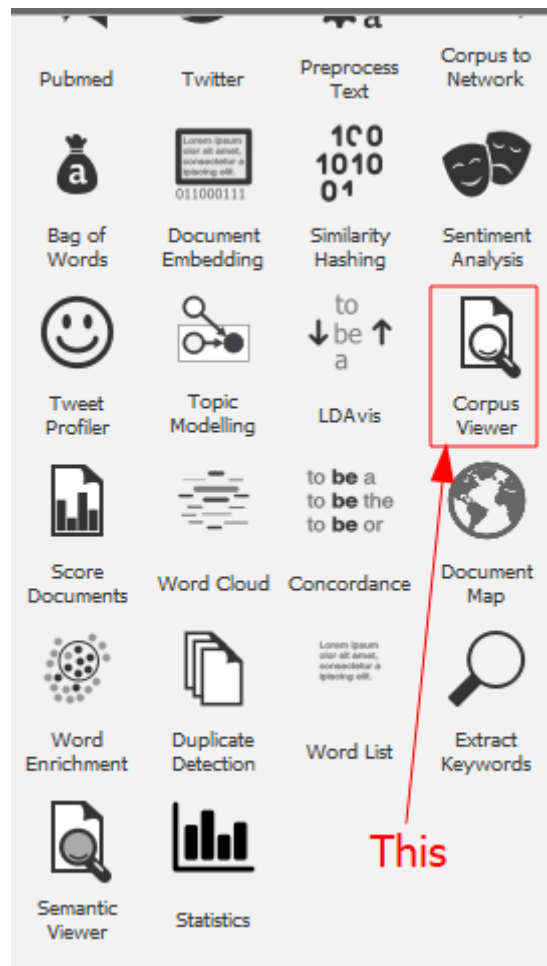
- Drag and drop a Corpus widget to the workspace. It can be found in the newly added Text Mining section.



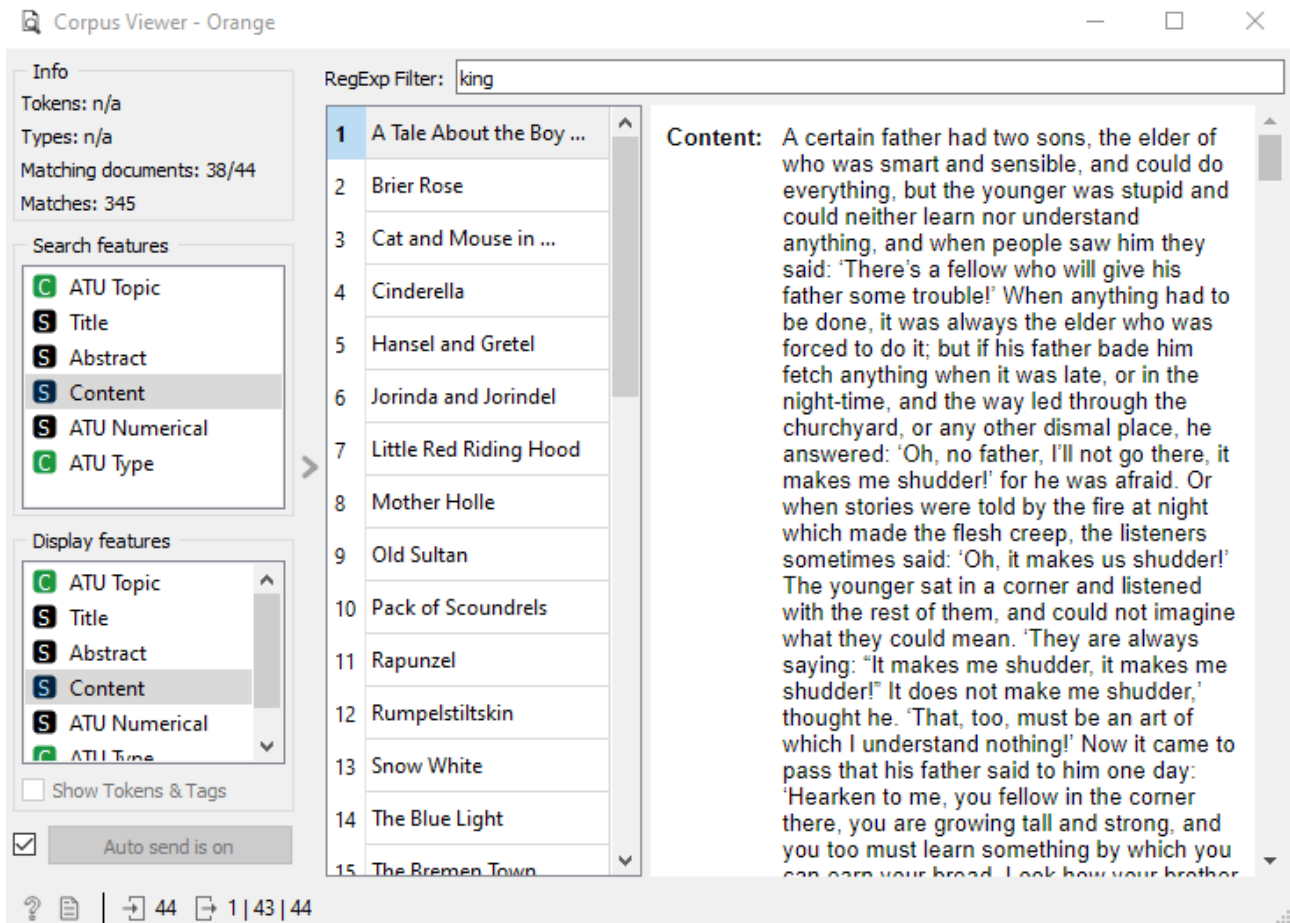
- Double click the Corpus widget and select the `grimm-tales-selected.tab` corpus file.



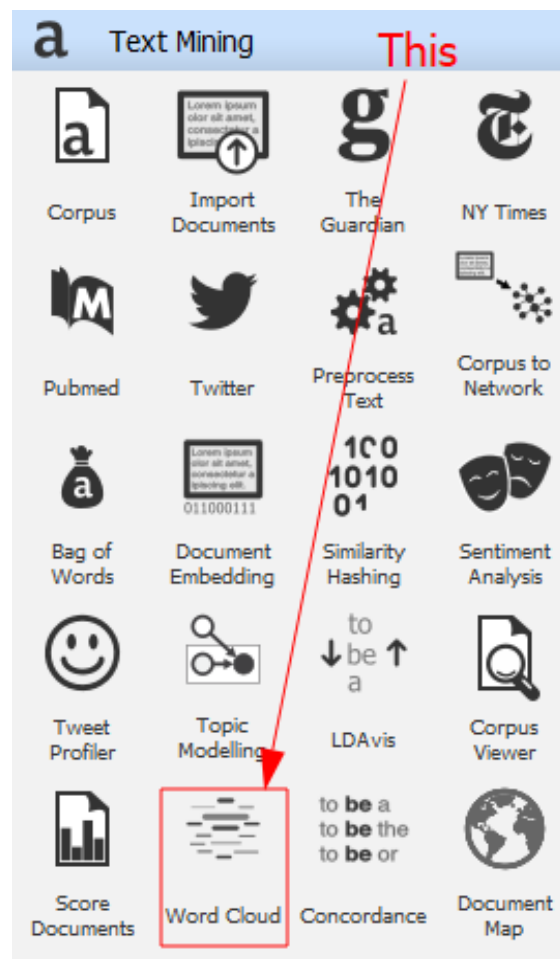
- Drag and drop a Corpus Viewer widget and connect it to the Corpus widget. The Corpus Viewer widget can be found in the Text Mining section.

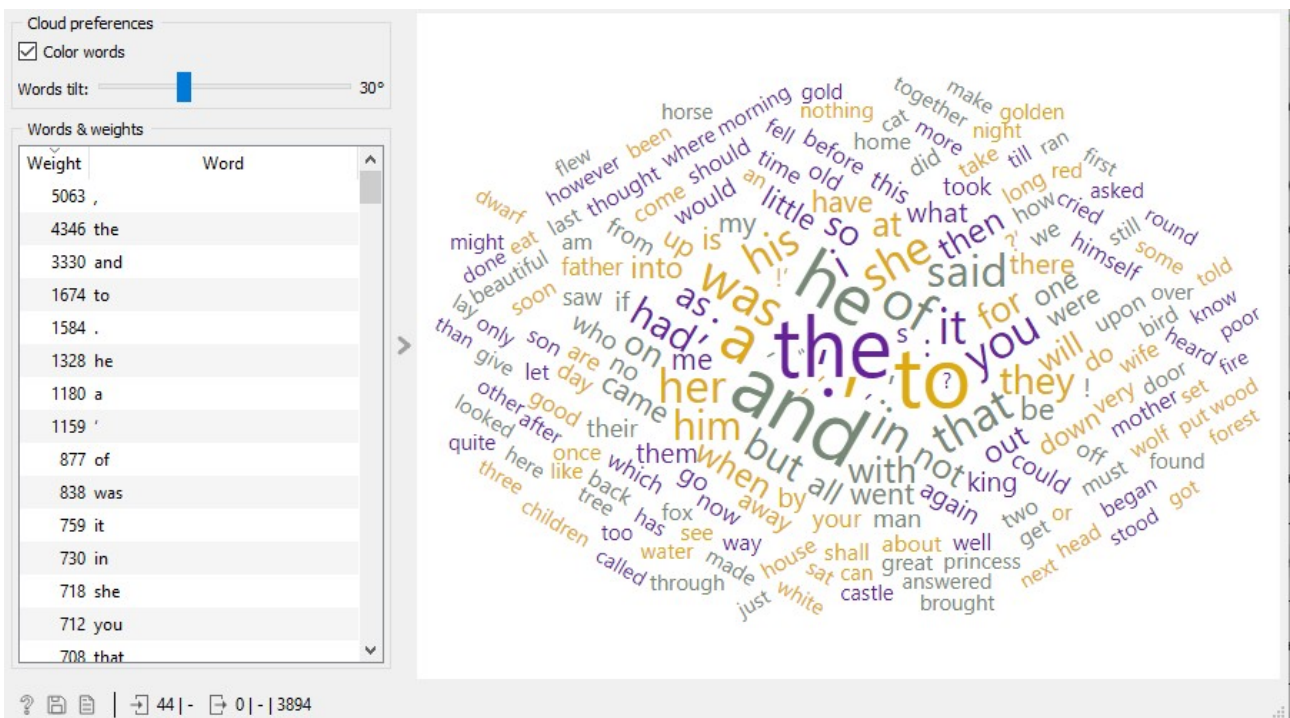
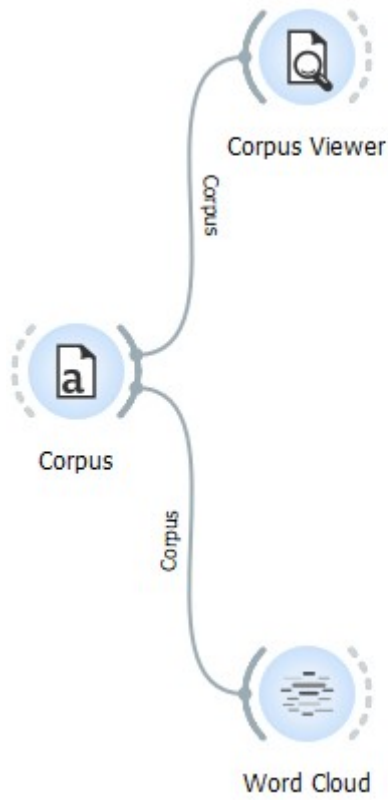


- Double-clicking the Corpus Viewer allows us to peek through the corpus and also allows us to filter text.

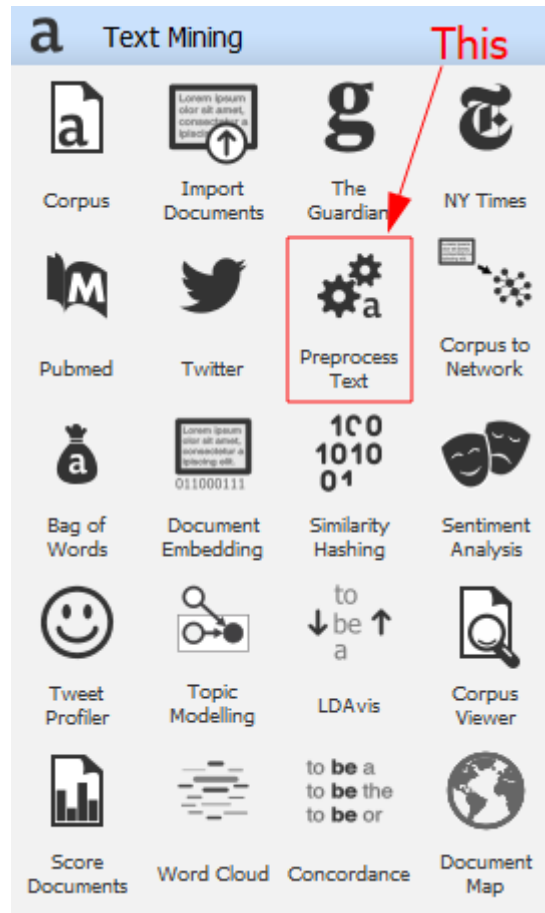


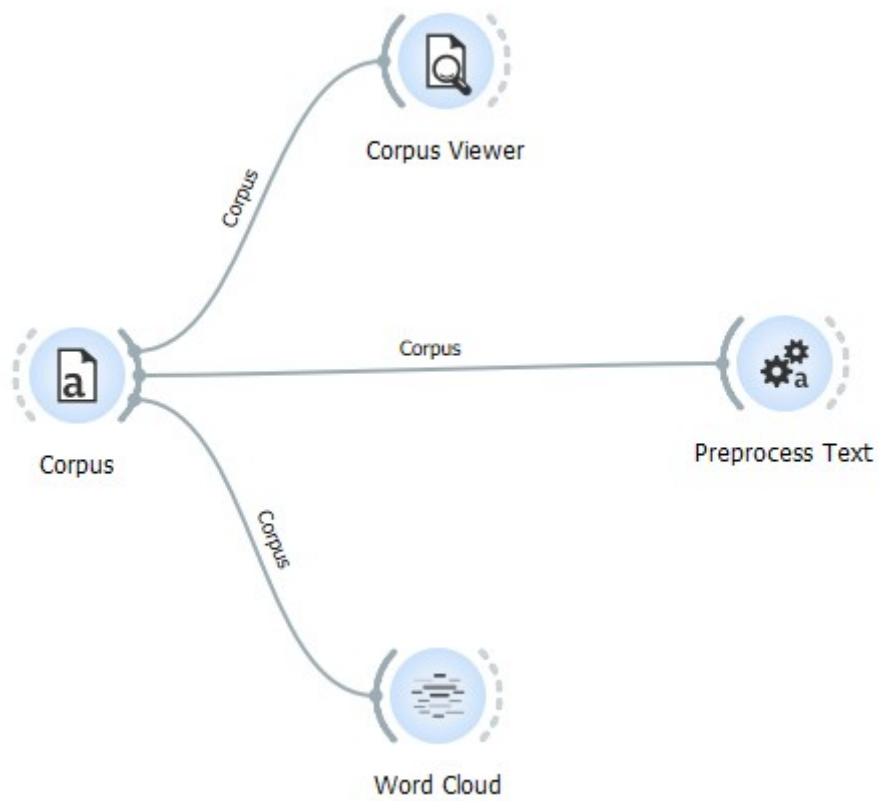
- Another method of visualizing data is through Word Clouds. Drag and drop a Word Cloud widget from the Text Mining section and connect it to the Corpus widget.





- We see that a lot of punctuation marks and uninformative words have made their way into the Word Cloud. In order to eliminate this, we have to use the Preprocess Text widget. Drag and drop a Preprocess Text widget from the Text Mining section onto the workspace and connect it to the Corpus widget.





- We now add another Word Cloud widget to see our updated cloud.

